Received: 3 March 2021 | Revised: 16 April 2021 | Accepted: 19 April 2021

DOI: 10.1002/btpr.3158

RESEARCH ARTICLE

BIOTECHNOLOGY PROGRESS

Genetic rearrangement during site specific integration event facilitates cell line development of a bispecific molecule

Barbara Tevelev[1] | Himakshi Patel[2] | Kathleen Shields[2] | Wei Wei[1] | Cecilia Cooley[1] | Sam Zhang[1] | Gabrielle Bitzas[3] | Weili Duan[3] | Lam Khetemenee[3] | Ryan Jackobek[3] | Aaron D'Antona[3] | Annette Sievers[3] | Amy King[3] | Amy Tam[3] | Yan Zhang[3] | Eric Sousa[3] | Justin Cohen[3] | Lila Wroblewska[3] | Jeffrey Marshall[2] | Martha Jackson[2] | John J. Scarcelli[1]

[1]Cell Line Development, Biotherapeutics Pharmaceutical Sciences, Pfizer Inc., Andover, Massachusetts, USA

[2]Analytical Research and Development, Biotherapeutics Pharmaceutical Sciences, Pfizer Inc., Andover, Massachusetts, USA

[3]BioMedicine Design, Pfizer Inc., Andover, Massachusetts, USA

Correspondence
John J. Scarcelli, Cell Line Development, Biotherapeutics Pharmaceutical Sciences, Pfizer Inc., Andover, MA 01810, USA.
Email: john.scarcelli@pfizer.com

Abstract

Site specific integration (SSI) expression systems offer robust means of generating highly productive and stable cell lines for traditional monoclonal antibodies. As complex modalities such as antibody-like molecules comprised of greater than two peptides become more prevalent, greater emphasis needs to be placed on the ability to produce appreciable quantities of the correct product of interest (POI). The ability to screen several transcript stoichiometries could play a large role in ensuring high amounts of the correct POI. Here we illustrate implementation of an SSI expression system with a single site of integration for development and production of a multi-chain, bi-specific molecule. A SSI vector with a single copy of all of the genes of interest was initially selected for stable Chinese hamster ovary transfection. While the resulting transfection pools generated low levels of the desired heterodimer, utilizing an intensive clone screen strategy, we were able to identify clones having significantly higher levels of POI. In-depth genotypic characterization of clones having the desirable phenotype revealed that a duplication of the light chain within the landing pad was responsible for producing the intended molecule. Retrospective transfection pool analysis using a vector configuration mimicking the transgene configuration found in the clones, as well as other vector configurations, yielded more favorable results with respect to % POI. Overall, the study demonstrated that despite the theoretical static nature of the SSI expression system, enough heterogeneity existed to yield clones having significantly different transgene phenotypes/genotypes and support production of a complex multi-chain molecule.

KEYWORDS

bioprocess development, CHO cells, gene expression system, multi-specific molecule, site-specific integration

1 | INTRODUCTION

Chinese hamster ovary (CHO) cells are the preferred host platform for development and production of biotherapeutic proteins.[1,2] CHO cells are the most dependable host for the industrial production of pharmaceutical proteins due to their ability to correctly fold, assemble and post-translationally modify recombinant proteins, all of which are needed for desirable pharmacokinetic and pharmacodynamic properties.[3–5] Additionally, CHO cells have attributes such as fast growth, high cell density, high expression, and adaptability to serum-free chemically-defined media[6] that make them well-suited for biopharmaceutical production.

The two most common strategies employed for CHO cell biotherapeutic transgene integration are random integration (RI) and site specific (or targeted) integration (SSI). RI is the most commonly employed strategy. The RI system is capable of generating highly productive cell lines, but requires a great deal of effort to screen hundreds to thousands of clones in order to identify commercially viable cell lines.[7] Since it is not possible to predict the orientation, location or copy number of the recombinant transgene integration, the RI system introduces variability that makes it challenging to select clones with predictable and stable phenotypic and genotypic profiles.

An alternative to the RI system is the SSI approach. One method of integrating transgenes into specific loci is via CRISPR/Cas9 technology.[8,9] This method has been applied for insertion of small DNA fragments and long target genes in CHO cells to construct recombinant industrial cell lines.[9–11] Another method of site-specific transgene integration utilizes recombinase mediated cassette exchange (RMCE).[12,13] This approach is facilitated by a CHO host engineered with a landing pad harboring counter-selection markers flanked by mutually-exclusive recombinase sites, ideally at a position within the genome known to allow for high-level and stable gene expression. There are several advantages for an RMCE SSI system: predictability that ensures phenotypic and genotypic stability, efficiency that allows for a less resource intensive clone screening process, and comparability between material generated from transfection pools and clones, which enables harmonization of expression systems between research and development groups.[14–16] A potential drawback of this system is relatively low productivity, as it may rely on expression of only a single copy of the transgene(s). Recent improvements in our SSI platform alleviates this concern, which have led to a second generation SSI expression system (with CHOK1SV GS-KO SSI host) which routinely yields clones with monoclonal antibody titers equal to or better than that of similarly developed RI cell lines.[17]

Multi-specific modalities are becoming increasingly popular in the biopharmaceutical industry.[18] These modalities represent a plethora of different strategies and mechanism of actions that can be utilized to increase molecule efficacy and specificity, reduce toxicity, and optimize clearance.[19–21] These advantages are achieved through the ability to simultaneously bind multiple targets, an attribute facilitated by multiple unique heavy and/or light chains (LCs), which allows for novel therapeutic potential.

Multi-specific molecules can be challenging to produce, since the product of interest (POI) can often be contaminated with unwanted product-related impurities such as homodimeric forms of a desired heterodimeric bispecific molecule. Efforts to optimize heterodimeric product output and quality may come at the sacrifice of development speed and potential product titer. One commonly employed strategy to create stable bispecific molecules is to separately develop, express, and partially purify two different antibodies having charge-based point mutations which facilitate heterodimerization when mixed under appropriate redox conditions.[21,22] This approach is time-consuming and resource intensive as it requires development of multiple cell lines, production processes, and purification streams. Another strategy for the generation of multi-specific molecules involves a single production cell line approach based upon point mutations within CH3 domains that result in different heavy chains (HCs) that are sterically complimentary, thus promoting heterodimerization.[21,23] This approach is also challenging as it is heavily dependent upon the cell line achieving optimal expression ratios of the various transgenes in order to produce high overall titer and % POI. While the nature of the RI cell line generation approach lends itself quite well to screening for various transgene expression ratios, the SSI approach could be limited in this regard, where both integration locus and transgene copy number/expression levels should theoretically be static. Up-front analytics capable of detecting process-related impurities typically are used as aids in clone screening, however the benefit of these tools can be called into question for an SSI cell line approach, as clones should theoretically not exhibit dramatic differences in product quality attributes related to peptide mis-pairing.

The purpose of this work was to determine the suitability of an SSI cell line for the production of a bispecific molecule composed of three unique polypeptide chains. Our finding of an unexpected gene duplication event required for high % POI expression may have implications for future expression vector design enhancements to optimize SSI expression of multi-chain, multi-specific modalities.

2 | METHODS AND REAGENTS

2.1 | Reagents

Chemicals were purchased from Sigma-Aldrich (St. Louis, MO) and cell culture reagents were sourced from Gibco/ThermoFisher (Waltham, MA). CD-CHO media was purchased from Thermo Fisher Scientific Inc. (Waltham, MA). DNA size markers, enzymes and PCR reagents were purchased from NEB (Ipswich, MA). DNA extraction kits were purchased from Qiagen (Germantown, MD). Oligonucleotide primers and probes were commercially synthesized by IDT (Coralville, IA).

2.2 | Vector construction

An SSI 2.0 vector backbone was used to construct the different expression vectors for the bi-specific 045 molecule. The vector contains designed cloning sites for insertion of two genes of interest

(GOI) in a single reaction. Two gene fragments are prepared by PCR and combined with two pre-digested vector fragments through an isothermal assembly reaction. Three- and four-chain vectors were constructed in two steps. In step one, intermediate single or two-chain vectors were generated. In step two, fragments spanning the genes of interest as well as appropriate regulatory elements (e.g., promoters) were amplified from the intermediate vectors and cloned into an empty SSI2.0 backbone using isothermal assembly. All of the expression vectors were confirmed by full vector sequencing and restriction digest diagnostics.

2.3 | Cell culture and generation of pools

A total of 50 μg of targeting and FLPe expression vectors pDNA was re-suspended in 100 μl TE pH 8.0. Three replicate reactions were carried out for each pool and combined for total volume of 300 μl. On the day of transfection, 12×10^6 cells were pelleted and re-suspended in 0.8 ml CD-CHO with 6 mM L-glutamine. The cell suspension was transfected with the vector mixture via electroporation. Post-electroporation the contents of the cuvette were transferred to 20 ml of CD-CHO with 6 mM L-glutamine and transferred to a fresh T75 flask. Cultures were maintained at 36.5°C, 5% (v/v) CO_2 in air using a humidified incubator. Twenty-four hours after transfection the cells were pelleted in a 50 ml falcon tube by centrifugation at 150x g for 8 min (at ambient temperature). The pellet was washed in CD-CHO before re-suspending in 20 ml CD-CHO and transferred to a fresh T75 flask. The cells were then cultured at 36.5°C, 5% (v/v) CO_2 in air using a humidified incubator. Once each culture in a T75 flask reached acceptable viability, it was transferred into suspension culture in a 250 ml vented shake-flask containing CD-CHO medium such that the total culture volume was 50 ml. The recovered transfection pools were cryopreserved and later thaw for purposes of cloning.

2.4 | Evaluation of pools in 1 L shake-flasks

Pools were cultured in one-liter shake-flasks under a 12-day fed-batch production process using Pfizer's proprietary chemically defined medium and feeds and maintained at 36.5°C, 5% (v/v) CO_2 in air using a humidified incubator. Counts and feeds were completed daily. The HC fusion (HCF)-LC-HC pools were passaged in Corning shake flasks (431147) while in the experiment assessing different 2xLC vector configurations, pools were passaged in ThermoFisher shake flasks (4113–1000). Daily viabilities and viable cell concentrations were determined using a Vi-CELL™ automated cell viability analyzer. Antibody concentration in the medium was determined by Protein A HPLC.

