

Heterologous protein production using euchromatin-containing expression vectors in mammalian cells

Katalin Zboray^{1,†}, Wolfgang Sommeregger^{2,3,†}, Edith Bogner¹, Andreas Gili³, Thomas Sterovsky³, Katharina Fauland³, Beatrice Grabner¹, Patricia Stiedl¹, Herwig P. Moll⁴, Anton Bauer⁵, Renate Kunert^{2,*} and Emilio Casanova^{1,4,*}

¹Ludwig Boltzmann Institute for Cancer Research (LBI-CR), Vienna, 1090, Austria, ²Vienna Institute of BioTechnology, Department of Biotechnology, University of Natural Resources and Life Sciences, Vienna, 1190, Austria, ³Polymun Scientific GmbH, Klosterneuburg, 3400, Austria, ⁴Institute of Pharmacology, Center of Physiology and Pharmacology, Comprehensive Cancer Center, Medical University of Vienna, Vienna, 1090, Austria and ⁵The Antibody Lab, Vienna, 1210, Austria

Received January 09, 2015; Revised April 27, 2015; Accepted April 29, 2015

ABSTRACT

Upon stable cell line generation, chromosomal integration site of the vector DNA has a major impact on transgene expression. Here we apply an active gene environment, rather than specified genetic elements, in expression vectors used for random integration. We generated a set of Bacterial Artificial Chromosome (BAC) vectors with different open chromatin regions, promoters and gene regulatory elements and tested their impact on recombinant protein expression in CHO cells. We identified the *Rosa26* BAC as the most efficient vector backbone showing a nine-fold increase in both polyclonal and clonal production of the human IgG-Fc. Clonal protein production was directly proportional to integrated vector copy numbers and remained stable during 10 weeks without selection pressure. Finally, we demonstrated the advantages of BAC-based vectors by producing two additional proteins, HIV-1 glycoprotein CN54gp140 and HIV-1 neutralizing PG9 antibody, in bioreactors and shake flasks reaching a production yield of 1 g/l.

INTRODUCTION

Recombinant protein production in mammalian cells is the predominant way of nowadays biologic drug production (1,2). Due to the ever increasing demand for protein therapeutics, new methods decreasing the time and costs in the generation of high-yield producer cell lines are extremely demanded. Regarding time considerations, large scale tran-

sient gene expression (TGE) (3–5) is the most rapid way to produce multiple recombinant proteins in tens to hundreds milligram quantities, typically from 1 to 10 l culture volumes. TGE is an attractive method to use mainly for pre-clinical studies and basic research. Nevertheless, for large scale production of the therapeutic proteins stable cell line generation is still the method of choice. Overall protein yield of a producer cell line is highly affected by its health status, longevity and metabolism (6). The recent sequencing of the CHO-K1 genome (7) combined with ‘Omics’ based systematic approaches (i.e. transcriptomics, proteomics and metabolomics) will result in a better understanding of these processes (Reviewed in (8–10)). Results of these studies will help to engineer CHO cells with improved culture characteristics, increased lifespan and production (11). Another aspect strongly impacting protein yield and stability is the nature of the expression vector used to generate producer cell lines (11). During the generation of stable cell lines, plasmid DNA integrates in a random manner into the host cell’s genome. Since the site of integration has a major impact on transcriptional activity of the incoming construct (so-called ‘chromatin positional effects’ (12)), expression levels can be low and unpredictable resulting in a high variability between individual clones. Consequently, there is a need to screen very high numbers of clones to identify efficient and stable producers (13). Moreover, high-yield producer cell line generation via transgene amplification using selection systems such as the DHFR or GS system (14) is time-consuming, induces genomic instability and may lead to silencing of transgene expression (15). To circumvent chromatin positional effects, expression vectors can be flanked by cis-regulatory elements which reduce the positional ef-

*To whom correspondence should be addressed. Tel: +43 1 40160 71210; Fax: +43 1 40160 931300; Email: emilio.casanova@lbicr.lbg.ac.at
Correspondence may also be addressed to Renate Kunert. Tel: +43 1 476546595; Fax: +43 1 476546242; Email: rene.kunert@boku.ac.at

†These authors contributed equally to the paper as first authors.

fects and allow stable expression of the transgene. Indeed, ubiquitous chromatin opening elements (UCOEs) (16–18), scaffold/matrix attachment regions (S/MARs) (19,20) and antirepressor (21) elements have been reported to have a beneficial effect on protein expression levels and stability. Alternatively, it is possible to insert the gene of interest (GOI) into a pre-defined locus in the host cell by using recombinase mediated cassette exchange (RMCE) methods (22). A well-chosen locus or so-called ‘hot-spot’ containing euchromatin can insure long-term stable protein production levels. However, only a single copy is integrated using this method which may limit the maximum achievable expression levels, although high antibody yields (up to 1 g/l) have been reported using the RMCE technology (23). In contrast to the previously described strategies, we aimed to create a system where we can combine the advantages of targeted integration in a hot-spot and the flexibility of random integration methods. We reasoned that large expression vectors harboring whole loci containing euchromatin (hot-spots) will not be affected by positional effects and will confer high and stable expression levels. To this end, we explored Bacterial Artificial Chromosomes (BACs) as expression vectors for recombinant protein production in CHO cells. BACs have a large cloning capacity (200–300 kilobase (kb)) and therefore they can accommodate an entire locus with most if not all of the elements that control the expression of a gene. Indeed, BACs have been widely used in the mouse transgenic field because they ensure positional effect independent and copy number dependent expression of a transgene (24–26). According to this, BAC vectors should be ideal tools applied to heterologous protein production in mammalian cells. BAC-based expression vectors containing carefully chosen loci should combine the beneficial effects of a stable genetic environment with the possibility to integrate several vector copies in the cell host, thus boosting the transgene expression and making transgene amplification unnecessary.

In this study we aimed to optimize the use of BAC-based vectors for recombinant protein production in CHO cells. To do this, we first generated a series of BAC-based vectors combining different open chromatin loci, promoters and other gene regulatory elements and tested them with an ‘easy to express’ single polypeptide chain protein: the Fc fragment of human IgG1 (IgG-Fc). Thereafter, we challenged our most efficient constructs and demonstrated the usefulness of BAC-based expression vectors producing two additional complex proteins: CN54gp140, a difficult to produce, heavily glycosylated single polypeptide derived from HIV-1 and PG9, a broadly neutralizing anti-HIV-1 antibody, as a complex protein model assembled from two different polypeptide chains. By using these three different protein models, we could demonstrate that BAC-based expression vectors dramatically reduce the time needed to generate high yield and stable cell lines.

MATERIALS AND METHODS

Generation of conventional and BAC-based expression vectors

All plasmids were constructed in pBluescript KS vector (Stratagene) by changing the multi cloning site to

a new polylinker region with unique restriction enzyme sites. Restriction enzymes and T4 ligase were purchased from Thermo Fisher Scientific. Cloning was performed in DH10B *E.coli* bacteria. All plasmid sequence information is available upon request.

BAC recombineering. BAC recombineering was performed as previously described (27,28). DH10B *E.coli* harboring the corresponding BACs were electroporated with the temperature sensitive pSC101-BAD-gbaA plasmid which carries the recombinase proteins necessary for homologous recombination. Cells harboring the pSC101-BAD-gbaA plasmid were selected in tetracycline (5 µg/ml) at 30°C overnight. Bacterial cells derived from one single positive colony were cultured overnight at 30°C and transferred to 50 ml of fresh medium next day. At an optical density (OD₆₀₀) of 0.2, the expression of the recombinogenic proteins was induced by the addition of L-arabinose (to a 0.3%–0.4% final concentration) and by shifting the temperature to 37°C. After one additional hour, cells were harvested and electro-competent cells were prepared by a double wash with ice cold distilled water. Cells were resuspended in 10% ice cold glycerol solution and aliquoted to Eppendorf tubes (100 µl) prior snap freezing in liquid N₂ or directly used for electroporation. 50–100 ng of linearized plasmid DNA carrying the insert for BAC modification (Figures 1b and 4a plasmid-controls) flanked by 120–200 base pair (bp) homology regions (HR) to the respective BAC was electroporated to these cells with a Gene Pulser Xcell (BioRad) electroporator. Cells were incubated for 70 min at 37°C without antibiotics in LB medium and plated on LB-agar containing 15 µg/ml kanamycin at 37°C overnight. BAC DNA was isolated from resistant colonies and then analyzed for correct recombination by polymerase chain reaction (PCR) using primer pairs within the incoming construct and the BAC. Original BAC clones were obtained from the BAC PAC Resources Children’s Hospital Oakland Research Institute (CHORI). *Rosa26* BAC: RP24–85I15, *Hprt* BAC: RP23–13N1, *Actb* BAC: RP23–249B5, *Rps21* BAC: RP23–16P3.

Generation of CHO-DUKX-B11 cell pools and single cell clones

Cell culturing. CHO-DUKX-B11 (ATCC CRL-9096) cells were maintained in ProCHO5 medium (Lonza) supplemented with 4 mM L-glutamine (PAA), Pen/Strep (Life Technologies), HT-supplement (Invitrogen) and phenol red (Sigma-Aldrich). Cells were cultured in a 37°C, 5% CO₂, humidified incubator without shaking.

