Two-color selection for amplified co-production of proteins in mammalian cells

Zahra Assur a, Ira Schieren b, Wayne A. Hendrickson a,b,*, Filippo Mancia a,*

a Department of Biochemistry and Molecular Biophysics, Columbia University, New York, NY 10032, USA
b Howard Hughes Medical Institute, Columbia University, New York, NY 10032, USA

Received 17 March 2007, and in revised form 26 April 2007
Available online 16 May 2007

Abstract

Structural and functional analyses for many mammalian systems depend on having abundant supplies of recombinant multi-protein complexes that can be produced best, or only, in mammalian cells. We present an efficient fluorescence marking procedure for establishing stable cell lines that overexpress two proteins in co-ordination, and we validate the method in the production of monoclonal antibody Fab fragments. The procedure has worked without fail on all seven of seven trials on Fabs, which are being used in the crystallization of G-protein coupled receptors. This manner of efficient selection may readily be adapted for the co-production of other complexes of two or more proteins.

© 2007 Elsevier Inc. All rights reserved.

Keywords: Antibodies; Protein overexpression; Mammalian cells; Protein complexes; Fluorescence

Recent proteomic studies have strengthened the concept that functional units within a cell typically consist of macromolecular assemblages as opposed to individual proteins [1]. Accordingly, there is a justified and increasing interest in structure determination and biophysical characterization of macromolecular complexes [2]. These interests require the production of stable complexes, which for the most part depends on recombinant expression in appropriate hosts. Since reconstitution from separately expressed proteins is often frustrated by instability of the isolated components, there are obvious advantages for simultaneous overexpression of multiple proteins in the same cell. Bacteria, given their cost-effectiveness and robustness, have been by far the most successful hosts for protein production to support biophysical studies, and strategies for the co-expression of genes in Escherichia coli are available [2,3]. Bacteria, and yeast or insect cells for that matter, might be poorly suited for the production of mammalian proteins [4,5]. These proteins may be toxic to the host, may require a better matched folding machinery, and may need post-translational modifications unavailable in the heterologous host.

Expression of recombinant proteins in mammalian cells can be achieved through transient transfection, viral infection or stable integration of expression constructs into the host genome. Transient co-transfection of separate plasmids carrying the individual genes represents possibly the most frequently employed technique for functional studies. Unfortunately, this has yet to translate into a general approach for producing abundant amounts of material as required for applications such as structure determination. Stable cell lines have obvious advantages for amplified recombinant protein production. However, the integration of the expression constructs into the genome of the host cell inevitably leads to a wide spectrum in protein synthesis levels within the same batch of transfected cells, depending primarily on the number of integrants and their sites of integration. The method therefore relies on being able to select for cells capable of producing the most protein.

* Corresponding authors. Address: Department of Biochemistry and Molecular Biophysics, Columbia University, 630 West 168th Street, New York, NY 10032, USA. Fax: +1 212 305 7952.
E-mail addresses: wayne@convex.hhmi.columbia.edu (W.A. Hendrickson), fm123@columbia.edu (F. Mancia).
Through coupling of the expression of each polypeptide chain with that of different antibiotic-resistant markers, targeting a specific genomic locus for integration, or incorporating all the components in a single plasmid one can increase the frequency of stable transfectants able to synthesize all the desired components. These provide only marginal relief, however, in the task of identifying the best expressing cells, which has typically involved time-consuming rounds of screening to yield lines with the required characteristics [6].

Here we present and validate procedures for the rapid selection of mammalian cells that co-express multiple proteins. We do this by coupling the expression of each protein chain of a complex to a separate fluorescent marker, and we test the system in applications to the high-level expression of antibody Fab fragments. Proper functionality of the expressed complexes was demonstrated by assessing correct assembly of the antibody fragments and their ability to recognize the antigen, a 5HT2c serotonin receptor.

Materials and methods

Cloning of Fv chains

Cloning of Fv regions, inclusive of signal peptide sequences, was performed with appropriate kits and degenerate primers (Novagen) following standard procedures and guidelines. The amplified fragments were cloned into pGEM-T vector (Promega). A second PCR was used to introduce appropriate restriction sites for cloning into expression vectors.

Construction of expression vectors

pFM1.2 [7] was used to generate two vectors for the separate expression of the two Fab chains. For light chain expression, GFP was substituted for RFP, taken from pires2-DsRed-Express (Clontech). The CH-His6 and CL regions of D1.3 anti-lysozyme antibody were removed from pASK84 [8], and cloned into the MCS regions of the respective vectors. The heavy and light chains of the Fv regions were then cloned into the matching vectors to generate the final expression constructs. The vectors for expression of heavy and light chains were named pFM-FabH and pFMFabL, respectively.