2.5 | Evaluation of clones in ambr®15

Clones were cultured in ambr®15 mini-bioreactors under a 12-day fed batch production process using Pfizer's proprietary chemically defined medium and feeds, including appropriate control of other parameters such as dissolved oxygen, pH, and temperature. Cell culture performance was analyzed as previously.[17]

2.6 | FACS cloning

For the vector HCF-LC-HC, single-cell cloning was performed on a FACSAria™ cell sorter equipped with FACSDiva v6.0 software with an air-cooled laser emitting at 488 nm. Cells of interest were identified in a forward scatter (FSC) versus side scatter (SSC) dot plot (gate one). Both FSC and SSC gating was used to identify cells of interest based on size and granularity (complexity). The FSC parameters indicated cell size whereas SSC parameters provide insight to the internal complexity or granularity of the cell. The population of interest was gated based on the combination of these two measured parameters. Further gating was performed to discriminate for doublets by plotting the heights versus width for both FSC and SSC (gate two and three).

The sorting for single-cell cloning targeted GFP-negative cells (i.e., expressing molecule 045 integrated within the host landing pad). A positive control (GFP-negative) from another SSI project and a host as negative control (GFP-positive) were utilized. The GFP sort analysis was based on doublet discrimination gating as well as the gating for the applicable cell population. As per manufacturer recommendations, gates one through three are the gates typically used for doublet discrimination. Gate four was established based on the GFP-negative and GPF-positive controls. Only cells present in gate four were sorted. Selection of single cells using this host has been describe previously.[17]

2.7 | Scale-up process and selection of cell lines

After cloning, the cells were left to recover in 96-well plates containing CD-CHO with 1 g/L cellastim (Invitria). Once the well's confluence reached ~50%, the static culture was moved to 48-well plates (containing 250 μl of CD-CHO), and then consequently to 24-well (containing 500 μl of CD-CHO) and finally to six-well plates (containing 3 ml of CD-CHO).

2.8 | Genomic DNA purification

Genomic DNA (gDNA) was isolated using Blood & Cell Culture DNA Mini Kit from QIAGEN as per the manufacturer's instructions. The gDNA was then eluted in water and quantified via NanoDrop (Thermo Fisher).

2.9 | Southern blot analysis

For Southern blot analysis, gDNA was isolated from end of production (EOP) cells collected at the harvest day of ambr®15 12-day

production process. Five μg for each sample were digested with either *EcoRI* or *HindIII*. The digested DNA samples were size-fractionated on 1% agarose gels, transferred to positively charged nylon membranes and cross linked by UV irradiation. Membranes were pre-hybridized in a formamide-based hybridization buffer at 45°C. They were subsequently hybridized in the same solution with a target-specific denatured [32P]-labeled probe at 45°C.

Probes for target genes (HC, LC, and fusion element) was designed using PCR primers flanking the full-length sequence. The probes were end-labeled with [γ-^{32}P] dATP (222 TBq/mmol, Perkin Elmer), followed by removal of free nucleotides using the Qiagen nucleotide removal kit. The labeled probes were used in the hybridization reaction to assess the restriction fragments for each sample. Hybridization occurred overnight at 55°C, followed by three washes with varying stringency of SSC and SDS at 45°C and room temperature to remove unbound γ-^{32}P. Washed membranes were exposed to phosphoimager screen and scanned for signal on Typhoon FLA 9000 Gel Imaging Scanner (GE Amersham).

2.10 | Determination of gene copy number

Copy number of HC, LC, and fusion-only genes per cell was quantified by digital droplet PCR (ddPCR) using the QX200 system (Bio-Rad Laboratories). Genomic DNA was purified from either passaging day three or ambr®15 end of production (EOP) samples. Approximately 1 μg of gDNA was digested with *CviQ1* then diluted 10-fold. Each sample dilution was run in triplicate. Specific primer/probe sets were used to determine the gene copy number for each target gene. A specific primer/probe set for a housekeeping gene, Cog1, which was previously shown to have two copies in the SSI host genome, was used as a normalizer. Along with samples, un-transfected host and no-template-control were used as negative controls.

2.11 | 3′ and 5′ on-target PCR

Forward and reverse primers were designed for end-point PCR to test 5′ and 3′ integration of the vector HCF-LC-HC within the host cell landing pad. To confirm 5′ integration, the forward primer was designed within the genomic sequence while the reverse primers was designed within the GS region. To confirm 3′ integration, the forward primer was designed downstream of the HC sequence and the reverse primer was designed in the genomic sequence. The Q5 polymerase and standard reaction and thermal cycling conditions as per manufacturing instructions (NEB; Ipswich, MA) were used to amplify the expected 1.4 kb bands. PCR products were assessed using agarose (1%) gel electrophoresis and ethidium bromide staining. A previously tested sample from a cell line expressing another molecule with confirmed genotype was used a positive control and un-transfected host was used as a negative control.

2.12 | Target capture sequencing

Targeted sequence capture analysis was performed. Pre-capture library was prepared using target sequence capture probes that were designed in the regions outside of the genes of interest but within the SSI 2.0 landing pad. Post-capture amplification and pac-bio library was prepared. Sequence analysis and data mining was performed, followed by re-analysis and generation of the revised model.

2.13 | Analytical tools for product quality evaluation

2.13.1 | Enzyme-linked immunosorbent assay

The antigen specific for the Fab domain of bi-specific molecule 045 was coated onto a 96 well enzyme-linked immunosorbent assay (ELISA) plate. After coating, the plate was blocked with reagent to prevent non-specific binding. In a separate polypropylene plate, conditioned media from clones were diluted in assay buffer and the sample dilutions were then transferred onto the wells of the assay plate. After incubation, the plate was washed to remove unbound 045. Next, biotin labeled antibody that specifically binds to the fusion element of the HCF was added and after incubation the plate was washed to remove unbound antibody. Bound biotinylated antibody was detected using streptavidin-HRP followed by a final wash and development with a solution of 3,3′,5,5′-tetramethyl benzidine substrate. The peroxidase reaction was stopped by the addition of sulfuric acid and the absorbance in each well was recorded at 450 nm. The concentration of bound bi-specific was determined by interpolation from a standard curve prepared using a purified bi-specific molecule 045.

Sodium dodecyl (lauryl) sulfate-polyacrylamide gel electrophoresis (SDS-page) and Western blot using near-infrared fluorescence detection was designed to detect the heterodimer by probing with a mouse antibody that is specific for the fusion element. The samples underwent protein denaturation followed by gel electrophoresis and Coomassie staining. Separated proteins were then transferred onto a polyvinylidene fluoride membrane and the membrane was blocked to prevent non-specific binding of antibodies. The membrane was then probed with a mouse antibody specific for the fusion element followed by detection with a IRDye goat anti-mouse secondary antibody. The western blot was imaged using the Li-Cor Odyssey® Imager.

For size exclusion HPLC (SEC), approximately 50 mg of each protein sample was injected onto a YMC-Pack Diol-200 size-exclusion column (300 × 8 mm, Waters Catalog number DL20S053008WT) maintained at 30°C. The high molecular mass species (HMMS), monomer, and LMMS were separated using isocratic elution with a mobile phase containing 20 mM sodium phosphate and 400 mM sodium chloride, pH 7.2.

Non-reduced capillary gel electrophoresis (nrCGE) analysis was used to assess transfection pools from 12-day fed-batch shake-flask process, 52 clones from a 4-day shake-flask passage cultures, and five

clones from the ambr®15 12-day production process. The analysis evaluated purity and heterogeneity of the material. Samples were denatured with SDS and heated with or without a reducing agent. The protein samples were reduced into fusion, heavy and LCs which were electrophoretically separated in a capillary containing sieving medium and detected using UV absorbance. The separation allows quantitation of the resolved chains as well as size related impurities.

The four-day shake-flask passage culture samples and 12-day fed-batch shake-flask pool samples were assessed on Caliper LifeSciences instrument (LabChip GXII) according to the manufacturer's instructions and the ambr®15 12-day production process harvest EOP clone samples were assessed on a Beckman Coulter instrument (PA800, PA800 Enhanced or PA800 plus).

For Caliper LifeSciences (PerkinElmer) instrument, HT Protein Express LabChip was prepared as per manufacturer's instructions. The protein concentration of sample was adjusted to 1.0 mg/ml with purified water. Approximately 5 μl of sample was mixed with 35 μl of denaturation buffer containing iodoacetamide (350 μl sample buffer +35 μl of 0.25 M Iodoacetamide (IAM, freshly prepared) in a 0.5 ml microfuge tube. The sample mixture was heated at 70°C for 10 min. After heating, samples were spun for 1 min at 3000 rpm and 70 μl of purified water was added. Samples were mixed thoroughly and spun at 3000 rpm for 5 min. Approximately 44 μl of each sample was transferred to a 96-well microtiter plate for analysis.

For Beckman Coulter instrument, IgG Purity/Heterogeneity Assay kit (PN A51967AB) was used as per manufacturer's instructions. Approximately 200 μg of sample (maximum volume of 90 μl) was mixed with 10 μl of 250 mM iodoacetamide and 100 μl of sample buffer. The total volume of sample was adjusted to 200 μl with purified water (MilliQ water) for a final protein concentration of 1.0 mg/ml.