Transfections. Lipofection: BACs and control plasmids were used for transfection with Freestyle MAX Reagent (Invitrogen). Briefly, 6.4 µg DNA and 16 µl Freestyle MAX Reagent (1:2.5 DNA : Freestyle MAX Reagent ratio) were incubated together in Opti-MEM (Life Technologies) serum free medium for 12 minutes and added to 2 × 10⁶ CHO-DUKX-B11 cells plated in 2 ml of ProCHO5 medium (Lonza) with supplements (see above) in 6-well tissue culture plates (Greiner Bio-One). BAC DNA was linearized

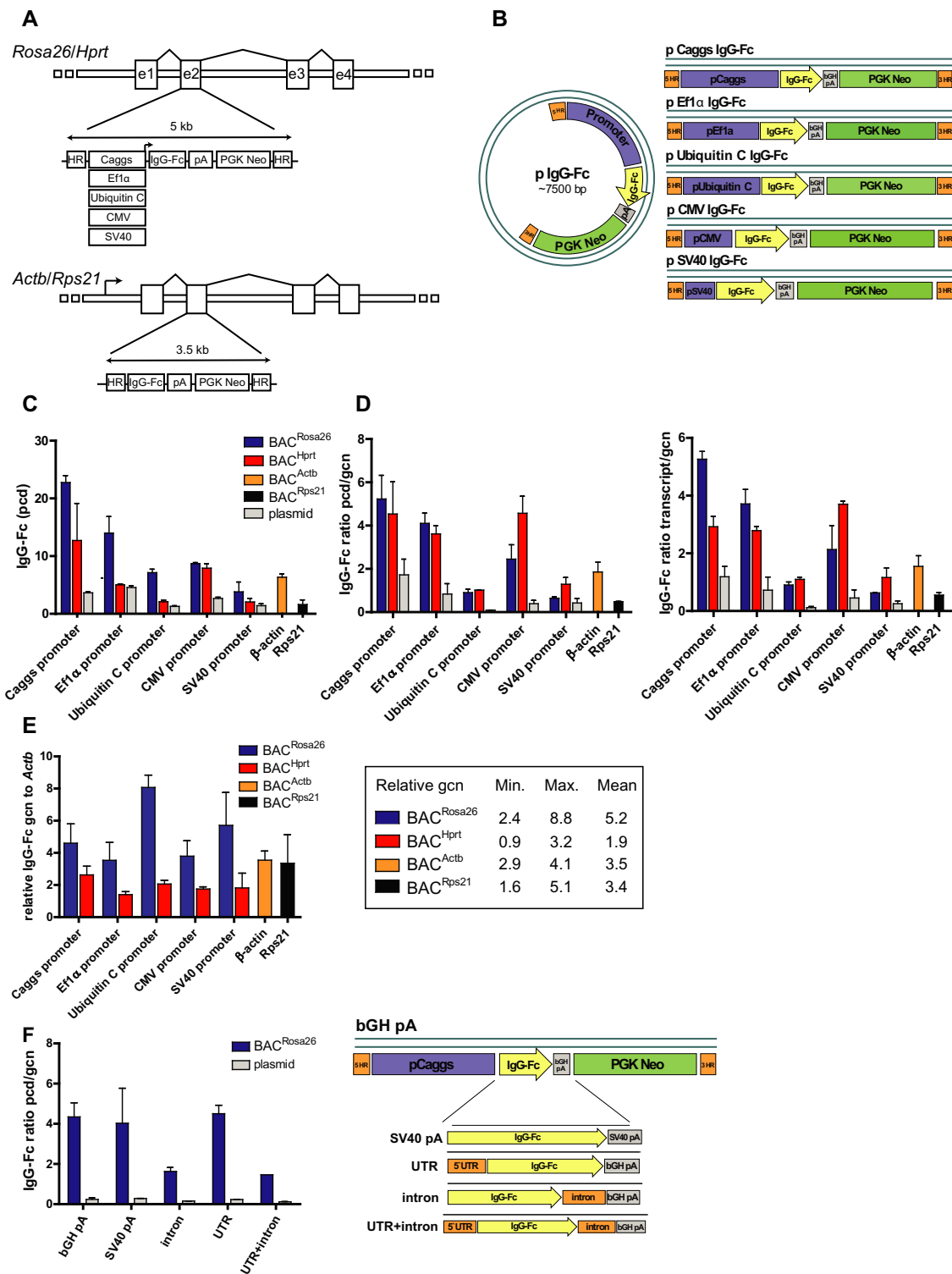


Figure 1. Optimization of BAC-based vector design. IgG-Fc was used as a model protein in every experiment. (A) Schematic view of BAC constructs. Incoming constructs were inserted into the second exon (e2) of the *Rosa26/Hprt* genes. In BACs containing the *Rosa26* or *Hprt* locus one of the following promoters was used: Caggs, Ef1α, Ubiquitin C, CMV, SV40. pA: polyadenylation signal, PGK Neo: neomycin selection cassette. HR: homology regions for BAC recombining. In BACs containing the *Actb/Rps21* loci the endogenous promoters were used to drive IgG-Fc expression. (B) Schematic view of plasmid controls; (C) Specific productivity (in pg/cell/day [pcd]) of IgG-Fc producing stable cell pools; (D) Specific productivity and relative transcript levels normalized to relative integrated vector (gene) copy numbers (gcn). gcn was calculated relative to genomic *Actb*; (E) Integrated gene copy numbers (gcn) relative to *Actb* in the different BAC derived cell pools. Observed minimum (Min.), maximum (Max.) and mean gcn values of the four BAC vectors are shown in a table format. (F) Normalized specific productivity of Caggs:BAC^{Rosa26} promoter cell pools containing the bGH polyA, SV40 polyA, an *Rps21*-derived 5' UTR or an intron (Clontech) sequence in the 3' position or the combination of both. Schematic view of these constructs is depicted on the right side of the figure, *n* = 2 independent transfections, error bars represent SEM.

prior to transfection by overnight digestion with the P1-SceI homing endonuclease (New England Biolabs). Antibiotic selection was started 2 days after transfection using 200 $\mu\text{g/ml}$ geneticin disulfate G418 (Carl Roth) and increased to 400 $\mu\text{g/ml}$ 14 days after transfection. Stable cell pools were analyzed 21 days after transfection.

Electroporation: CHO-DUKX-B11 cells were washed once with sterile 1x DPBS (Gibco) and then were resuspended in electroporation buffer (BioRad). 10 μg of plasmid or BAC DNA was electroporated into 4×10^5 CHO-DUKX-B11 cells in 200 μl total volume with a Gene Pulser Xcell (BioRad) electroporator. For pulsing the following protocol was used with 2 mm gap cuvettes: 160 V, 15 ms, square wave pulse. After electroporation cells were quickly transferred to 2 ml prewarmed growth media and were cultured as described above.

Cell counting and viability. CHO-DUKX-B11 cells were treated for 10 min, 37°C with Accumax (Sigma-Aldrich) solution prior to counting in order to dissociate cell clumps. Cells were counted then with CASY Cell Counter (Model TT, Roche).

Clone selection and stability. Single cell clones were obtained by limiting dilution technique. Cells from stable cell pools were plated to 96-well tissue culture plates (Thermo Fisher Scientific) in a 0.5 cell per well density. As sub-cloning medium 50% of 0.22 μm sterile filtered culture supernatant derived from CHO-DUKX-B11 cells and 50% fresh ProCHO5 medium with supplements (see above) containing 400 $\mu\text{g/ml}$ G418 was used. Culture supernatants of growing cells were pre-screened by ELISA. Clones from wells showing a high product concentration were subsequently expanded and analyzed for their specific productivities.

Clonal stability of different CHO-DUKX-B11 cell lines was monitored in T-25 flasks (Greiner Bio-One). Cells were cultured as described above without any antibiotic pressure. Cells were passaged into fresh medium every 3–4 days to 1×10^5 cells/ml. Specific and volumetric productivity was measured once per week.

Generation of CHO-K1/PG9 and CHO-K1/CN54gp140 cell lines

The two production clones CHO-K1/PG9 and CHO-K1/CN54gp140 were generated from a serum-free adapted host cell line derived from CHO-K1 (ATCC CCL-61). Transfection was performed using Lipofectin (Life Technologies). Briefly, 5 μg of P1-SceI (New England Biolabs) linearized BAC DNA was complexed with 25 μl Lipofectin in CD-DG44 medium (Life Technologies) containing 8 mM L-glutamine and 0.18% Pluronic F68 in a total volume of 200 μl . One million host cells, grown in CD-CHO medium (Life Technologies) supplemented with 8 mM L-glutamine were harvested in exponential growth phase and resuspended in 1.8 ml CD-DG44 medium with supplements. DNA complexes were added drop-wise to the cells and the transfection mixture was incubated in a humidified CO₂ shaking incubator (ISF1-X; Kuhner) for 24 hours at 37°C, 5% CO₂ and 125 revolutions per minute

(rpm). 18 ml CD-CHO medium supplemented with 8 mM L-glutamine and 500 $\mu\text{g/ml}$ G418 was added to the transfection mix and the culture was seeded in 96-well culture plates (Nunc, Thermo Fisher Scientific) at a volume of 100 μl per well. Plates were incubated in a humidified incubator at 37°C and 5% CO₂ until growth was observed (14–21 days). Positive wells were analyzed by ELISA. Clones from wells showing a high product concentration were subsequently expanded to a volume of 10 ml in T-25 cell culture flasks (Greiner Bio-One). The clones were monitored in routine culture for 3 passages before the best clone in terms of specific productivity and growth was used for single-cell dilution sub-cloning. Sub-cloning medium was prepared using 50% of 0.22 μm sterile filtered culture supernatant and 50% fresh CD-CHO medium supplemented with 8 mM L-glutamine and 500 $\mu\text{g/ml}$ G418. Cells were diluted in sub-cloning medium and one cell per well was seeded in a 384-well cell culture plate (Corning) in a volume of 50 μl per well. Plates were incubated in a humidified incubator at 37°C and 5% CO₂. Grown wells were expanded to 96-well culture plates (Nunc, Thermo Fisher Scientific). The culture supernatant of single-clones was evaluated according to the product concentration and best performing clones were further expanded, banked and analyzed.

Cell counting and viability. The cell density of CHO-K1 cell lines was calculated by counting the cell nuclei of lysed cells with the particle counter Multisizer 4 (Beckman Coulter). Cell viability was determined by trypan blue exclusion method using a Neubauer cell counting chamber. Growth rate μ [1/d] was calculated according to Equation (1) where X [cells] represents the total viable cell number and t [d] the time in days.

$$\mu = \ln(X_1/X_0) * 1/(t_1 - t_0) \quad (1)$$

Specific productivity qP [pg/cell/day] was calculated according to Equation (2) where P [pg] represents the product amount and X [cells] the total viable cell number.

$$qP = ((P_1 - P_0)/(X_1 - X_0)) * \mu \quad (2)$$

Clone stability. The clone stability of the CHO-K1/CN54gp140 and CHO-K1/PG9 recombinant cell lines was monitored in shake flasks (Corning) over a period of 20 passages (71–73 days) after thawing in ActiCHO SM medium (GE Healthcare Life Sciences) without selection pressure. Cells were passaged into fresh medium every 3–4 days to 2×10^5 cells/ml. The cell concentration, viability as well as product concentration was monitored before each splitting.

Fed-batch. Fed-batch experiments were performed using the ActiCHO media system including feed A and B (GE Healthcare Life Sciences). Fed-batches were carried out in shake flasks (Corning) with a starting volume of 45 ml and in the DASGIP Bioblock Bioreactor System (4 \times 1.2L, DASGIP, Eppendorf AG) with a starting volume of 650 ml and a starting cell concentration of $2\text{--}3 \times 10^5$ cells/ml. ActiCHO P medium supplemented with 8 mM L-glutamine served as batch medium, while feed A and feed B were fed to the culture daily starting on day 4. Feed B feed rate was calculated according to the current volume of the culture and

was 0.28% (vol/vol) per day. Feed A feed rate was calculated according to the D-glucose concentration of feed solution A and the current D-glucose concentration of the culture and set daily to 6.5 g/l using feed A. The cell concentration, viability, product concentration as well as the glucose concentration was analyzed daily.

Product concentration

The product concentrations of culture supernatants were determined using enzyme linked immuno-sorbent assays (ELISA).