Collection of electropherogram data and analysis was performed using Waters Empower Chromatography and Results Management software. For data analysis and purity assessment, all peaks excluding buffer related peaks in the electropherogram are integrated and relative percent time corrected area (% TCA) of each species (main, fragments, other, etc.) is reported. Relative % TCA of each species is calculated using following equation:

$$\%\text{TCA of Main Peak} = \frac{\text{Time Corrected Area (TCA) of Main Peak}}{\text{Total TCA}} \times 100.$$

3 | RESULTS AND DISCUSSION

3.1 | Evaluation of HCF-LC-HC transfection pools expressing the 045 bispecific in an SSI expression system

Molecule 045 is a bi-specific heterodimer, comprised of two identical LCs and two different HCs (Figure 1(a)). The variable domains of the two HC's are identical, with one chain being a typical mAb HC, while the other HC is fused to a receptor-recognizing element at its C-terminus and is referred to as HC fusion (HCF). The CH3 domains have been engineered to create either a "knob" or a "hole" to promote heterodimerization and facilitate the post-translational assembly of the molecule.[23]

Initial experiments performed using a transient expression system showed low % POI with high levels of HC homodimer composed of two non-fused HC, suggesting that expression of the HCF was limiting (data not shown). To mitigate this for stable expression, a vector was constructed which includes a single copy of each of the relevant chains in the following orientation: HCF, LC, HC (Figure 1(b)).

We evaluated the ability of the HCF-LC-HC vector to yield desirable % POI in a stable cell line expression system using a targeted integration approach with a single landing pad. Positive selection was achieved by culturing the CHO cells in medium which lacked glutamine. The HCF-LC-HC transfection pools' viability was monitored for several days. Approximately 15 days after transfection, the viability declined to 20% and subsequently recovered to >95% culture viability by day 22 (Figure 1(c)).

Once the HCF-LC-HC transfection pools reached >95% viability, they were evaluated in a 12-day shake-flask fed-batch production process (Figure 1(d)). Growth, viability, productivity, and product quality parameters were assessed. The HCF-LC-HC pools reached peak cell density on day 10 and then the growth decreased over the duration of the production process. The HCF-LC-HC pools productivity ranged from 180 to 350 mg/L as determined by the Protein A (ProA) titer results on harvest day (Figure 1(e)). It is noteworthy to add that the ProA titer analysis does not discriminate between properly formed heterodimeric molecules and improper HC-HC or HCF-HCF homodimeric molecules, or molecules that did not properly assemble (e.g.,: HC-LC half molecules). Following purification over a ProA affinity column, size exclusion chromatography (SEC) analysis showed that the average % POI for the HCF-LC-HC pools (A, B, C) was only 33%, with the remaining material consisting of HC-HC homodimers, half molecules, or higher order aggregates (Figure 1(f)).

3.2 | Evaluation of clonal cell lines derived from the HCF-LC-HC transfection pools

At this point, we rationalized that there were two possibilities accounting for HCF-LC-HC pool performance with respect to % POI. One explanation could be that every cell within the pool was producing approximately 33% of the intended heterodimeric product, dependent entirely upon the transcriptional and translational efficiency provided by the transfected vector. Another possibility was that each pool was comprised of cells producing several different % POI outcomes, the sum of which was reflected in overall pool performance. This possibility would be consistent with expectations for the RI approach, where high heterogeneity among clones is typically observed. If this were the case, we reasoned that we should be able to clone from the HCF-LC-HC transfection pools and find a range of % POI outcomes, including clones that produce % POI significantly

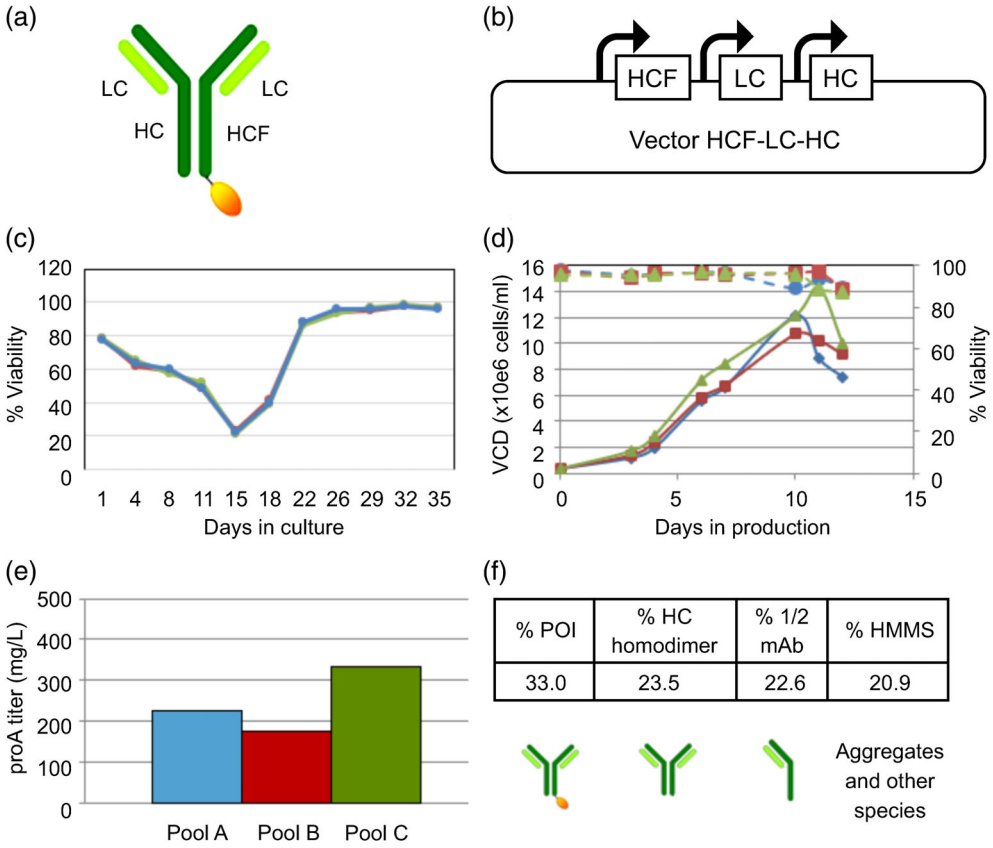


FIGURE 1 Bi-specific molecule 045 structure, configuration and pool performance. (a) Depiction of the structure of bi-specific molecule 045. LC-light chain, HC-heavy chain, HCF -heavy chain fused to receptor binding element. (b) Topology of vector HCF-LC-HC used for transfection, illustrating a single copy of each of the genes of interest. (c)Transfection/selection recovery of the HCF-LC-HC pools as represented by % viability across days in culture after transfection of the vector HCF-LC-HC into the SSI 2.0 host cell line. Blue: Pool A; Red: Pool B; Green: Pool C. (d) Viable cell density and % viability for HCF-LC-HC pools over days in a 12-day shake-flask fed-batch production process. (e) Harvest proA titer (mg/L) from a 12-day shake-flask fed-batch production process for HCF-LC-HC pools. (f) Assessment of high and low molecular mass species of proA purified harvest material from one of the representative transfection pools from a 12-day shake-flask fed-batch production process

greater than 33%. To this end, we carried out a single cell cloning and screening campaign to determine the % POI from clones derived from the transfection pools. An overview of the clone screening approach is shown in Figure 2.

For cloning, single non-fluorescent cells, (cells which were no longer harboring the GFP gene present in the SSI host cell landing pad) were deposited in 96-well plates using flow cytometry. Approximately 2 weeks post-cloning, over 300 clones were scaled-up and conditioned media from these clones was screened for expression of material bound to ProA (Figure 3(a)). As noted above, the ProA titer assay detected the desired heterodimeric molecule as well as product-related impurities including HC-HC and HCF-HCF homodimers. We found 240 clones that had ProA titer values in the range of 1–18 ug/ml. These clones were scaled up to 24-well plates. Five days post-seeding, conditioned media from clones was assessed via enzyme-linked immunosorbent assay (ELISA) as depicted in Figure S1. ELISA results were quantified using a standard curve of purified heterodimer (POI). The ELISA detects HC-HCF heterodimer and HCF homodimer but does not detect HC homodimers. Clones producing material which did not have a signal above background were

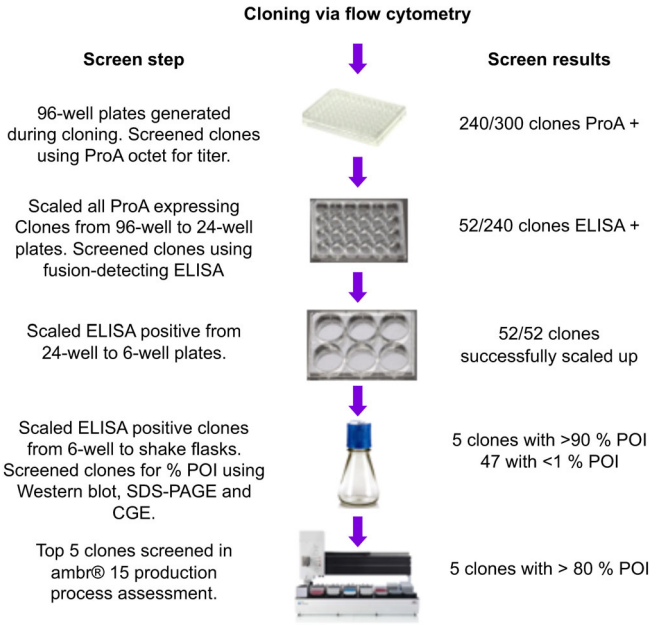


FIGURE 2 Clone screen for cells expressing molecule 045

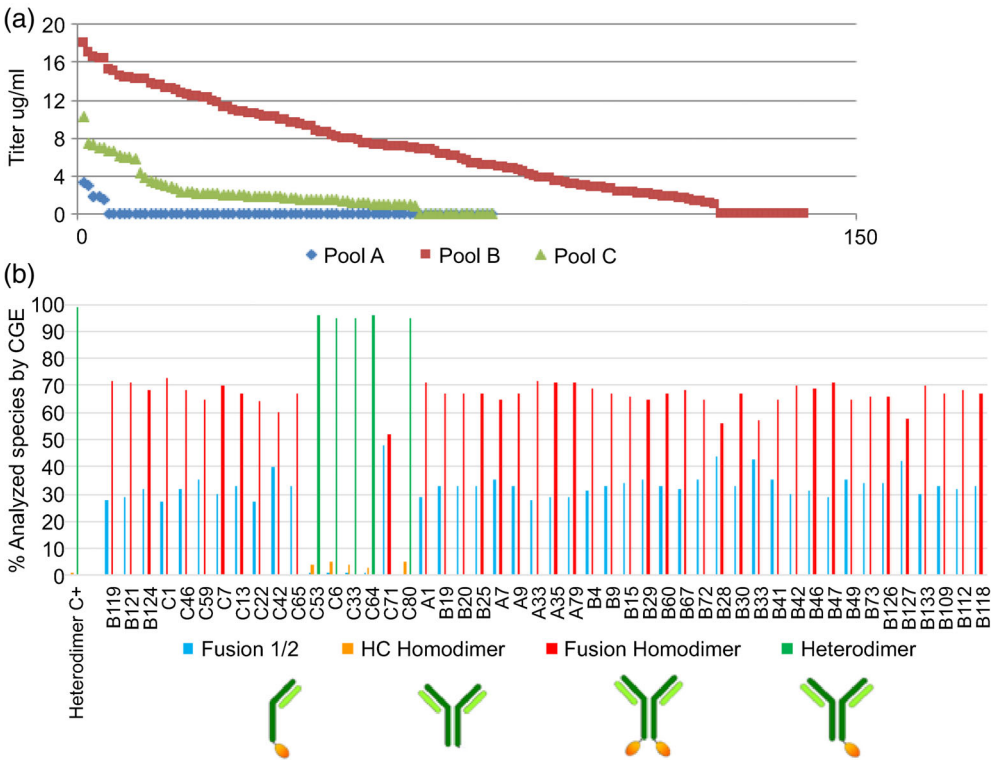


FIGURE 3 Evaluation of clones expressing molecule 045. (a) ProA titer screen at the 96-well stage, depicts titer across clones from the three HCF-LC-HC transfection pools: A, B, and C. Clones having less than 1 ug/ml were not progressed forward. (b) Non-reduced capillary gel electrophoresis (nrCGE) results for 52 ELISA-positive (+) clones as analyzed by the four-day shake-flask passage cultures

discarded (negative by ELISA). Since the fusion element of the HCF was used for detection, it was possible that the ELISA might bias selection of clones producing the HCF-HCF homodimer. However, we reasoned that it would be an efficient screening tool based on the evaluation of material expressed from the cellular transfection pools in which we did not observe appreciable levels of the HCF homodimer and in which most of the impurities consisted of HC homodimer and half molecule.

Approximate titer values by ELISA were divided by the proA titer values to yield a ratio used to further evaluate clones. Clones with ratios close to or greater than one were hypothesized to be making primarily heterodimer or HCF homodimer. Of the 240 clones analyzed, material from 52 clones were within the considered ratio range and these clones were expanded to 30 ml cultures in shake flasks. On day four, the condition media from the shake flasks was analyzed further. From transfection pool A, six clones moved forward with the ratio ranging from 0.6 to 0.7 and titer spread of 3–10 ug/ml; for transfection pool B, 30 clones with ratio of 0.9–1.7 and titer spread of 2–24 ug/ml; and for transfection pool C, 16 clones with ratio of 0.9–1.5 and titer range of 2–12 ug/ml (See Table S1).

After establishment of growth in shake flasks, material from the clones at day four of passaging was subjected to size/structure analysis by SDS PAGE, which allows for the discrimination of product related impurities from intact heterodimer based on size, followed by western blot analysis using fluorescently labeled receptor that specifically binds to HCF (data not shown). Western blot was used to assess for the presence or absence of the desired heterodimer by probing with antibody raised against the receptor binding element. Of the 52 clones, 47 did not display the desired heterodimer (POI). For these clones, the material from four-day shake-flask passage cultures contained predominantly HCF homodimer and HCF half molecule based on the banding pattern of fusion homodimer control as illustrated by representative clones in Figure S2a (A1, A35, B42, B108, C1, and C7). Of the 52 ELISA-positive clones, only five clones from transfection pool C (C6. C33, C53, C64, and C80) displayed the same size band as the positive heterodimer control and were producing material of intended quality (Figure S2a).

These five clones were further examined using SDS-PAGE to confirm the size of the intended protein and to ascertain presence of all possible impurities. For the five clones having a desirable phenotype, the results of the SDS-PAGE confirmed presence of the bi-specific 045 heterodimer (POI), as the band size resembled a positive control of the same predicted size (Figure S2b). The results also indicated the presence of impurities corresponding to different size species than those observed in the transfection pools: heterodimer with a single LC, HC-half molecule, HC, and LC at a relatively low level of the total material analyzed (Figure S2b).

We used nrCGE to better quantify the amount of % POI relative to the total material bound to ProA. The nrCGE results confirmed the western blot and the SDS-PAGE data, as only five clones from transfection pool C were identified containing the desired % POI at >90% (Figure 3 (b)). As indicated by the western blot, SDS-PAGE, and nrCGE results, the remaining clones from HCF-LC-HC pool A, B, and C displayed % POI at <1% and consisted mostly of HCF-HCF homodimers or HCF half-molecules (Figure 3(b)). Out of the 300 clones that were initially analyzed, only five clones from the targeted integration cell expression system could produce bispecific antibody molecules with high % POI.

To determine if the high % POI phenotype observed from material harvested from four-day shake-flask passage cultures would

remain consistent in a 12-day AMBR fed-batch process, clones with high POI (C6, C33, C53, C64, and C80) as well as several clones with low POI (A1, A35, B42, B108, C1, and C7) were selected for further assessment. These 11 selected clones were subjected to a 12-day AMBR fed-batch production process similar to the process previously used to analyze the transfection pools. For most clones, the viability was maintained above 80% throughout the process, with exception of C53 for which the viability dropped to 60% on day 12 (Figure 4(a)–(c)). On harvest day of the 12-day AMBR fed-batch process, the concentration of ProA-bound product secreted into culture medium from the five clones with high % POI ranged from ~2700 to 1400 mg/L with specific productivity (Qp) ranging from 10 to 17 pg/cell/day (Figure 4(d)). The remaining clones displayed very similar titer and Qp of ~1200 mg/L and 3 pg/cell/day, respectively. Interestingly, the VCD for the five clones which produced material of high % POI was generally lower than the clones with low % POI, but their Qp was significantly higher.

Material harvested from the 12-day AMBR fed-batch cultures were purified by ProA affinity chromatography and product quality was assessed by nrCGE. The results confirmed the material produced by clones C6, C33, C53, C64, and C80 consisted predominantly of the desired heterodimer while the remaining clones produced material <1% POI (Table 1). A slight decrease in % POI for C6, C33, C53, C64, and C80 was observed between results from four-day shake-flask passage cultures (Figure 3(b)) versus 12-day AMBR fed-batch production process harvest cultures (Table 1). This slight discrepancy in % POI could be attributed either to culturing conditions, or due to differences in the analytical assays used (high-throughput nrCGE vs. standard nrCGE). Assessment of clones producing low % POI (A1, A35, B42, B108, C1, and C7) in the 12-day AMBR fed-batch production process confirmed their inability to produce high % POI, as the results (Table 1) were similar to those observed in the initial screen performed on material from four-day shake-flask passage cultures (Figure 3(b)). The 12-day AMBR fed-batch product quality data confirmed the results of the clone screen and demonstrated the consistency and stability of the % POI phenotype between cell passaging conditions and fed-batch process conditions for the selected five clones. The qualitative expression patterns are consistent for all clones and conditions tested, suggesting that the ability of a given clone to secrete low or high % POI material is largely independent of specific culture conditions. Further, the high POI phenotype was found to be consistent over several months of culture passaging (Figure S3).

Out of 300 clones, the screening strategy employed identified five clones from the HCF-LC-HC transfection pools that produced significantly greater than the 30% POI observed for the parental transfection pools. The remaining clones were unable to produce material with the desired product quality demonstrating significant cellular heterogeneity in the transfected pools.

3.3 | Genotypic characterization of clonal cell lines derived from the HCF-LC-HC transfection pools

We hypothesized that the difference between clones capable of producing high and low % POI was related to transgene genotype. In-depth genotypic characterization was performed on the five clones capable of producing high % POI. Several other clones with low % POI were also partially evaluated to better understand how the undesirable phenotype correlated with genotype. An *in silico* prediction of the vector integration site topology is shown in Figure 5(a). Genomic DNA was extracted from cells harvested at the end of the 12-day AMBR fed-batch production process and subjected to several assays to determine integration topology.

Southern blotting was used to assess the internal topology of the landing pad. As shown in Figure 5(a), *Eco*RI digestion of genomic DNA from cells harboring the HCF-LC-HC vector payload was expected to yield fragments containing HCF, LC, and HC at approximately 8.5 kilobases (kb), 3.0 kb, and 6.2 kb, respectively. The *Hin*dIII restriction digest was expected to liberate a fragment containing the LC gene at approximately 5.5 kb in size. The Southern blot results for the restriction endonuclease *Eco*RI revealed the expected band for blots hybridized with fusion element and HC probes, suggesting proper integration of the payload within the landing pad for those transgenes at both 3′ and 5′ end (Figure 5(b),(c)). In addition, the Southern blot probed with LC for the *Eco*RI restriction digests indicated intact coding region of the LC downstream of the HCF and upstream of HC at the expected ~3.0 kb size marker (Figure 5(d)).

However, as shown in Figure 5(e), *Hin*dIII digestion of genomic DNA from cells harboring the HCF-LC-HC vector payload was expected to yield only one restriction fragment hybridizing to the LC gene probe at approximately 5.5 kb. Surprisingly, the *Hin*dIII Southern blot revealed two bands for five clones (C6, C33, C53, C64, and C80) capable of producing >85% POI: the expected band at ~5.5 kb and an additional unexpected band of ~3.0 kb. The topology for clones with poor % POI revealed different restriction fragment lengths when probing for LC; a single unexpected band at ~3.0 kb for clone A1, and two unexpected bands at ~3.0 and ~9.0 kb for clones B42 and B108 (Figure 5(e)).