Anti-human IgG-Fc ELISA. 96-well immuno-sorbent plates (Nunc MaxiSorp, Thermo Fisher Scientific) were coated with 100 μ l of goat anti-human IgG-Fc (Sigma-Aldrich) diluted to 1 μ g/ml in 1x Dulbecco's phosphate-buffered saline (DPBS) (Gibco, Life Technologies) and were incubated overnight at 4°C without shaking. Next day wells were blocked with 5% bovine serum albumin (BSA) (Sigma-Aldrich) in 1x DPBS after a washing step. 1x DPBS containing 0.05% Tween 20 (Sigma-Aldrich) was used as washing buffer. After a further washing step, standard (Fc from human plasma, Calbiochem) and samples were added to the blocked microwells and incubated for one hour. As detection antibody horseradish peroxidase conjugated protein A (ProtA-HRP) (Sigma-Aldrich) diluted 1:80 000 in 2% BSA in 1x DPBS was used. As a colorigenic substrate 3,3',5,5'-tetramethylbenzidine (TMB) was used (Sigma-Aldrich). The reaction was stopped after 20 minutes by addition of 30% H₂SO₄ and absorbance was measured at 450 nm (620 nm reference wavelength) in a microplate reader (TECAN Infinite 200 Pro). A sigmoidal curve from standard was calculated and plotted with GraphPad Prism5.

Anti-human IgG- λ ELISA. We used the same protocol as described above for IgG-Fc, except that we used goat anti-human lambda HRP conjugate (Southern Biotech) as a detection antibody in a 1:2000 dilution. IgG1 lambda from human myeloma plasma (Sigma-Aldrich) was used as a standard.

CN54gp140 ELISA. 96-well immuno-sorbent plates (Nunc MaxiSorp, Thermo Fisher Scientific) were coated with 100 μ l Lectin from *Galanthus nivalis* (Sigma-Aldrich) diluted to 5 μ g/ml in coating buffer at 4°C overnight. 50 μ l of samples and standards, diluted in dilution buffer in a 1:2 dilution series, were transferred to the pre-coated plate and incubated for one hour on a shaker at 300 rpm. A₂₈₀ quantified affinity purified CN54gp140 was used as standard at a starting concentration of 100 ng/ml. Next, the plate was incubated with 50 μ l of monoclonal anti-gp41 (HIV-1) IgG1 5F3 (Polymun Scientific, Austria) diluted to 1 μ g/ml in dilution buffer for one hour. Afterward, the plate was incubated with 100 μ l of HRP conjugated, polyclonal, γ -chain specific goat anti-human antibody. The color reaction was induced by adding 100 μ l of staining buffer (1.5 mg/ml o-phenylene diamine diluted in 0.15 M citric acid buffer, pH 5.0 containing 0.01% H₂O₂) to each well. After incubation for 5 min, the reaction was stopped by the

addition of 100 μ l 25% H₂SO₄ to each well. Absorption was measured at 492 nm (620 nm reference wavelength) using a microplate reader, and mean concentrations were calculated relative to the standard curve.

PG9 binding assay. To test the binding specificity of the produced PG9 antibody, 96-well immuno-sorbent plates (Nunc MaxiSorp, Thermo Fisher Scientific) were coated with 100 μ l gp120 ZM109 (kindly provided by Dr. Lukas Mach, University of Natural Resources and Life Sciences, Vienna) diluted to 1 μ g/ml in coating buffer overnight at 4°C. After each incubation step the plates were washed three times using washing buffer. PG9 antibody as well as negative control 2F5 (Polymun Scientific, Austria), a monoclonal anti-gp41 HIV-1 antibody, were diluted to a starting concentration of 1 μ g/ μ l in dilution buffer and further diluted in a 1:2 dilution series. 50 μ l of the dilutions were transferred to the pre-coated plates and incubated for one hour on a shaker at 300 rpm. Afterward, the plates were incubated with 100 μ l of HRP conjugated, polyclonal, γ -chain specific goat anti-human antibody (Life Technologies) and stained as above described for CN54gp140 ELISA.

Real-time PCR analysis

Sample preparation for analysis. Genomic DNA and RNA were isolated from stable transfected cells. RNA was isolated with TRIzol Reagent (Life Technologies) according to the manufacturer's instructions. RNA integrity was assessed by loading each sample to a MOPS gel. RNA was treated with DNaseI (Thermo Fisher Scientific) prior to reverse transcription by RevertAid First Strand cDNA Synthesis Kit (Thermo Fisher Scientific). Genomic DNA was isolated by QIAamp DNA Blood Mini kit (Qiagen) according to the manufacturer's instructions.

IgG-Fc DNA and cDNA were detected by a TaqMan assay. As a genomic control *Actb* was used. TaqMan probes as well as qPCR Core kit were ordered from Eurogentec. PG9 antibody heavy and light chain DNA and cDNA were detected by a Sybr Green assay. We used Taq polymerase from 5PRIME for this assay. Primers were ordered from Eurofins Genomics. As a template 5 ng genomic DNA or cDNA derived from 50 ng total RNA was used per reaction in a 25 μ l final volume. See probe and primer sequences under the headline 'Primer and probe list'.

Calculation of integrated gene copy numbers and mRNA expression levels. Integrated gene copy numbers (gcn) as well as relative transcript levels were calculated with the $2^{-\Delta\Delta C_T}$ method (29). IgG-Fc/ PG9 HC/ PG9 LC integrated copy numbers and mRNA expression (cDNA) levels were measured by real-time PCR using the genomic *Actb* and *Actb* cDNA as a reference for relative quantification. All gcn and relative transcript values in this study are indicated in x-fold values relative to genomic *Actb* and *Actb* transcript. PCR amplification efficiency ranged from 96% to 100%.

Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGEs were performed using the XCell Sure Lock Minicell electrophoresis system and precast gels (Life Technologies). The increase of product during fed-batches was determined on NuPAGE 4–12% bis-tris gels by applying 15 μ l of reduced (10% (vol/vol) 0.5 dithiothreitol (DTT)) or non-reduced culture supernatants as well as specific standards. Gels were stained using SYPRO Ruby protein gel stain (Life Technologies) and fluorescence was detected using a biomolecular imager (ImageQuantTM LAS4000, GE Healthcare Life Sciences). The quality of CN54gp140 was determined on a NuPAGE 3–8% tris-acetate gel by applying 20 μ g of non-reduced, purified product. The gel was stained using Coomassie brilliant blue.

Primer and probe list

Real-time PCR analysis

TaqMan probes:

TaqMan IgG-Fc: 5' 6-FAM – ACAAGTGCAAGGT-GAGCAACAAGGC - BHQ-1 3'

TaqMan ACTB: 5' 6-FAM – CCATCCTGGCCTCACT-GTCCACCT – TAMRA 3'

Primers:

IgG-Fc for: CACCAGGATTGGCTGAATG

IgG-Fc rev: TGGAGATGGTCTTCTCGATG

ACTB for TaqMan: TGAGCGCAAGTACTCTGTG

ACTB rev TaqMan: TTGCTGATCCACATCTCCTG

PG9 HC for Sybr: ACCGTCAGTCTTCTCTTCC

PG9 HC rev Sybr: TCCTTGCCATTCAGCCAGTC

PG9 LC for Sybr: CTGCAATGGAACCAGCAATG

PG9 LC rev Sybr: TCTCGTGCTTGTCAGAGACT

ACTB for Sybr: CGTACCCTGGCATTGTGAT

ACTB rev Sybr: GGCAACATAGCACAGCTTCT

The following primers have been used to amplify the homology regions for BAC recombination by ET-cloning:

Rosa 3 HR exon2 for: TTTTTTCTCGAGGTCTTGTC
ACACTCAAGAAGCTACC

Rosa 3 HR exon2 rev: TTTTTTTTAATTAAGCTGAT
TCTGCAGGCCGACTTCAG

Rosa 5 HR exon2 for: TTTTTTGCGATCGCCTGCAT
GCAGTGGAACACTCTTGTC

Rosa 5 HR exon2 rev: TTTTTTCCTTAGGGGAGATG
TTTCTAAAAGGACCAACAGTC

Rosa 3 HR exon9 for: TTTTTTCTCGAGCAGATATG
CCATTTGGAAAAAGGT

Rosa 3 HR exon9 rev: TTTTTTTTAATTAAGTGTCC
AAGAAGCAAGTGCTTCTC

Rosa 5 HR exon9 for: TTTTTTGCGATCGCCTTGGG
AAAATATATTGAAATATC

Rosa 5 HR exon9 rev: TTTTTTCCTTAGGATCCACA
GAAGCAGTTCTCAGTG

Rosa 3 HR Thumpd for: TTTTTTCTCGAGGGAGTG
CCGTGCGCCGAGCCCG

Rosa 3 HR Thumpd rev: TTTTTTTTAATTAAGTGTCC
TTTCCCCCGGGCCCGGTCGTG

Rosa 5 HR Thumpd for: TTTTTTGCGATCGCCGCT
CAGAGACTCACGCAGCCCT

Rosa 5 HR Thumpd rev: TTTTTTCCTTAGGGCGAG
ACTCGAGTTAGGCCCAACG

Hprt 3 HR for: TTTTTTCTCGAGGGCTAGATGCC
CATGAGGGCGGCG

Hprt 3 HR rev: TTTTTTTTAATTAAGTGTCC
CTGAAGCCGCCCTCCG

Hprt 5 HR for: TTTTTTGCGATCGCGCCAGCG
GAGCCTCCGGGGACGGAGCC

Hprt 5 HR rev: TTTTTTCCTTAGGGGACCGGTC
GGCTCGCGGCAAAAAGCGG

Actb 3 HR for: TTTTTTCTCGAGGATATCGCTGCG
CTGGTCGTCGACAAC

Actb 3 HR rev: TTTTTTTTAATTAAGTGTCC
AACAGCCTTCTTAGCAC

Actb 5 HR for: TTTTTTGCGATCGCCTGCCCTAGG
TCCGCTCCGGGCC

Actb 5 HR rev: TTTTTTATTTAAATGGCGAACTAT
CAAGACACAAAAGAAG

Rps21 3 HR for: TTTTTTCTCGAGACGACGCCGG
CGAGTTTGTGGACC

Rps21 3 HR rev: TTTTTTTTAATTAATCATCTGG
ATGGACGCGTGGTCCTTG

Rps21 HR for: TTTTTTGCGATCGCGCTGGACA
AGTGAGCAGGTAGGCC

Rps21 HR rev: TTTTTTATTTAAATCCTGAGGCCG
TCCCTGCAGAGTTAC

Sequence information for genetic elements

Rps21 5' UTR (*Mus Musculus*):

CCCTCCTGCCCAATAGTGTCTGCGGAGGCACGA
GCTACTTCCTTTCTGCGCTCTCGCTGGACAAGTG
AGCAGGTAGGCCTCGGTGCGGGCTGTGTGCGA
GCGGGGCCGGGCTGGGCCGGGCTGGGGCGCA
GTTGGGGAGGTCGCGGTCACATGGCCGAGAGT
GCTTGCGGGCCATGGGTCTGGTTGGGCCCGCA
GACTCGGAGTGCGGGCCTGACCAGTCTCAAGA
CCTGCGTGGTAACTCTGCAGGGACGGCCTCAG
GATG

Artificial intron (Clontech):

GAATTAATTCGCTGTCTGCGAGGGCCAGCTGT
TGGGGTGAGTACTCCCTCTCAAAGCGGGCAT
GACTTCTGCGCTAAGATTGTCAAGTTTCCAAA
ACGAGGAGGATTTGATATTCACTGGCCCGCG
GTGATGCCTTTGAGGGTGGCCGCGTCCATCTG
GTCAGAAAAGACAATCTTTTTGTTGTCAAGCT
TGAGGTGTGGCAGGCTTGAGATCTGGCCATAC
ACTTGAGTGACAATGACATCCACTTTGCCTTTCT
CTCCACAGGTGTCCACTCCAGGTCCAAGTGC
AGGTGC

RESULTS

Optimization of BAC-based vector design

We used two strategies to design BAC-based expression vectors appropriate for recombinant protein production. First, we searched for BACs containing a permissive chromatin environment and combined them with strong ectopic promoters. Second, we searched for BACs containing loci highly transcribed in mammalian cells and expressed the protein of interest under the endogenous promoter. For the first approach, we decided to use two BACs harboring the

murine *Rosa26* and *Hprt* (hypoxanthine phosphoribosyltransferase) loci. The *Rosa26* and *Hprt* loci are well known to contain permissive chromatin and are widely used in the mouse transgenic field for targeted transgene integration (30–33). We constructed a series of plasmids combining different conventional promoters (Caggs, Efl α , Ubiquitin C, CMV and SV40) expressing the Fc part of human IgG1 (IgG-Fc) as a model protein and recombined them into the *Rosa26* and *Hprt* BACs (Figure 1a and b). Regarding the second approach, we used BACs harboring the murine *Actb* (β -actin) and *Rps21* (ribosomal protein 21) loci, which are highly transcribed in mammalian cells. The incoming (promoterless) constructs were inserted in frame to the translational start site of the *Actb* and *Rps21* genes by recombining (27,28), thus the IgG-Fc protein was expressed under control of the respective endogenous promoters (Figure 1a).

To test the performance of the different BAC-based expression vectors we transfected them into CHO-DUKX-B11 (34) suspension cells. Plasmids with identical expression units were used as controls. We selected transfected cells for three weeks with neomycin and analyzed secreted protein production levels in the stable cell pools. The use of the *Rosa26* BAC backbone (BAC^{Rosa26}) in combination with Caggs promoter resulted in the highest protein expression levels in the cell pools (22.7 ± 1.2 pg/cell/day [pcd], values represent mean \pm SEM) followed by the constructs combining the BAC^{Rosa26} with the Efl α promoter (13.9 ± 2.8 pcd) and the *Hprt* BAC backbone (BAC^{Hprt}) combined with the Caggs promoter (12.7 ± 6.4 pcd, Figure 1c). Of note, cell pools generated with the respective control plasmids showed considerably lower productivity. Cell pools generated with combinations of the BAC^{Rosa26} and the BAC^{Hprt} with the CMV, the Ubiquitin C and the SV40 promoters, as well as the *Actb* and *Rps21* constructs yielded only moderated expression levels (Figure 1c). Next, we performed a direct comparison of the strength of each construct by normalizing the secreted protein and transcript levels to the integrated vector copy number in the cell pools. The combination of the BAC^{Rosa26} and the Caggs promoter proved to be the best choice regarding normalized transcription compared to all the other constructs, but mostly we did not find striking differences between the *Rosa26* and the *Hprt* BACs (Figure 1d). Interestingly, we observed an overall lower copy number integration for the BAC^{Hprt} constructs (1.9 relative copies on average) compared to the BAC^{Rosa26} constructs (5.2 relative copies on average) in the stable cell pools (Figure 1e). This phenomenon could explain why in general the productivity of the cell pools is higher with the BAC^{Rosa26} compared to the BAC^{Hprt} constructs despite of their rather similar normalized protein productivity and transcript levels.

Having identified the BAC^{Rosa26} in combination with the Caggs promoter (Caggs:BAC^{Rosa26}) as the best genetic tool, we tested if other genetic elements may improve its performance. We replaced the bovine growth hormone (bGH) polyA by a Simian virus 40 (SV40) polyA, added an artificial intron sequence derived from Clontech and a 5' untranslated region (UTR) sequence derived from the murine *Rps21* gene to the original construct. The use of these elements did not result in any beneficial effect. Interestingly,

the use of an intron sequence downstream of the transgene decreased the expression levels alone and in combination with the 5' UTR (Figure 1f).

Next, we investigated the clonal distribution within the cell pools generated with the Caggs:BAC^{Rosa26}, the Caggs:BAC^{Hprt} and Caggs:plasmid control. We isolated single cell clones by limiting dilution. 250 single-cells were plated of which 45–55 clones (18–22%) were growing and were pre-screened for IgG-Fc production. The best producing 20 clones from each pool were further characterized (Figure 2a). The Caggs:BAC^{Rosa26} derived clones had a productivity ranging from 15 to 87 pcd, while Caggs:BAC^{Hprt} derived clones ranged from 13 to 26.5 pcd. Productivity of the plasmid derived clones was considerably lower (1.8 to 10 pcd). Thus, the BAC-based expression vectors represent a considerable improvement compared to plasmid-based expression vectors. Indeed, stable cell pools generated by electroporation of the Caggs:BAC^{Rosa26} construct expressing eGFP showed that all clones of the pool expressed the transgene (Figure 2b). In addition, the mean fluorescence intensity of the Caggs:BAC^{Rosa26} derived cells was considerably higher compared to the Caggs:plasmid control. We observed two distinct cell populations within the Caggs:plasmid pool. This could reflect an oligoclonal population after cell transfection or that the eGFP negative cells could lose the eGFP expression via transgene silencing which is often observed in plasmid based transgenesis. Furthermore, unlike in the case of plasmids (35,36), copy number analysis of Caggs:BAC^{Rosa26} clones showed a direct correlation between the specific productivity and the integrated vector copy number, ranging from 0.5 to 25.5 copies relative to genomic *Actb* (Figure 2c).

Fed-batch production of CN54gp140 using an optimized BAC-based vector

To gain further insight into the usefulness of the Caggs:BAC^{Rosa26} vector applied to industrial processes, we produced the HIV-1 (CN54(37)) envelop glycoprotein CN54gp140 in CHO-K1 cells using fed-batch cultures. Of note, the CN54gp140 protein is highly glycosylated (38) and is considered to be difficult to express (39,40). Fermentation of a cell line generated with the Caggs:BAC^{Rosa26} vector using a bioreactor or shake flask yielded up to 1 g/l (Figure 3a and b). For comparison, previous cell lines expressing gp140 (GS amplified) yielded maximum in the single digit mg/l range (39,40). The Caggs:BAC^{Rosa26} CN54gp140 cell line (CHO-K1/CN54gp140) was tested for stability in terms of specific and volumetric productivity by routine cultivation in ActiCHO SM medium without selection pressure for 20 passages (71 days) (Figure 3c). The specific and volumetric productivity stayed constant over the whole period of time with a mean \pm SEM of 2.56 ± 0.16 pcd and 13.15 ± 1.24 mg/l. The growth rate was also constant over time and averaged 0.89 ± 0.02 per day. Furthermore, the quality of the affinity purified CN54gp140 was assessed on a non-reducing SDS-PAGE (Figure 3d). Most of the purified product was present in trimers, but also mono-, di- and higher oligomers were detected. These results highlight the stability and high yield

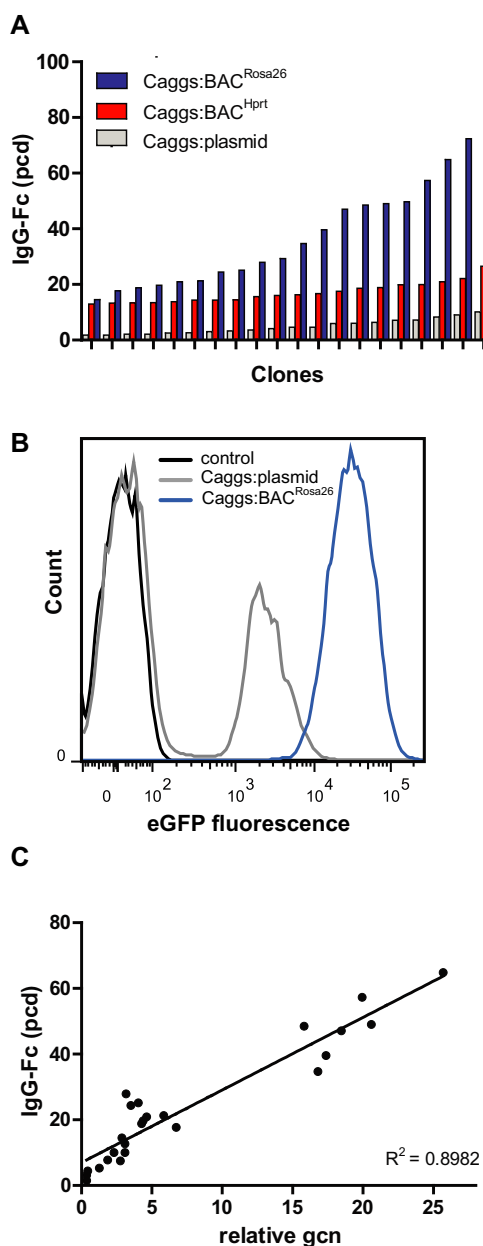


Figure 2. Clonal distribution of IgG-Fc cell pools. (A) Specific productivity (pcd) of the 20 best clones out of 50 pre-screened clones derived from stable cell pools generated with the Caggs:BAC^{Rosa26}, Caggs:BAC^{Hprt} and Caggs:plasmid vectors. (B) FACS analysis shows eGFP fluorescence of Caggs:BAC^{Rosa26} and Caggs:plasmid stable cell pools expressing eGFP compared to empty plasmid control cells. (C) Correlation between specific productivity of Caggs:BAC^{Rosa26} clones and their integrated vector (gene) copy numbers (gcn) relative to *Actb* gcn, $n = 25$.

of BAC-vectors derived cell lines and provides a proof of concept for its use at industrial level.