To further analyze the transgene integrations, a PCR assay was designed to confirm correct 5′ and 3′ integration of the vector within the landing pad by generating amplicons that span the junction of vector and endogenous host DNA. Based on the *in silico* model, the PCR product for 5′ and 3′ integration assessment was predicted to result in an amplicon of 1.4 kilobases (kb) (Figure 6(a)). For both the 5′ and 3′ PCR assays, a single amplicon of the expected size was observed for clones C6, C33, C53, C64, and C80. (Figure 6(b),(c)). Taken together, the results demonstrated proper integration of the vector within the landing pad at both the 5′ and 3′ ends for the five clones with high % POI. For clones A1, A35, B42, B108, C1, and C7 the results revealed the expected amplicon for the confirmation of the 5′ integration site but no product was detected for the 3′ integration site, indicating a potential difference between the in silico predicted and actual genotype for these poor % POI clones downstream of the LC.

To assess transgene copy number, sequence specific primer-probe sets were designed and droplet digital polymerase chain reaction (ddPCR) assays were developed. The primer-probe set designed for HC detected both the HC and HCF. For the five clones producing high % POI, the ddPCR results for the HC were within expectation of

TABLE 1 Product quality analysis for selected cell lines

Cell line	LC %	1/2 HC %	1/2 fusion %	HC homodimer %	Fusion homodimer %	Heterodimer (POI) %
A1	1.13	0	45.21	0	56.09	0
A35	1.17	0	44.76	0	54.06	0
B42	3.22	0	49.85	0	46.93	0
B108	1.55	0	48.07	0	50.37	0
C1	0	0	45.00	0	55	0
C7	0	0	49.46	12.84	36.08	0
C6	0.54	3.11	0	13.94	0	82.41
C33	0.66	3.4	0	11.52	0	84.43
C53	0.93	2.39	0	7.96	0	88.73
C64	0.74	2.88	0	10.88	0	85.5
C80	0.69	3.92	0	12.24	0	83.15

Note: Protein integrity/fragmentation analysis of material collected from a 12-day AMBR fed-batch harvest using nrCGE.

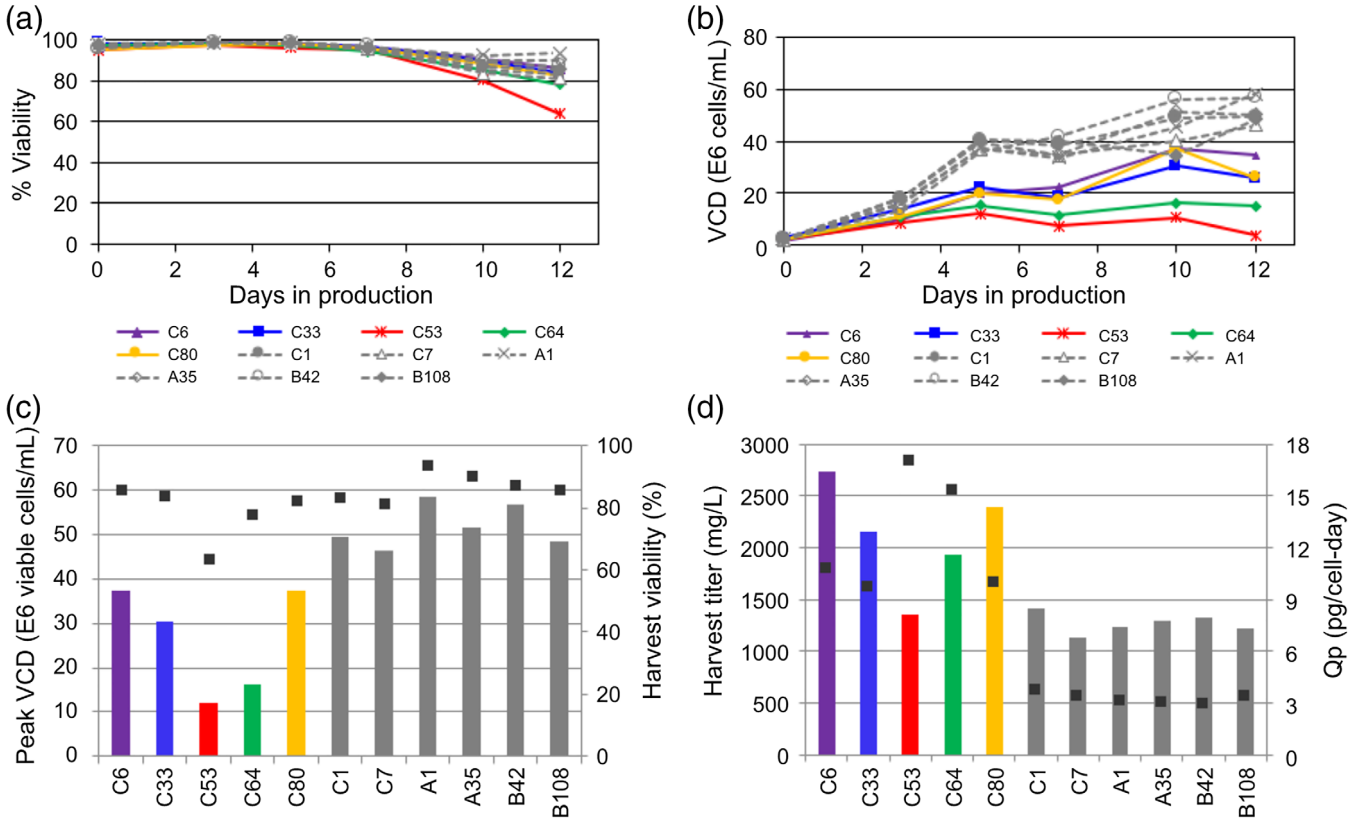


FIGURE 4 Phenotypic assessment of clones in a 12-Day AMBR fed-batch production process. (a) % Viability and (b) Viable cell density (VCD) over 12 days in AMBR fed-batch process. Five clones producing high amounts of POI (C6, C33, C53, C64, and C80) represented in color and a set of representative clones for each transfection having less than 1% POI (C1, C7, A1, A35, B42, and B108) represented in gray. (c) Peak VCD is depicted as one million cells per ml and % viability at harvest. Five clones with desirable phenotype represented in color and clones having less than 1% POI represented in gray. (d) Harvest titer of total protein A – bound material and specific productivity (Qp)

two copies detected per cell (Figure 6(d)). This demonstrates that a single copy of both the HC and HCF gene were integrated for the clones with high % POI. However, for the LC, the ddPCR showed two LC copies per cell, instead of the expected one LC copy per cell (Figure 6(d)).

Gene copy analysis on clones producing low levels of the heterodimeric molecule revealed a series of different genotypes. Based on the ddPCR results, clones A1, A35, C1, and C7 only harbor the HCF transgene (Figure 6(d)). This genotype was in agreement with protein analytics, as these clones produced protein with high

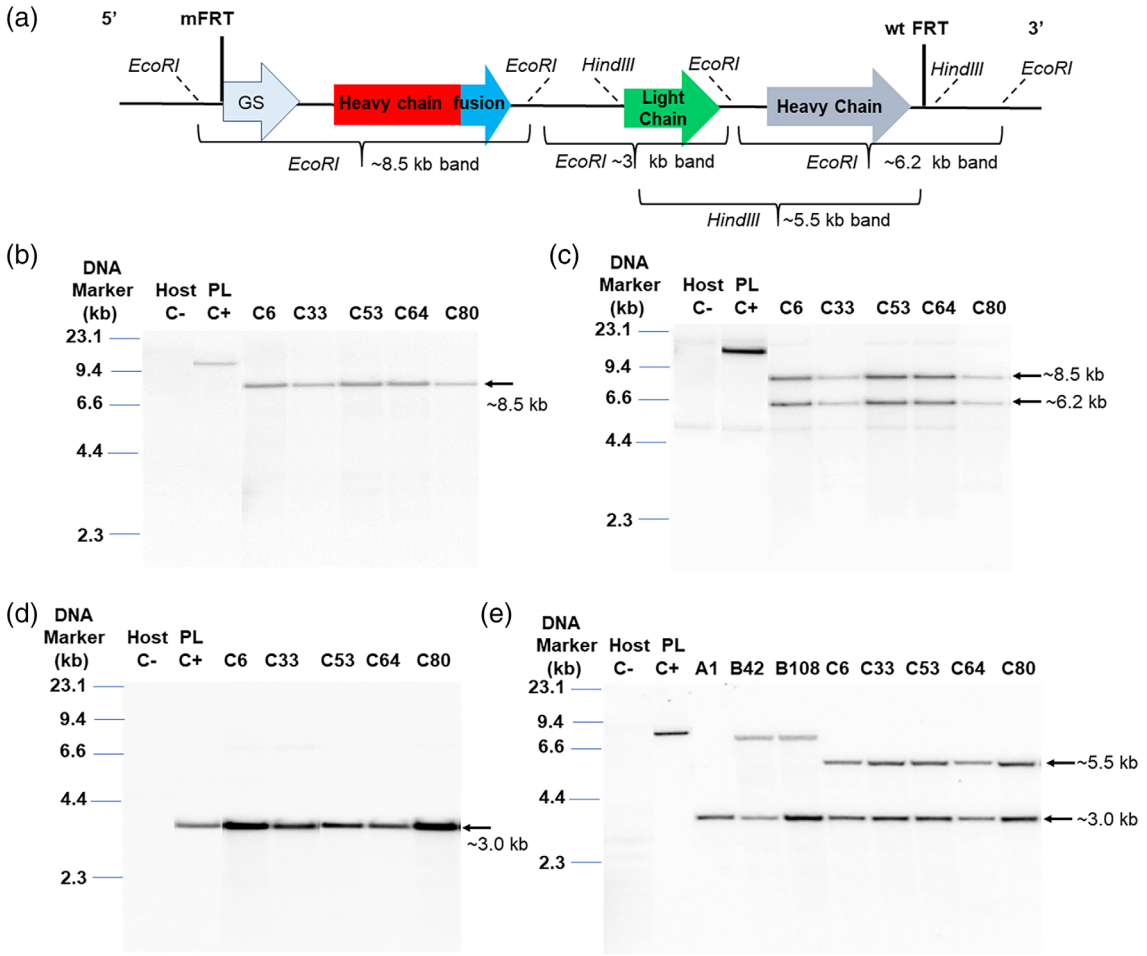


FIGURE 5 Southern blot characterization of clones expressing bispecific molecule 045. (a) A depiction of the vector HCF-LC-HC targeted integration *in silico* model. The *EcoRI* and *HindIII* digestion sites are marked and the expected restriction fragments are highlighted. ((b)–(e)) Southern blot analysis for untransfected host (negative control; C−); transfection vector (positive control; PL C+); and the five C clones producing high levels of POI: C6, C33, C53, C64, and C80. (b) Samples digested with *EcoRI* and Southern blot hybridized with fusion element gene fragment probe. (c) Samples digested with *EcoRI* and southern blot hybridized with HC gene probe. (d) Samples digested with *EcoRI* and Southern blot hybridized with LC gene probe. (e) Samples digested with *HindIII* and Southern blot hybridized with LC gene probe