Production of an anti-HIV-1 antibody as a complex protein model using BAC-based vectors

To optimize the BAC-based expression system for production of complex proteins, we chose an antibody as a model protein. PG9 (IgG1- λ) is a neutralizing antibody which tar-

gets the variable regions 1 and 2 (V1/V2) of the HIV-1 envelope glycoprotein 120 (gp120) (41). We explored three different strategies to produce an antibody using the BAC expression system. First, we generated two independent expression cassettes for the PG9 light chain (LC) and heavy chain (HC) and placed them at different positions into the BAC^{Rosa26}. We tested three different insertion sites to minimize transgene silencing due to transcriptional interference between the expression cassettes. In the BAC 10, BAC 20 and BAC 30 vectors, heavy chain (HC) and light chain (LC) constructs were recombined 10, 20 and 30 kb apart from each other (Figure 4a). Second, we linked the expression of PG9 LC and HC using either the FMDV 2A peptide or an Internal Ribosome Entry Site (IRES) and inserted the resulting expression cassette into a single location into the BAC^{Rosa26} (Figure 4a). Third, PG9 LC and HC expression cassettes were inserted into two different BAC^{Rosa26} vectors and cell lines were generated by co-transfecting both BACs (Figure 4a).

We generated stable cell pools by transfecting all the BAC^{Rosa26} constructs and corresponding control plasmids into CHO-DUKX-B11 cells followed by antibiotic selection. Antibody production in the stable cell pools showed higher yields using BAC vectors compared to plasmids with relatively similar values between different BAC constructs (Figure 4b). All BAC pools produced between 3.52 and 4.59 pcd, with the exception of the EMCV IRES BAC (1.72 pcd). Analysis of transcript expression levels normalized to the integrated vector copy number revealed that all the BAC-based constructs expressed higher levels of HC and LC than their respective control plasmids. Interestingly, we observed that the LC transcripts were more abundant than the HC in all the BACs containing two distinct expression cassettes (Figure 4c).

We then analyzed the clonal distribution of the cell pools generated with the BAC HC and BAC LC (BAC HC/LC), BAC 10, BAC 20, and control plasmid p HC and p LC constructs (p HC/LC). Out of 300 plated single cells, 51–111 clones (17–37%) were growing. 50 clones per construct were pre-screened for antibody production and the best 30 were used to analyze the specific productivity. While single clones derived from the BAC 10 and BAC 20 pools showed little variability regarding their productivity, clones derived from the co-transfection of HC and LC BAC constructs (BAC HC/LC) showed a heterogeneous distribution and the highest expression levels, with the best clone producing 57 pcd (Figure 4d). Of note, the majority of single clones derived from the pool generated with the plasmid control (co-transfection, p HC/LC) did not express detectable levels of the antibody, reaching a maximum specific productivity of 7.18 pcd.

LC expression levels were consistently higher than the HC in cell pools generated with two distinct expression cassettes (Figure 4c). Since unbalanced HC/LC expression ratio may limit the assembly and secretion of antibodies, we analyzed the HC and LC transcript levels of single cell clones derived from the BAC HC/LC pool and correlated to the antibody productivity. We found a strong correlation of the specific productivity of clones with their HC transcript levels ($R^2 = 0.9497$; $P < 0.0001$) but not with their LC expression levels ($R^2 = 0.4203$; $P = 0.0589$) (Figure 4e). This

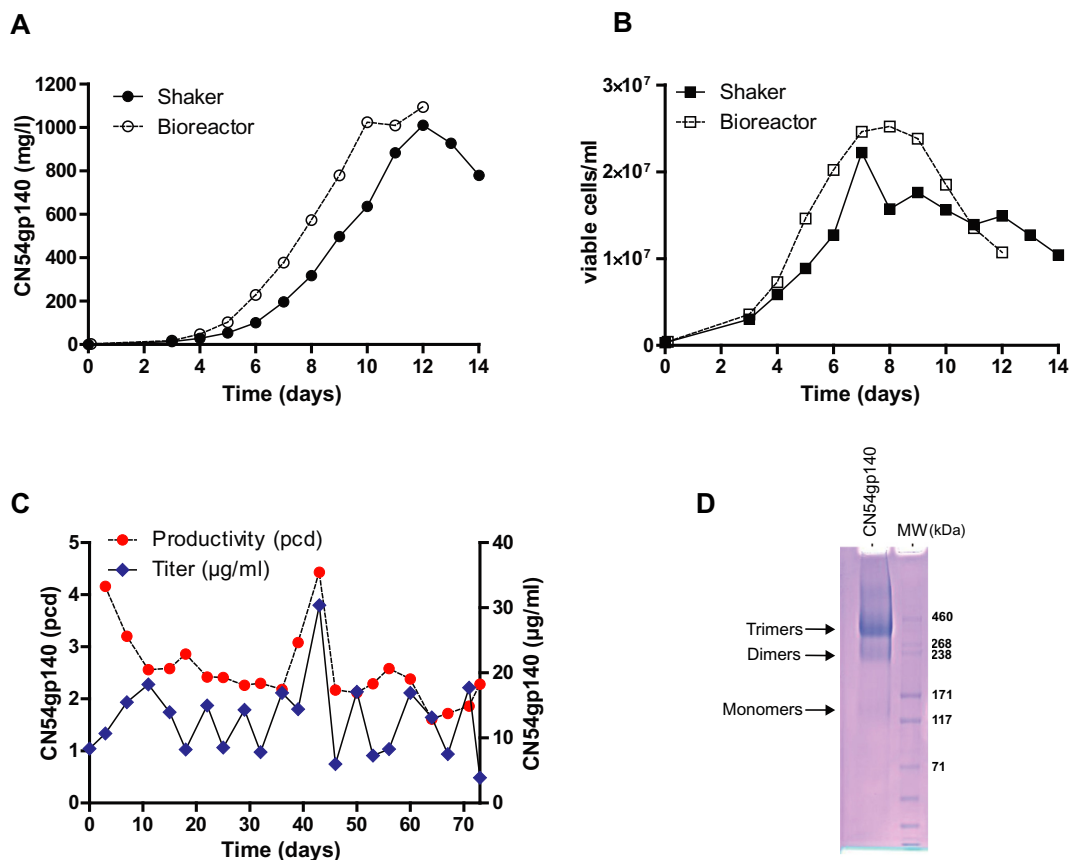


Figure 3. Fed-batch production of CN54gp140 using an optimized BAC-based vector. Fed-batch production of CN54gp140 with the Caggs:BAC^{Rosa26} vector in CHO-K1 cells. (A) Product accumulation (mg/l) and (B) viable cell densities (viable cells/ml) are compared in shake flasks and in bioreactors. (C) Long-term production stability of the Caggs:BAC^{Rosa26} CN54gp140 cell line without selection pressure. (D) Quality control of purified CN54gp140 in a non-reducing SDS-PAGE.

suggests that the HC expression is a limiting factor for the production of the PG9 antibody.

We monitored long-term production stability in the absence of antibiotic selection pressure for seven weeks in BAC (BAC HC/LC) and plasmid (p HC/LC) derived clones. Specific and volumetric productivity stayed relative stable during this time period, only a few highly expressing BAC derived clones showed a slight decrease in specific productivity (Figure 5a and b). We further examined the long-term stability of BAC HC/LC clones regarding their integrated vector copy numbers. Transgene copy numbers of HC and LC did not substantially change during this time (Figure 5c).

Finally, we assessed the suitability of the BAC-based expression vector system by producing the PG9 antibody using bioreactors and shake flasks. For this purpose, we generated a PG9 producing cell line in CHO-K1 cells using the BAC 20 strategy. This cell line was extensively screened and assessed for its usability in industrial application. PG9 antibody production reached 1.12 g/l in shake flasks and 0.75 g/l in lab-scale bioreactors with 0.8×10^7 and 1.1×10^7 maximal viable cell densities in 13 and 11 days fed-batch fermentation processes, respectively (Figure 6a and b). Specific and volumetric productivity of the cell line was stable over 20 passages (73 days) without selection pressure (Fig-

ure 6c). SDS-PAGE analysis of supernatants confirmed that PG9 was correctly assembled and secreted. Furthermore, secreted PG9 antibody was functional as shown by a binding assay to its target gp120 (ZM109) (Figure 6e and d). These results offer a proof of principle for industrial production of antibodies using the BAC expression system.

DISCUSSION

While BAC vectors are commonly used in the mouse transgenic field as expression vectors, their use for recombinant protein production in mammalian cells is still not well established.

Here we describe a comprehensive analysis of BAC-based expression vectors applied to recombinant protein production in CHO cells and report that BAC vectors containing open chromatin regions highly increased recombinant protein production for all proteins tested in this study. However, BAC vectors increase protein production by increasing the transcription of the GOI, thus, they may not offer a significant advantage if non-transcriptional related processes (i.e. translation, folding, transport or secretion) are the limiting factors in the production of the protein of interest.

We found that BAC vectors containing highly transcribed genes such as *Actb* and *Rps21* can be used to express a GOI from the endogenous promoter, albeit resulting in moder-

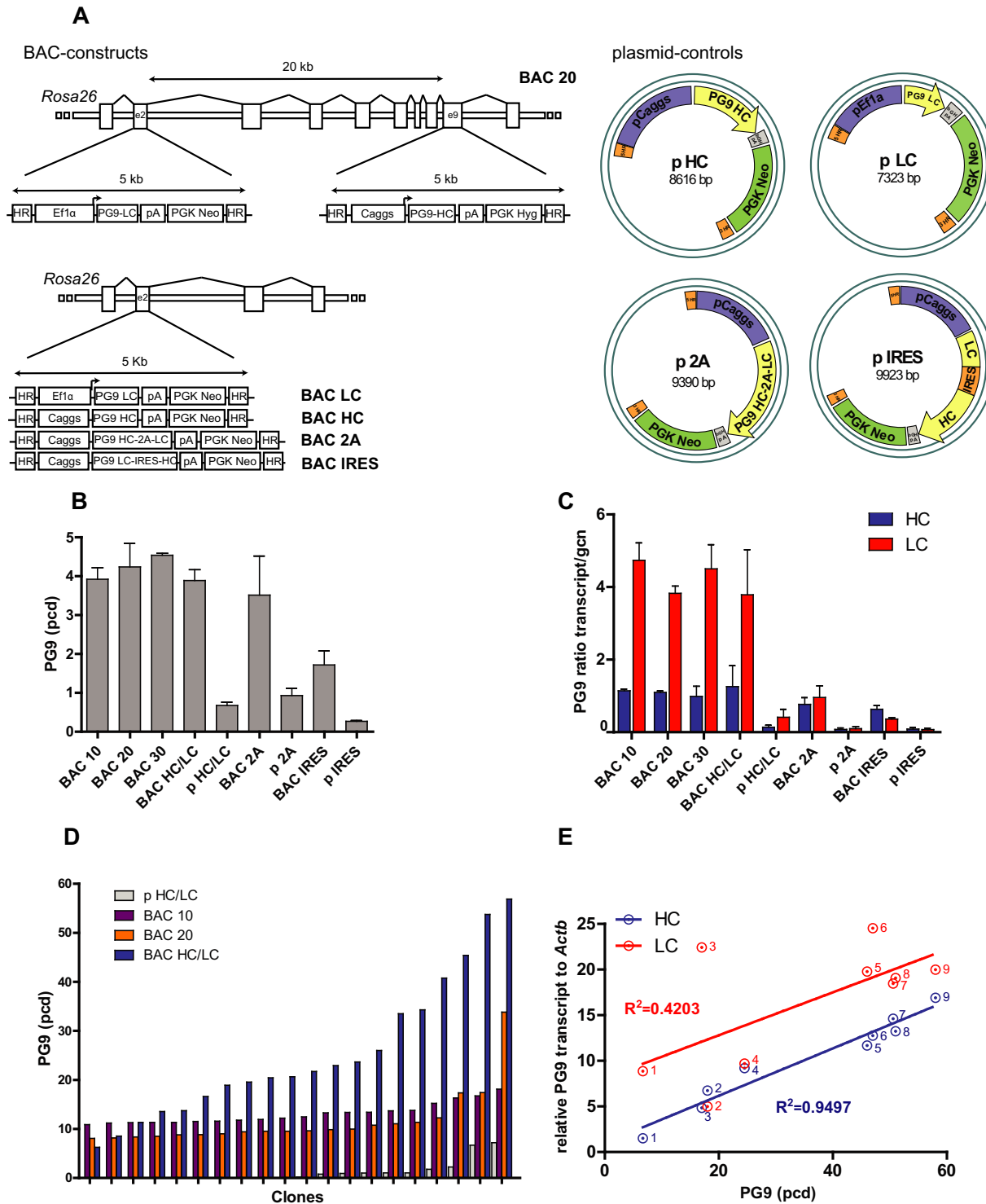


Figure 4. Strategies for production of the anti-HIV-1 antibody PG9 as a complex protein model using BAC-based vectors. (A) Schematic view of human IgG1-λ PG9 antibody constructs: BACs and plasmid controls. In the case of BAC 10, BAC 20 and BAC 30 constructs, heavy chain (HC) and light chain (LC) expression units were inserted into the same BAC^{Rosa26} at different locations 10, 20 and 30 kb apart from each other (example shows BAC 20). Lower panel, BAC LC and BAC HC vectors: LC and HC expression cassettes were recombined into two independent BAC^{Rosa26} vectors. Stable cell pools were generated by co-transfection of these two independent BACs (BAC HC/LC). BAC 2A and BAC IRES: LC and HC expression cassette was inserted into a single BAC. LC and HC expression was linked by a bicistronic mRNA using the FMDV 2A peptide or an IRES element. (B) Specific productivity of stable PG9 antibody expressing cell pools (pcd) generated with the constructs described above (p HC/LC, p 2A and p IRES are respective plasmid controls). (C) Normalized HC and LC transcript levels to the gene copy number relative to genomic *Actb* (gcn) of integrated vectors in the stable cell pools. (D) Specific productivity in pcd of the 20 best clones from 50 pre-screened clones derived from the BAC HC/LC, BAC 10, BAC 20 and plasmid HC/LC (p HC/LC) stable cell pools. $n = 2-4$ independent transfections, error bars represent SEM. (E) Correlation between HC and LC transcript levels (normalized to *Actb*) and specific productivity (pcd) of nine PG9 BAC HC/LC-derived clones. Specific productivity correlates with HC transcript with an $R^2 = 0.9497$.

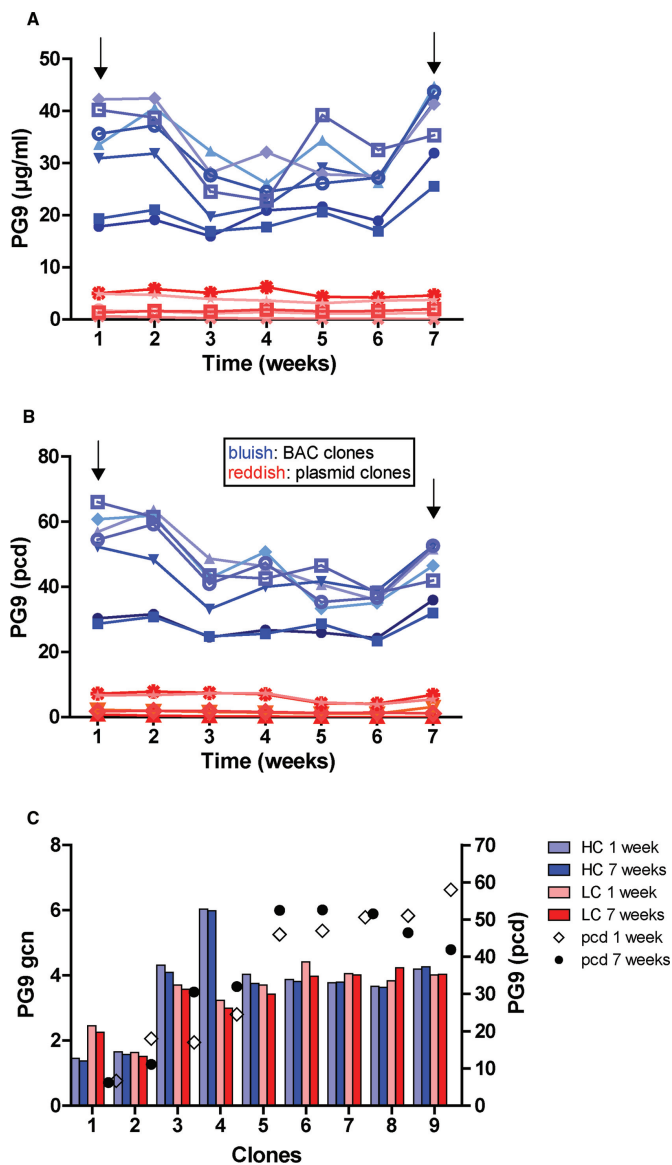


Figure 5. Long-term stability of PG9 antibody producing clones without selection pressure. (A) Volumetric ($\mu\text{g/ml}$) and (B) specific productivity (pcd) of BAC (in bluish color) and plasmid (in reddish color) clones generated from BAC HC/LC and p HC/LC stable pools during a 7 weeks culturing period. Arrows indicate sample collection for genomic DNA isolation. (C) LC and HC integrated vector copy changes were monitored in 9 BAC^{Rosa26} HC/LC clones. Gene copy numbers relative to genomic *Actb* (gcn) are indicated as bars. Light bars indicate gcn at the first week, dark bars indicate gcn at the seventh week (bluish color, HC; reddish color, LC); specific productivity of each clone is indicated on a dot plot.

ated expression levels. On the other hand, the combination of BACs containing permissive chromatin regions with a potent (heterologous) promoter resulted in cell lines expressing high protein levels. Interestingly, transgene expression levels were highly dependent on the combination of open chromatin containing loci and the heterologous promoters. For example, as shown in Figure 1c and d, the SV40 and Ubiquitin C promoters (which are well-known to be 'potent' promoters) gave poor results when combined with the *Rosa26* or *Hprt* BACs. This illustrates the complexity of

chromatin regulation and underscores that even loci which are considered to be permissive or 'open chromatin' may not be compatible with some ectopic promoters. Similar observations were obtained by an RMCE study in which the transgene expression levels were dependent not only on the targeting constructs but also on their combination with the genomic integration sites and even on the relative orientation of the transgene (42). This highlights the importance of testing different promoter sequences with different chromatin contexts, since promoter strength can vary broadly between different constructs and cell types (43).

We identified BAC^{Rosa26} in combination with the Caggs promoter as the BAC-based expression vector producing the highest protein and transcript levels. The positive effects mediated by the chromatin environment of *Rosa26* gene are not restricted to one particular cell type or to a particular protein. Beneficial effects have been reported in mouse studies (30), in human cells (HEK 293 (44)) and in hamster cell lines (CHO (45)). By using IgG-Fc as a model protein, we could demonstrate that it is possible to obtain CHO cell pools producing 24 pcd in three weeks using the Caggs:BAC^{Rosa26} vector. In addition, single cell clones producing up to 87 pcd could be isolated by analyzing as few as 50 clones. Furthermore, unlike in the case of plasmids (35,36), the productivity of BAC-based clones was directly proportional to their integrated transgene copy numbers in the host cell. This phenomenon indicates that BACs are able to overcome chromatin positional effects upon stable integration into the CHO cell genome. In addition to IgG-Fc, we showed the advantages of the BAC vectors in the production of HIV-1 envelope glycoprotein CN54gp140. We could reach a yield of 1 g/l in fed-batch cultures by using the BAC^{Rosa26} vector. CN54gp140 is a highly glycosylated protein and considered to be difficult to express. Indeed, productivity of cell lines expressing recombinant gp140 are reported to be mainly in the single digit mg/l range (39,40). Moreover, we could show that the BAC^{Rosa26} vector conferred long-term production stability and optimal protein quality.

We have demonstrated that complex proteins such as antibodies can be successfully expressed from a single BAC-based vector containing two separate expression units placed at different locations within the BAC sequence. However, according to our results, co-transfection of two BAC vectors carrying each one of the polypeptide chains seems to be more beneficial. This could be due to promoter interference when both expression units are placed in the same vector (46) although the distance was 30 kb or may reflect the complexity of antibody production. By placing the HC and LC of the PG9 antibody in the same BAC vector, the ratio between HC and LC expression is fixed which may not be optimal for appropriate antibody folding and secretion (47). On the other hand, expressing HC and LC from two independent BACs may result in different HC/LC expression ratios depending on the number of copies integrated of each BAC. Therefore, it could be easier to find clones with the appropriate HC/LC ratio and antibody expression. Indeed, our results suggest that the expression of the HC is the limiting factor in the production of the PG9 antibody and this should be taken into account in future design of BAC vectors for antibody production.