percentages of the fusion homodimer and half molecule and below 1% of the heterodimer. The finding of a single copy of the HC, which is HCF, is also in agreement with the previous PCR results that suggested sequence alterations 3′ of the LC. Three copies of LC were observed for clones B42 and B108 and although gene copy analysis showed that both the HC and HCF were present, these clones also produced protein predominantly consisting of HCF fusion homodimer and half molecule. Interestingly, we did not observe any clones harboring a single LC copy by ddPCR. The gene copy analysis revealed that no clones analyzed had the expected genotype (one copy of each transgene), and that the differences observed in genotype among the clones could at least partially explain differences observed in protein product quality. The Southern blot results were in agreement with the observation from copy number experiments, as the blots revealed clear differences in banding topology and hence clone genotype, and confirmed correlation between two copies of LC with specific integration topology and high % POI. This unexpected genotype likely occurred post-transfection as the expression vector was sequenced verified prior to transfection. DNA sequencing and restriction enzyme mapping confirmed that the expression plasmid HCF-LC-HC was as expected and contained a single LC copy (Figure S4).

Pacbio SMRT long read sequencing was performed to further asses the genotypic integration topology of two high % POI producing clones (C33 and C53). Emphasis was placed on more in-depth understanding of the top % POI producing clones and determining the location of the second copy of the LC. Sequence capture probes were designed for non-GOI regions within the vector to enrich long reads correspondent to the exogenously introduced DNA fragments. Assembly of payload associated reads was thereby performed to confirm or revise the transfection vector-based model for the analyzed clones. An *in silico* validation was performed to ensure that the original SMRT long reads could be mapped back to the putative assembly without significant trimming. Such truncation manifests sharp decline or discontinuity in

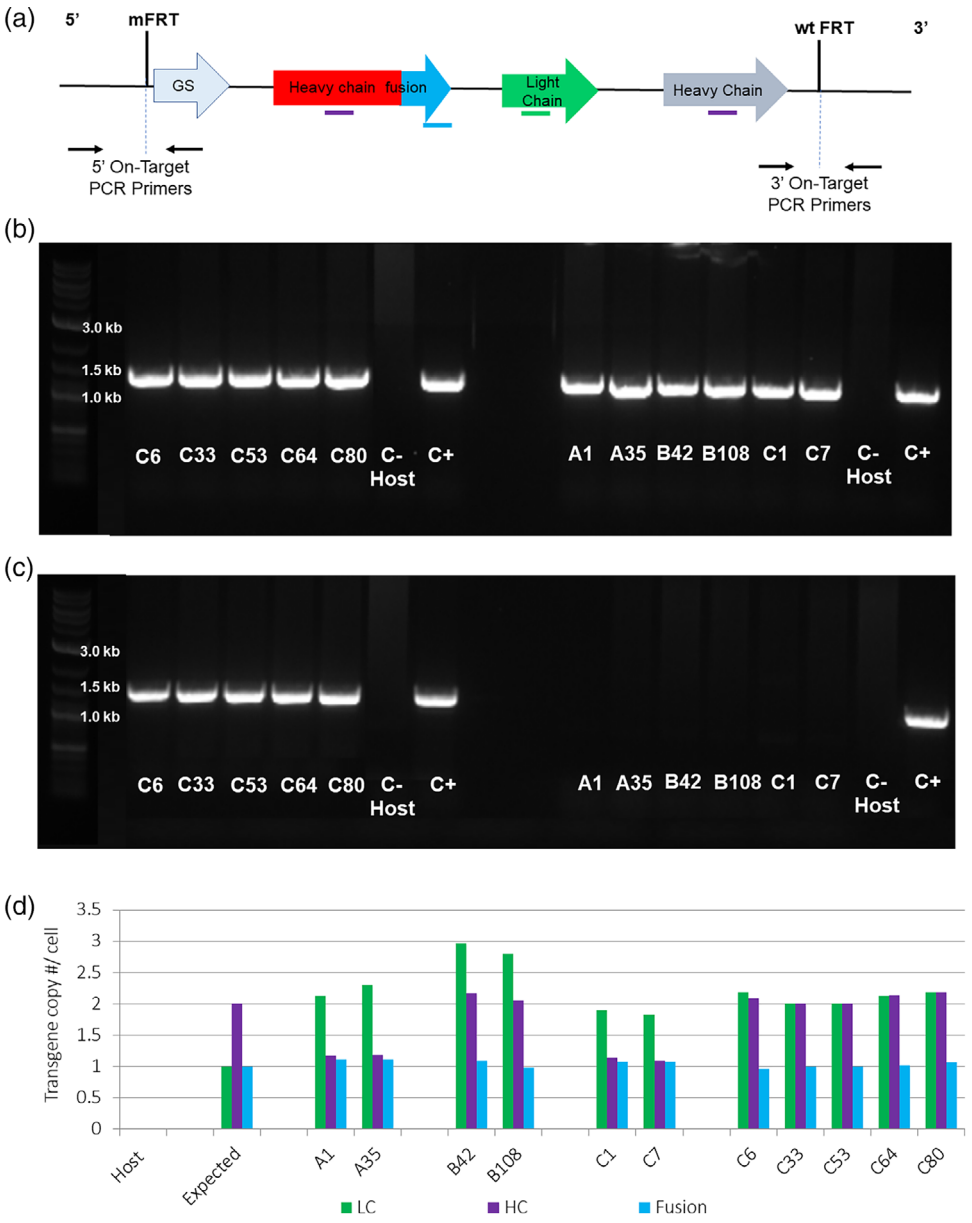


FIGURE 6 Genotypic Characterization of the Cell Lines from Vector HCF-LC-HC Transfections. (a) The vector HCF-LC-HC targeted integration in silico model. The black arrows indicate location of primers used to confirm 5′ and 3′ integration of the vector within the landing pad. The purple, blue and green lines indicate location of primers-probe sets used to assess gene copy number via digital droplet PCR (ddPCR). (b) Ethidium bromide gel showing PCR results with expected amplicon of 1.4 kb for 5′ integration for the five C clones producing high % POI (C3, C33, C53, C63, and C80), and six clones with poor % POI (A1, A35, B42, B108, C1, and C7), un-transfected host (Host), and previously tested positive control (C +). (c) Ethidium bromide gel showing PCR results with expected amplicon of 1.4 kb for 3′ integration for the five C clones producing high levels of POI (C3, C33, C53, C63, and C80), and six clones with poor % POI (A1, A35, B42, B108, C1, and C7), un-transfected host (Host), and previously tested positive control (C +). (d) Expected and experimental transgene copy per cell ddPCR results for un-transfected host; clones with poor % POI (A1, A35, B42, B108, C1, and C7); and five C clones producing high levels of % POI (C6, C33, C53, C64, and C80). Green: LC; Purple: HC; Blue: HC Fusion

the depth of coverage of the reference template, implicating necessity of repeating the assembly with adjusted parameters. A continuous coverage pattern was observed for both C33 and C53, suggesting the confidence of the assembly (Figure 7(a),(b), upper panel).

Retrospective analyses have further suggested that the Pacbio assemblies are in line with the experimental observations. The *in silico* *Hind*III digestion of the reconstructed payload yielded comparable digestion pattern to that from Southern blot analysis (Figure 5). Annotation of the assemblies revealed that re-arrangement occurred within the landing pad for C33 and C53 (Figure 7(a),(b), lower panel). Consistent with the gene copy number data, this rearrangement involved a duplication event where the LC along with its transcriptional elements, such as promoter, intron, and termination signal, was replicated in the correct orientation next to the original sequence.

3.4 | Evaluation of 2xLC transfection pools expressing the 045 bispecific in an SSI expression system

To further explore the utility of the SSI system to properly express molecule 045, several new vector topologies were designed to test the hypothesis that two copies of LC were required for cell survival and for proper PQ attributes. Different vector configurations harboring two copies of LC (referred to as 2xLC) were generated and transfected into the SSI host cell line (Figure 8(a)). One of the 2xLC vectors, the HCF-LC–LC-HC, was designed to reflect the orientation that was achieved through the selective pressure when using the vector HCF-LC-HC (1xLC control) in the initial transfections. Along with the new 2xLC vectors, the original HCF-LC-HC 1xLC vector was also transfected as a comparative control.