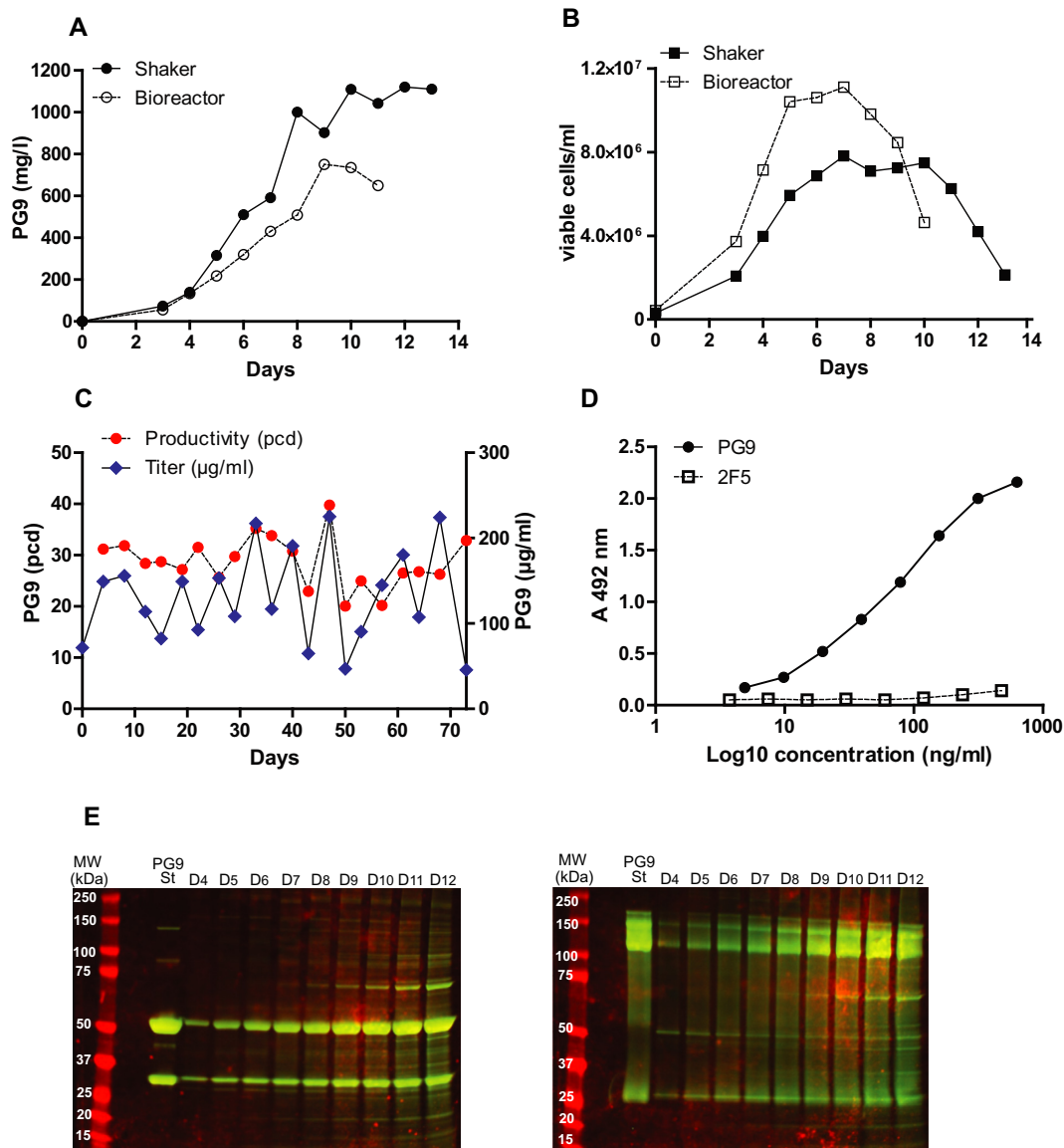


Figure 6. BAC-based PG9 antibody fed-batch production in shake flasks and bioreactors. (A) Product accumulation (mg/l) and (B) viable cell densities (viable cells/ml) in shake flask and bioreactor fed-batch cultures. (C) Specific (pcd) and volumetric ($\mu\text{g/ml}$) productivity of CHO-K1/PG9 cell line in routine cultivation without selection pressure over 20 passages (73 days). (D) Binding curve of PG9 and negative control 2F5 to gp120 (ZM109), PG9 EC₅₀ was calculated to be 64 ng/ml. (E) SDS-PAGEs of fed-batch culture supernatants under reducing (left panel) and non reducing (right panel) conditions from CHO-K1/PG9 cell line at different days (D4-D12, PG9 St: antibody standard).

Nevertheless, the use of BAC^{Rosa26} to express the PG9 antibody was much superior to a conventional vector allowing us to isolate single cell clones with a productivity of 57 pcd and more than 1 g/l in bioreactor and shake flasks. Furthermore, binding analysis demonstrated the functionality and quality of the produced PG9 antibody. In addition, we have identified two additional insertion sites within the BAC^{Rosa26} vector that can be used not only to produce complex proteins but also to co-express polypeptides to improve the quality of the end product (e.g. glycosylation modifying enzymes), the cell viability (e.g. antiapoptotic factors), the cell metabolism (e.g. mTOR (48)), or the secretion efficiency (e.g. XBP-1 (49)). Loss of productivity during long-term culture is a common phenomenon in recombinant CHO

cells generated with techniques based on random integration and it is attributed both to gene copy number loss and epigenetic silencing effects (36). The use of BACs successfully stabilized long-term protein expression in the absence of selection pressure in all of the tested target proteins. This suggests that BAC^{Rosa26} provides a chromatin environment protecting against transgene silencing. On the other hand, CHO cells are known to exhibit genotypic drift during long-term culture which may result in transgene rearrangement and loss of expression (50). However, we did not observe changes in BAC integrated copy numbers during the long-term culture, suggesting that integrated BAC vectors are genetically stable.

Besides all the advantages described above there are certain points to consider when using the BAC-system. Firstly, the BAC size: large genomic sequences present in the BAC serve as permissive environment for the transgene; however, the large size of the BACs is a disadvantage during cell transfection. DNA size affects the transfection efficiency and the uptake of large DNA molecules may be compromised in some cell types. Secondly, possible BAC toxicity: the open chromatin region present in the BAC has to be chosen carefully. BAC vectors derived from genomic libraries may contain several loci and the over-expression of some of these extra genes can be toxic to the cells. Thus, for certain BACs high copy number integration may be not supported by the host cells. In this regard, we observed that the BAC^{Rosa26} constructs integrated with a mean of 5.2 relative copies in stable cell pools (ranging from 2.4 to 8.8) while the BAC^{Hprt} stable pools harbored only 1.9 relative copies on average (ranging from 0.9 to 3.2) regardless of the heterologous promoter used. In addition, the narrow distribution of IgG-Fc productivity in Caggs:BAC^{Hprt} single cell clones compared to the Caggs:BAC^{Rosa26}-derived clones suggests as well that there is a limiting factor for production, presumably the number of vector copies integrated in the cell host. Thirdly, heterogeneous expression: while the BAC-system can help to boost overall protein production, it is very important to match the expression technique to the study needs. The use of BAC-based vectors generates a heterogenic stable pool population with cells harboring different integrated copy numbers and consequently different expression levels. Therefore, if the aim is to study the characteristics, activity or functionality of a protein, targeted integration to an active genomic site is the appropriate method to choose, since it can ensure equal levels of expression in different cell clones. This can be done for example by RMCE technology (22) or by using the CRISPR/Cas9 system (51).

The Caggs:BAC^{Rosa26} construct confers a reasonable (approximately 5 pcd/gcn) specific productivity per relative copy. Based on the RMCE technology others have identified genomic hot-spots in CHO cells supporting a productivity of up to 13 pcd per single transgene integration (23). This indicates that there can be more efficient hot-spots in the CHO genome than the ones used in our study. A future development of the BAC expression technology would be to perform a genomic screening, searching for hot-spots allowing high expression levels in CHO cells (or to use previously identified ones), cloning such genomic regions into BAC vectors and use them as BAC-based expression vectors. This future generation of BAC vectors should be useful tools for recombinant protein production.

In summary, BAC-based vectors with an optimal chromatin environment are valuable tools for recombinant protein production in mammalian cells, since they confer position independent, copy number dependent and high and stable expression. Moreover, high producing stable cell pools and single cell clones can be obtained without the need of transgene amplification in three and six weeks, respectively. Therefore, BAC-based expression vectors provide a fast and efficient alternative to conventional cell line-based production.

ACKNOWLEDGEMENTS

We thank L. Mach for kindly providing us with gp120 (ZM109). Furthermore, we like to thank M. Chang and Polymun QC for product analysis.

FUNDING

Austrian Research Promotion Agency (FFG) [FFG BRIDGE 829611 to E.C., K.Z. and E.B.]; Austrian Science Fund (FWF) [Project W1224 to W.S., P 25599 to E.C.]; Polymun Scientific GmbH, Klosterneuburg, Austria and the University of Natural Resources and Life Sciences, Vienna. Funding for open access charge: Austrian Science Fund (FWF) Project W1224.

Conflict of interest statement. AB and EC hold a patent on the BAC-expression system.

REFERENCES

- Zhu, J. (2012) Mammalian cell protein expression for biopharmaceutical production. *Biotechnol. Adv.*, **30**, 1158–1170.
- Walsh, G. (2014) Biopharmaceutical benchmarks 2014. *Nat. Biotechnol.*, **32**, 992–1000.
- Baldi, L., Hacker, D.L., Adam, M. and Wurm, F.M. (2007) Recombinant protein production by large-scale transient gene expression in mammalian cells: state of the art and future perspectives. *Biotechnol. Lett.*, **29**, 677–684.
- Wurm, F. and Bernard, A. (1999) Large-scale transient expression in mammalian cells for recombinant protein production. *Curr. Opin. Biotechnol.*, **10**, 156–159.
- Geisse, S. and Fux, C. (2009) Recombinant protein production by transient gene transfer into Mammalian cells. *Methods Enzymol.*, **463**, 223–238.
- Hacker, D.L., De Jesus, M. and Wurm, F.M. (2009) 25 years of recombinant proteins from reactor-grown cells - where do we go from here? *Biotechnol. Adv.*, **27**, 1023–1027.
- Xu, X., Nagarajan, H., Lewis, N.E., Pan, S., Cai, Z., Liu, X., Chen, W., Xie, M., Wang, W., Hammond, S. *et al.* (2011) The genomic sequence of the Chinese hamster ovary (CHO)-K1 cell line. *Nat. Biotechnol.*, **29**, 735–741.
- Kildegaard, H.F., Baycin-Hizal, D., Lewis, N.E. and Betenbaugh, M.J. (2013) The emerging CHO systems biology era: harnessing the 'omics revolution for biotechnology. *Curr. Opin. Biotechnol.*, **24**, 1102–1107.
- Datta, P., Linhardt, R.J. and Sharfstein, S.T. (2013) An 'omics approach towards CHO cell engineering. *Biotechnol. Bioeng.*, **110**, 1255–1271.
- Dietmair, S., Nielsen, L.K. and Timmins, N.E. (2012) Mammalian cells as biopharmaceutical production hosts in the age of omics. *Biotechnol. J.*, **7**, 75–89.
- Kim, J.Y., Kim, Y.-G. and Lee, G.M. (2012) CHO cells in biotechnology for production of recombinant proteins: current state and further potential. *Appl. Microbiol. Biotechnol.*, **93**, 917–930.
- Wilson, C., Bellen, H.J. and Walter, G. (1990) Position effects on eukaryotic gene expression. *Annu. Rev. Cell Biol.*, **6**, 679–714.
- Bandaranayake, A.D. and Almo, S.C. (2014) Recent advances in mammalian protein production. *FEBS Lett.*, **588**, 253–260.
- Wurm, F.M. (2004) Production of recombinant protein therapeutics in cultivated mammalian cells. *Nat. Biotechnol.*, **22**, 1393–1398.
- Jun, S.C., Kim, M.S., Hong, H.J. and Lee, G.M. (2006) Limitations to the development of humanized antibody producing Chinese hamster ovary cells using glutamine synthetase-mediated gene amplification. *Biotechnol. Prog.*, **22**, 770–780.
- Antoniou, M., Harland, L., Mustoe, T., Williams, S., Holdstock, J., Yague, E., Mulcahy, T., Griffiths, M., Edwards, S., Ioannou, P.A. *et al.* (2003) Transgenes encompassing dual-promoter CpG islands from the human TBP and HNRPA2B1 loci are resistant to heterochromatin-mediated silencing. *Genomics*, **82**, 269–279.
- Benton, T., Chen, T., McEntee, M., Fox, B., King, D., Crombie, R., Thomas, T.C. and Bebbington, C. (2002) The use of UCOE vectors in combination with a preadapted serum free, suspension cell line allows