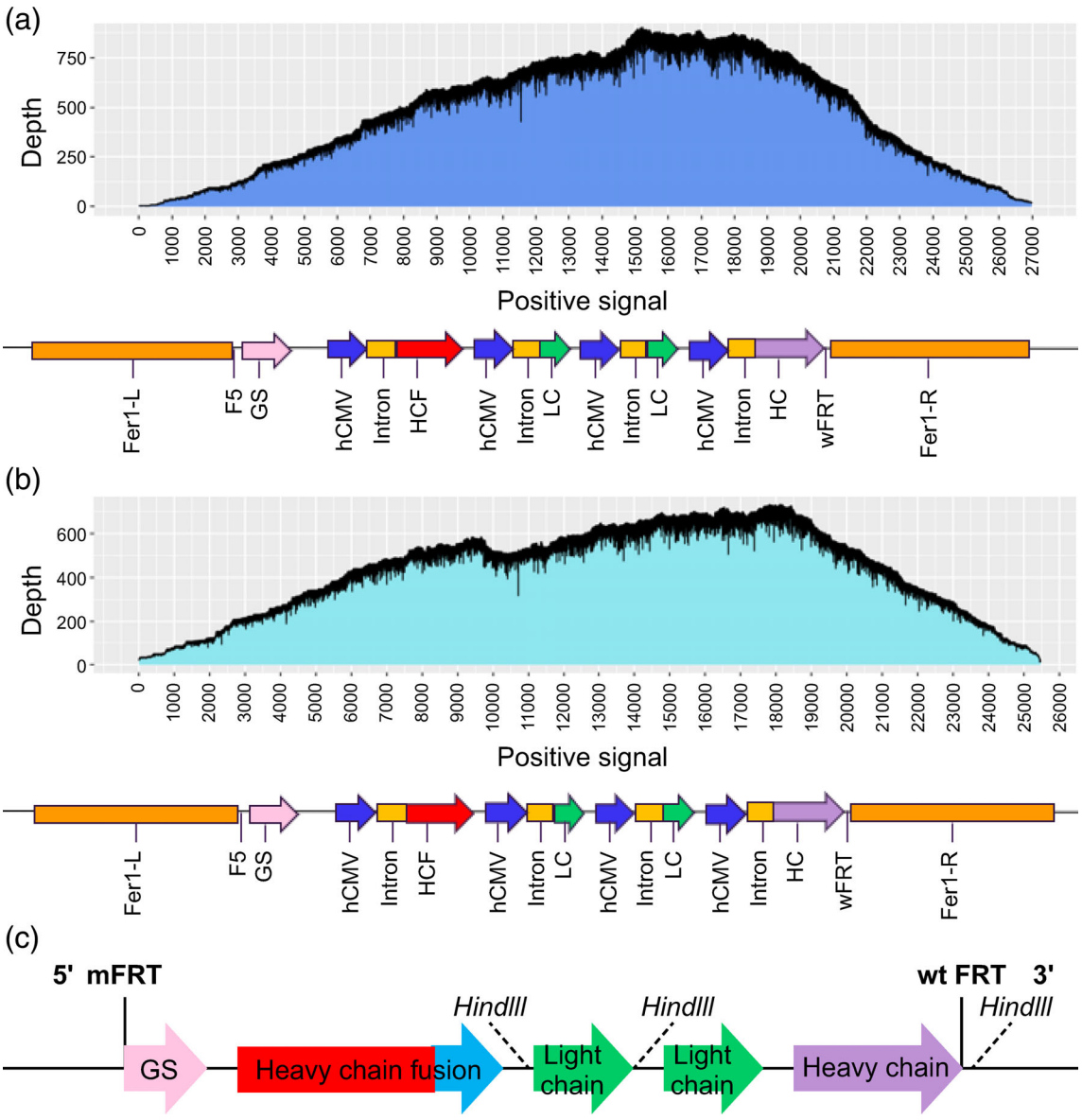
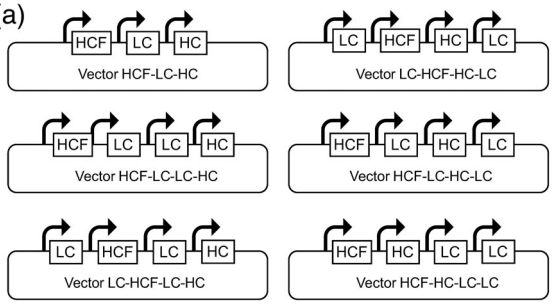


FIGURE 7 Targeted sequence capture analysis and generation of new model for the high % POI HCF-LC-HC Cell Lines. Coverage depth across the transgene genomic locus and schematic of the transgene genomic locus based on long-read sequence analysis for C33 and C53 clones ((a) and (b), respectively).The analysis confirmed duplication of the LC and its transcriptional components within the landing pad. (c) Depiction of the new model with predicted fragments liberated by *Hind*III digest and the presence of two LC

Once the transfection pools reached >95% viability, they were transferred to shake flasks and evaluated in a 12-day shake-flask fed-batch production process. Growth, viability, productivity, and product quality parameters were assessed for the 2xLC transfection pools and deemed comparable to HCF-LC-HC 1xLC control and to previous observations and performance (not shown). The productivity of the 2xLC pools ranged from 450 to 912 mg/ml while for the HCF-LC-HC 1xLC pool control was 44 mg/ml as determined by ProA HPLC (Figure 8(b)). Following purification over a ProA affinity column, size exclusion chromatography (SEC) analysis and nrCGE showed that the average % POI across the different 2xLC pools was significantly higher than for the HCF-LC-HC 1xLC control (Figures 1 and 8(b)). These results support our hypothesis that two copies of LC are required for proper formation and production of the 045 molecule. For one particular configuration, the HCF-HC-LC–LC, the % POI of the pool was 93% by nrCGE and 85% by SEC. Typically, a pool with high % POI is suggestive of potentially high rate of success when it comes to the percentage of clones with desirable attributes and of screening process that is time efficient and less effort consuming.

A discrepancy between SEC and nrCGE data for the HCF-LC-LC-HC configuration was observed, which indicated a 65% POI by SEC but only 27% POI by nrCGE. It is unclear as to why the molecule was unable to form more stably in the case of one particular vector configuration (HCF-LC-LC-HC) and not for others. The inconsistency between the nrCGE and SEC methods for these samples suggest that pools harboring the HCF-LC-LC-HC vector configuration were less

(a)



(b)

Transfection pools	Harvest titer mg/mL	% POI by nrCGE	% POI by SEC
HCF-LC-HC	44	18.9	21.41
HCF-LC-LC-HC	667	27.3	65.57
LC-HCF-LC-HC	696	65.6	73.27
LC-HCF-HC-LC	746	67.5	76.45
HCF-LC-HC-LC	450	72.9	66.38
HCF-HC-LC-LC	912	85.4	93.81

FIGURE 8 Vector configurations and pool performance in 12-day shake-flask fed-batch production process. (a) List of the different 2xLC vectors constructed and assessed in comparison to the 1xLC control (HCF-LC-HC) and the vector configuration which was selected during the initial transfecting process for the 045 project (HCF-LC–LC-HC). (b) % POI comparison of SEC and nrCGE results from the 2xLC pools and harvest titer as reported by HPLC

effective in producing material with proper disulfide bonding. Alternatively, pool heterogeneity and differences in production process might have played a role in the observed discrepancy. These results suggest a potential link between vector topology of LC:HC ratio and overall molecular stability and are consistent with observations from other groups.[24] Interestingly, this discrepancy between assay readouts was not observed during assessment of clonal cell lines harboring the HCF-LC-LC-HC topology either from 4-day shake-flask cultures or from 12-day AMBR fed-batch production process.

4 | CONCLUSION

A targeted integration host cell expression system combined with an extensive clone screen strategy was utilized to produce the 045 recombinant multi-chain bispecific molecule. Ultimately, we were successful in identifying several clone candidates with a desirable phenotype that moved forward. Throughout the course of this molecule's cell line development, several interesting findings were observed.

Initially, the transfection pools only produced 33% of the desired heterodimer with the remaining 67% consisting of product-related impurities. Using enhanced screening, we were able to find five clones (C6, C33, C53, C64, and C80) with significantly higher % POI than what was generated by the parental transfection pool. Three out of the five clones (C6, C64, and C80) were subjected to genetic generational stability studies in a 12-day AMBR fed-batch process. The data indicated consistent titer over 4 months of passaging, and more importantly unfluctuating % POI across the tested time points (Figure S3). For all clones with desirable % POI, two copies of the LC gene were present per cell. The two LC copies were located next to each other in a head to tail arrangement within the intended integration site. Previous groups have reported concatemerization of plasmid at a given integration site and demonstrated that the structural variation in the transgene integration region will impact expression, stability, and may contribute to potential rearrangements.[25,26] Our study indicates phenomenon unique to this, as only a specific portion of the vector (LC) was duplicated and was localized in the middle of the vector payload within the SSI landing pad. Furthermore, the rearrangement is specific to the intended integration site and no off-target events were observed.

Interestingly, of the clones that were analyzed genetically, none were found to have the expected genotype. The ELISA results suggest a bias of clones not expressing the HCF transgene (188 of 240 analyzed); perhaps due to the absence or loss of the HCF gene. Further, of the 52 clones with a positive ELISA signal for the fusion element, 47 predominantly produced the HCF homodimer, suggesting the absence of or inability to express the HC gene. In agreement with this, four of the six clones that were both predominantly producing HCF homodimer and analyzed genetically did not harbor the HC gene (Figure 6(d)). The negative PCR results of 3′ integration for all six clones analyzed with poor % POI suggest a partial deletion of the HC gene/transcriptional elements downstream of the LC, corroborating the nrCEG and ddPCR gene copy data (Figure 6(b),(c)).

We feel the most likely explanation for the observed transgene rearrangements is that the expected genotype of single copy of each of the GOIs was toxic to the cells, thus placing selective pressure on the population towards transgene rearrangements to evade the toxicity. This toxicity may have been the result of non-optimal LC:HC ratio caused in part by two copies of a HC gene (HC and HCF) per cell versus one copy of LC. Several groups have previously published reports describing HC toxicity and growth suppression.[27–29] Clones that were found to be either negative for ELISA signal or capable of producing only HCF homodimer may be products of genetic rearrangements to evade toxicity via loss of the HCF or HC gene/expression, respectively. Alternatively, LC duplication may have been the dominant mechanism allowing cell survival since all of the clones analyzed showed two LC copies. Clones producing high levels of the desired heterodimer may have evaded toxicity by duplication of the entire LC transcriptional unit to provide a necessary complement for the excess HC produced. It is likely that two copies of LC were required for survival through selection and production of the desired POI.

For multi-chain, multi-specific molecules expressed in SSI systems, an up-front evaluation of multiple vector configurations could be used to enhance the probability of identifying clones with desirable properties including appropriate % POI. Without an optimal vector strategy, cells may rearrange transgenes to avoid product-related toxicity. These rearrangements may or may not result in clones with

desirable % POI. Thus, for non-optimized vectors, significant clone screening may be required to identify clones capable of expressing appropriate product. Our results suggest that for molecule 045 the transfected vector was not optimal, and a better configuration would include at least two LC copies on the transfected plasmid. We tested this hypothesis by generating and transfecting multiple 2xLC vector orientations. The results from the transfection pools clearly illustrate that all vectors that contained two copies of the LC had significantly higher % POI than the one copy control. The % POI from the 2xLC vector increased to over 85% in comparison to the HCF-LC-HC 1xLC control of approximately 20%. In this case, up-front vector optimization may have resulted in an increased number of clones producing molecules with desirable product quality attributes and eliminated the need to carry out extensive clone screening. These benefits may also extend beyond cell line generation, as a larger selection of clones producing high % POI can equate to shorter, less-intensive purification process and analytical development efforts.

Would an RI approach have afforded better clone screen outcomes for this particular bi-specific molecule? Typically, cloning results from the RI system yield a wide range of genotypes and phenotypes, allowing for higher clonal diversity. In comparison to the SSI system, the RI system does offer the ability for transcriptional variability by using multiple expression vectors for different transgenes and altering the vector ratios used for transfection. Certain chains within the multispecific may exist at a higher molar ratio relative to other chains, which would require higher expression of these genes. Additionally, chains that are different than the naturally occurring heavy and LCs may not possess the same level of transcriptional and/or translational efficiency. By virtue of transgene insertions at arbitrary genomic locations, the RI system may provide a range of transcriptional efficiency for transgenes that can be harnessed through single cell cloning and screening for proper product formation. Given the static nature with respect to integration site and transgene copy number of an SSI host system, one may not expect a wide range of transcriptional variability, theoretically resulting in more objective outcomes for clones of certain vector configurations. This property is usually viewed as a positive for an SSI system. While our case study fortuitously suggests that this is not always true, two other properties of the SSI system should be realized for future development of bispecific molecules. The first of these is genetic stability of the transgene. An inherent drawback to RI remains potential instability across generational age since the transgene(s) are integrated at random and possibly in unstable chromosomal locations. To address this drawback, intensive screening over generational age is essential for the RI system, which could have been further compounded by the complexity of the 045 bi-specific. Second, in contrast to the RI, SSI allows for a quicker determination of the likelihood of success/failure during clone screen. As established by this study, the analysis of material generated from SSI transfection pools can generate meaningful results with respect to both product titer and product quality, whereas with RI, a cloning step is more likely to be required, caused in part by a higher percentage of non-expressing cells.

Even with an optimized vector strategy, it is not clear if a targeted integration system will be able to yield desirable outcomes for even more complex modalities, for example molecules comprised of greater than three unique chains. Vector size may become limiting with respect to ease of construction and stability within the host system. As the number of chains increase, so will the number of vector topologies in need of examination prior to cloning. Importantly, a single vector may not provide the appropriate transcript ratios required to attain a desirable outcome. To this end, a potential area of focus includes examination of cell lines possessing multiple landing pads capable of receiving different vectors.[30] These efforts coupled with an understanding of transcript expression and ratio within and across the inserted landing pads may lead to consistent, reproducible generation of cell lines producing protein with intended quality and generational stability. If successful, this could enable highly efficient cell line development for complex modalities on accelerated early-phase timelines.

ACKNOWLEDGMENTS

The authors would like to acknowledge and thank Martin Allen for useful advice during the preparation of this manuscript.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1002/btpr.3158.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Barbara Tevelev https://orcid.org/0000-0002-1162-7541
John J. Scarcelli https://orcid.org/0000-0002-3314-7602

REFERENCES

1. Li F, Vijayasankaran N, Shen AY, Kiss R, Amanullah A. Cell culture processes for monoclonal antibody production. *MAbs*. 2010;2(5):466-479.
2. Kim JY, Kim YG, Lee GM. CHO cells in biotechnology for production of recombinant proteins: current state and further potential. *Appl Microbiol Biotechnol*. 2012;93(3):917-930.
3. Zhong X, Ma W, Meade CL, et al. Transient CHO expression platform for robust antibody production and its enhanced N-glycan sialylation on therapeutic glycoproteins. *Biotechnol Prog*. 2019;35(1):e2724.
4. Wurm FM. Production of recombinant protein therapeutics in cultivated mammalian cells. *Nat Biotechnol*. 2004;22(11):1393-1398.
5. Zhu J. Mammalian cell protein expression for biopharmaceutical production. *Biotechnol Adv*. 2012;30(5):1158-1170.
6. Kelley B. Industrialization of mAb production technology: the bioprocessing industry at a crossroads. *MAbs*. 2009;1(5):443-452.
7. Gorman C, Bullock C. Site-specific gene targeting for gene expression in eukaryotes. *Curr Opin Biotechnol*. 2000;11(5):455-460.
8. Zhao M, Wang J, Luo M, et al. Rapid development of stable transgene CHO cell lines by CRISPR/Cas9-mediated site-specific integration into C12orf35. *Appl Microbiol Biotechnol*. 2018;102(14):6105-6117.
9. Yang H, Wang J, Zhao M, et al. Feasible development of stable HEK293 clones by CRISPR/Cas9-mediated site-specific integration for biopharmaceuticals production. *Biotechnol Lett*. 2019;41(8–9):941-950.
10. Lee JS, Grav LM, Pedersen LE, Lee GM, Kildegaard HF. Accelerated homology-directed targeted integration of transgenes in Chinese hamster ovary cells via CRISPR/Cas9 and fluorescent enrichment. *Biotechnol Bioeng*. 2016;113(11):2518-2523.

11. Inniss MC, Bandara K, Jusiak B, et al. A novel Bxb1 integrase RMCE system for high fidelity site-specific integration of mAb expression cassette in CHO cells. *Biotechnol Bioeng*. 2017;114(8):1837-1846.
12. Turan S, Zehe C, Kuehle J, Qiao J, Bode J. Recombinase-mediated cassette exchange (RMCE) - a rapidly-expanding toolbox for targeted genomic modifications. *Gene*. 2013;515(1):1-27.
13. Baumann M, Gludovacz E, Sealover N, et al. Preselection of recombinant gene integration sites enabling high transcription rates in CHO cells using alternate start codons and recombinase mediated cassette exchange. *Biotechnol Bioeng*. 2017;114(11):2616-2627.
14. Zhang L, Inniss MC, Han S, et al. Recombinase-mediated cassette exchange (RMCE) for monoclonal antibody expression in the commercially relevant CHOK1SV cell line. *Biotechnol Prog*. 2015;31(6):1645-1656.
15. Kim MS, Lee GM. Use of Flp-mediated cassette exchange in the development of a CHO cell line stably producing erythropoietin. *J Microbiol Biotechnol*. 2008;18(7):1342-1351.
16. Grav LM, Sergeeva D, Lee JS, et al. Minimizing clonal variation during mammalian cell line engineering for improved systems biology data generation. *ACS Synth Biol*. 2018;7(9):2148-2159.
17. Feary M, Moffat MA, Casperson GF, Allen MJ, Young RJ. CHOK1SV GS-KO SSI expression system: a combination of the Fer1L4 locus and glutamine synthetase selection. *Biotechnol Prog*. 2021;e3137. https://doi.org/10.1002/btpr.3137.
18. Liu H, Saxena A, Sidhu SS, Wu D. Fc engineering for developing therapeutic bispecific antibodies and novel scaffolds. *Front Immunol*. 2017;8:38.
19. Carter PJ, Lazar GA. Next generation antibody drugs: pursuit of the 'high-hanging fruit'. *Nat Rev Drug Discov*. 2018;17(3):197-223.
20. Brinkmann U, Kontermann RE. The making of bispecific antibodies. *MAbs*. 2017;9(2):182-212.
21. Spiess C, Zhai Q, Carter PJ. Alternative molecular formats and therapeutic applications for bispecific antibodies. *Mol Immunol*. 2015;67(2 Pt A):95-106.
22. Strop P, Ho WH, Boustany LM, et al. Generating bispecific human IgG1 and IgG2 antibodies from any antibody pair. *J Mol Biol*. 2012;420(3):204-219.
23. Ridgway JB, Presta LG, Carter P. 'Knobs-into-holes' engineering of antibody CH3 domains for heavy chain heterodimerization. *Protein Eng*. 1996;9(7):617-621.
24. Ho SC, Koh EY, van Beers M, et al. Control of IgG LC:HC ratio in stably transfected CHO cells and study of the impact on expression, aggregation, glycosylation and conformational stability. *J Biotechnol*. 2013;165(3-4):157-166.
25. Dhiman H, Campbell M, Melcher M, Smith KD, Borth N. Predicting favorable landing pads for targeted integrations in Chinese hamster ovary cell lines by learning stability characteristics from random transgene integrations. *Comput Struct Biotechnol J*. 2020;18:3632-3648.
26. Bandyopadhyay AA, O'Brien SA, Zhao L, Fu HY, Vishwanathan N, Hu WS. Recurring genomic structural variation leads to clonal instability and loss of productivity. *Biotechnol Bioeng*. 2019;116(1):41-53.
27. Kohler G. Immunoglobulin chain loss in hybridoma lines. *Proc Natl Acad Sci U S A*. 1980;77(4):2197-2199.
28. Haas IG, Wabl MR. Immunoglobulin heavy chain toxicity in plasma cells is neutralized by fusion to pre-B cells. *Proc Natl Acad Sci U S A*. 1984;81(22):7185-7188.
29. Wu K, Hoy MA. Clathrin heavy chain is important for viability, oviposition, embryogenesis and, possibly, systemic RNAi response in the predatory mite Metaseiulus occidentalis. *PLoS One*. 2014;9(10):e110874.
30. Gaidukov L, Wroblewska L, Teague B, et al. A multi-landing pad DNA integration platform for mammalian cell engineering. *Nucleic Acids Res*. 2018;46(8):4072-4086.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Tevelev B, Patel H, Shields K, et al. Genetic rearrangement during site specific integration event facilitates cell line development of a bispecific molecule. *Biotechnol Progress*. 2021;37:e3158. https://doi.org/10.1002/btpr.3158