- for rapid production of large quantities of protein. *Cytotechnology*, **38**, 43–46.
18. Ye, J., Alvin, K., Latif, H., Hsu, A., Parikh, V., Whitmer, T., Tellers, M., de la Cruz Edmonds, M.C., Ly, J., Salmon, P. *et al.* (2010) Rapid protein production using CHO stable transfection pools. *Biotechnol. Prog.*, **26**, 1431–1437.
 19. Zahn-Zabal, M., Kobr, M., Girod, P.-A., Imhof, M., Chatellard, P., de Jesus, M., Wurm, F. and Mermod, N. (2001) Development of stable cell lines for production or regulated expression using matrix attachment regions. *J. Biotechnol.*, **87**, 29–42.
 20. Girod, P., Nguyen, D., Calabrese, D., Puttini, S., Martinet, D., Regamey, A., Saugy, D., Beckmann, J.S., Bucher, P. and Mermod, N. (2007) Genome-wide prediction of matrix attachment regions that increase gene expression in mammalian cells. *Nat. Methods*, **4**, 747–753.
 21. Kwaks, T.H.J., Barnett, P., Hemrika, W., Siersma, T., Sewalt, R.G. a B., Satiijn, D.P.E., Brons, J.F., van Blokland, R., Kwakman, P., Kruckeberg, A.L. *et al.* (2003) Identification of anti-repressor elements that confer high and stable protein production in mammalian cells. *Nat. Biotechnol.*, **21**, 553–558.
 22. Wirth, D., Gama-Norton, L., Riemer, P., Sandhu, U., Schucht, R. and Hauser, H. (2007) Road to precision: recombinase-based targeting technologies for genome engineering. *Curr. Opin. Biotechnol.*, **18**, 411–419.
 23. Crawford, Y., Zhou, M., Hu, Z., Joly, J., Snedecor, B., Shen, A. and Gao, A. (2013) Fast identification of reliable hosts for targeted cell line development from a limited-genome screening using combined φ C31 integrase and CRE-Lox technologies. *Biotechnol. Prog.*, **29**, 1307–1315.
 24. Marshall, V.M., Allison, J., Templeton, T. and Foote, S.J. (2004) Generation of BAC transgenic mice. *Methods Mol. Biol.*, **256**, 159–182.
 25. Van Keuren, M.L., Gavrilina, G.B., Filipiak, W.E., Zeidler, M.G. and Saunders, T.L. (2009) Generating transgenic mice from bacterial artificial chromosomes: transgenesis efficiency, integration and expression outcomes. *Transgenic Res.*, **18**, 769–785.
 26. Yang, X.W. and Gong, S. (2005) An overview on the generation of BAC transgenic mice for neuroscience research. *Curr. Protoc. Neurosci.*, Chapter 5, Unit 5.20.
 27. Muylers, J.P.P., Zhang, Y., Testa, G. and Stewart, A.F. (1999) Rapid modification of bacterial artificial chromosomes by ET-recombination. *Nucleic Acids Res.*, **27**, 1555–1557.
 28. Zhang, Y., Buchholz, F., Muylers, J.P. and Stewart, A.F. (1998) A new logic for DNA engineering using recombination in *Escherichia coli*. *Nat. Genet.*, **20**, 123–128.
 29. Livak, K.J. and Schmittgen, T.D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods*, **25**, 402–408.
 30. Soriano, P. (1999) Generalized lacZ expression with the ROSA26 Cre reporter strain. *Nat. Genet.*, **21**, 70–71.
 31. Zambrowicz, B.P., Imamoto, A., Fiering, S., Herzenberg, L.A., Kerr, W.G. and Soriano, P. (1997) Disruption of overlapping transcripts in the ROSA beta geo 26 gene trap strain leads to widespread expression of beta-galactosidase in mouse embryos and hematopoietic cells. *Proc. Natl. Acad. Sci. U.S.A.*, **94**, 3789–3794.
 32. Bronson, S.K., Plaehn, E.G., Kluckman, K.D. and Maeda, N. (1996) Single-copy transgenic mice with chosen-site integration. *Proc. Natl. Acad. Sci. U.S.A.*, **93**, 9067–9072.
 33. Hatada, S., Kuziel, W., Smithies, O. and Maeda, N. (1999) The influence of chromosomal location on the expression of two transgenes in mice. *J. Biol. Chem.*, **274**, 948–955.
 34. Urlaub, G. and Chasin, L.A. (1980) Isolation of Chinese hamster cell mutants deficient in dihydrofolate reductase activity. *Proc. Natl. Acad. Sci. U.S.A.*, **77**, 4216–4220.
 35. Wirth, M., Bode, J., Zettlmeissl, G. and Hauser, H. (1988) Isolation of overproducing recombinant mammalian cell lines by a fast and simple selection procedure. *Gene*, **73**, 419–426.
 36. Chusainow, J., Yang, Y.S., Yeo, J.H.M., Toh, P.C., Asvadi, P., Wong, N.S.C. and Yap, M.G.S. (2009) A study of monoclonal antibody-producing CHO cell lines: what makes a stable high producer? *Biotechnol. Bioeng.*, **102**, 1182–1196.
 37. Su, L., Graf, M., Zhang, Y., Von Briesen, H., Xing, H., Köstler, J., Melzl, H. and Wolf, H. (2000) Characterization of a virtually full-length human immunodeficiency virus type 1 genome of a prevalent intersubtype (C / B') recombinant strain in china characterization of a virtually full-length human immunodeficiency virus type 1 genome of a prevalent. *J. Virol.*, **74**, 11367–11376.
 38. Pabst, M., Chang, M., Stadlmann, J. and Altmann, F. (2012) Glycan profiles of the 27 N-glycosylation sites of the HIV envelope protein CN54gp140. *Biol. Chem.*, **393**, 719–730.
 39. Jeffs, S.A., Goriup, S., Keble, B., Crane, D., Bolgiano, B., Sattentau, Q., Jones, S. and Holmes, H. (2004) Expression and characterisation of recombinant oligomeric envelope glycoproteins derived from primary isolates of HIV-1. *Vaccine*, **22**, 1032–1046.
 40. Zhang, C.W., Chishti, Y., Hussey, R.E. and Reinherz, E.L. (2001) Expression, purification, and characterization of recombinant HIV gp140. The gp41 ectodomain of HIV or simian immunodeficiency virus is sufficient to maintain the retroviral envelope glycoprotein as a trimer. *J. Biol. Chem.*, **276**, 39577–39585.
 41. McLellan, J.S., Pancera, M., Carrico, C., Gorman, J., Julien, J.-P., Khayat, R., Louder, R., Pejchal, R., Sastry, M., Dai, K. *et al.* (2011) Structure of HIV-1 gp120 V1/V2 domain with broadly neutralizing antibody PG9. *Nature*, **480**, 336–343.
 42. Nehlsen, K., Schucht, R., da Gama-Norton, L., Krömer, W., Baer, A., Cayli, A., Hauser, H. and Wirth, D. (2009) Recombinant protein expression by targeting pre-selected chromosomal loci. *BMC Biotechnol.*, **9**, 100.
 43. Qin, J.Y., Zhang, L., Clift, K.L., Huler, I., Xiang, A.P., Ren, B.-Z. and Lahn, B.T. (2010) Systematic comparison of constitutive promoters and the doxycycline-inducible promoter. *PLoS One*, **5**, e10611.
 44. Blaas, L., Musteanu, M., Eferl, R., Bauer, A. and Casanova, E. (2009) Bacterial artificial chromosomes improve recombinant protein production in mammalian cells. *BMC Biotechnol.*, **9**, 3.
 45. Mader, A., Prewein, B., Zboray, K., Casanova, E. and Kunert, R. (2013) Exploration of BAC versus plasmid expression vectors in recombinant CHO cells. *Appl. Microbiol. Biotechnol.*, **97**, 4049–4054.
 46. Eszterhas, S.K., Bouhassira, E.E., Martin, D.I.K. and Fiering, S. (2002) Transcriptional interference by independently regulated genes occurs in any relative arrangement of the genes and is influenced by chromosomal integration position. *Mol. Cell. Biol.*, **22**, 469–479.
 47. Schlatter, S., Stansfield, S., Dinnis, D., Racher, A., Birch, J. and James, D. (2005) On the optimal ratio of heavy to light chain genes for efficient recombinant antibody production by CHO cells. *Biotechnol. Prog.*, **21**, 122–133.
 48. Dreesen, I.A.J. and Fussenegger, M. (2011) Ectopic expression of human mTOR increases viability, robustness, cell size, proliferation, and antibody production of chinese hamster ovary cells. *Biotechnol. Bioeng.*, **108**, 853–866.
 49. Ku, S.C.Y., Ng, D.T.W., Yap, M.G.S. and Chao, S.-H. (2008) Effects of overexpression of X-box binding protein 1 on recombinant protein production in Chinese hamster ovary and NS0 myeloma cells. *Biotechnol. Bioeng.*, **99**, 155–164.
 50. Cao, Y., Kimura, S., Itoi, T., Honda, K., Ohtake, H. and Omasa, T. (2012) Construction of BAC-based physical map and analysis of chromosome rearrangement in Chinese hamster ovary cell lines. *Biotechnol. Bioeng.*, **109**, 1357–1367.
 51. Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A. *et al.* (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science*, **339**, 819–123.