Cell culture and generation of stable lines

HEK293 cells were maintained at 37 °C, in a humidified environment enriched with 5% CO2. HEK293 cells were grown in DMEM (Chemicon) supplemented with 10% FBS (Hyclone), Penicillin, Streptomycin, L-glutamine (Pen/Strep/L-glu; Sigma).1 293 GntI cells [9] carrying a 5HT2c serotonin receptor were grown in DMEM/F12 (Gibco) supplemented with 10% FBS, Pen/Strep/L-glu, 4 µg/mL blasticidin (Invitrogen) and 500 µg/mL G418. A plasmid carrying resistance to puromycin was mixed with the two expression vectors in a 1:5:5 ratio prior to transfection with lipofectamine (Invitrogen). Stable integrants were selected by addition of 5 µg/mL puromycin to the growth medium. Production cell lines from single double fluorescent colonies were selected either by FAC sorting in Autoclon mode and subsequent visual inspection of the resulting clones, to identify single, highly fluorescent colonies, or by manual picking of the most intense double fluorescent colonies after antibiotic selection.

FAC sorting

Data were collected using a Beckman Coulter Altra flow cytometer equipped with Autoclon. Untransfected (control) and stably integrated cells were passed through the cell sorter. The viable cell population was determined using the forward and side scatter characteristics of the cells. The fluorescent cells were excited at the 488 nM line of a krypton–argon laser. GFP emission was detected with a 525/30 nm band pass filter. Viable cells (100,000) were collected for each pool, according to their fluorescence profile. Typically, the top fluorescent cells of the double positive population (corresponding to 0.1% of the viable cell population) were chosen for cloning to single cell purity. This was achieved with the flow cytometer in Autoclon mode, in 96-well plates.

Fab purification

Na–Hepes, pH 7.5, NaCl and imidazole were added to the harvested medium to 20, 400 and 10 mM final concentrations, respectively. The medium was added to equilibrated Ni–NTA (Qiagen) on a column. The column was washed with 10 column volumes of 20 mM Na–Hepes, pH 7.5, 400 mM NaCl, 25 mM imidazole. Purified Fab was eluted with 200 mM imidazole in the same buffer.

Confocal microscopy on fluorescent cells

Cells were plated on 8-well CC2 treated glass slides (Nalge Nunc Int.) at a density of 1 x 10^5 per well. The next day, the cells were fixed with 4% paraformaldehyde in 0.1 M phosphate buffer for 10 min at room temperature. Cells were then washed 3 times with PBS and incubated for 30 min with the nuclear stain TOTO-3 iodide (Invitrogen) diluted at 1:2000 in PBS. Stained cells were viewed and photographed on a Zeiss confocal microscope at 40x magnification in three separate emission channels, at 520 nM (GFP), 570 nM (RFP) and 633 nM (TOTO-3).
Immunocytochemistry

293 GntI− cells carrying stable integration of an inducible expression cassette for 5HT2c were induced with tetracycline (2 μg/mL) and sodium butyrate (5 mM). Two days later, induced and uninduced (control) cells were plated on 8-well CC2 treated glass slides (Nalge Nunc Int.) at a density of $1 \times 10^5$ per well. After 24 h, cells were fixed with 4% paraformaldehyde in 0.1 M phosphate buffer for 10 min at room temperature. The fixed cells were blocked with 5% goat serum in PBS and incubated for 1 h with hybridoma supernatants (for full-length Ab), and medium from Fab-expressing cells (for Fab). The secondary antibody, used at a dilution of 1:1000, was Cy3-conjugated anti-mouse IgG (Jackson Immunoresearch). TOTO-3 iodide (Invitrogen) was used diluted at 1:2000 as a nuclear stain. Labeled cells were visualized and photographed by confocal microscopy as described in the previous paragraph.

Fab-antigen complex reconstitution and analysis

Purified Fab was concentrated to ~2 mg/mL. 5HT2c expressed in E. coli as a fusion to periplasmic maltose-binding protein (MBP) was purified on an immobilized streptavidin column by means of an SBP epitope tag [10] fused to the C-terminus of the receptor, by prior solubilization of the isolated membranes at 5 mg/mL membrane protein concentration with 1% (w/v) n-dodecyl-β-D-maltopyranoside (αDDM; Anatrace) and 0.2% cholesteryl hemisuccinate, (CHS; Anatrace). MBP was removed by cleavage of the fusion with TEV protease [11] at an engineered site in the linker region. Purified 5HT2c was concentrated to ~1 mg/mL. For reconstitution, detergent was added to the concentrated Fab to 0.1% final concentration, mixed with 5HT2c in approximately a 2:1 molar ratio and incubated on ice for 2 h. Samples were run on a 7.5 mm × 60 cm TSKgel G3000SW column (TOSOH Bioscience) at a flow rate of 0.3 mL/min in 20 mM Na–Hepes, pH 7.0, 200 mM NaCl, 1 mM EDTA, 0.025% αDDM.

Results

We first adapted a previously described fluorescence-selection vector for the high-level production of proteins in mammalian cells [7] for use with a second color. This vector carries a strong, constitutive promoter sequence, derived from cytomegalovirus (CMV [12]). Target gene expression is coupled via an IRES element [13] to that of green fluorescent protein (GFP [14]), and was modified to replace the coding sequence for GFP with that for red fluorescent protein (RFP [15]). The target genes for our current application are those for antibody Fab fragments corresponding to monoclonal antibodies (MAbs) against a rat G-protein coupled receptor (GPCR) for serotonin.
(5HT2c [16]), which were generated as previously described [17]. In this instance, we further modified the GFP- and RFP-containing vectors to include coding sequences for the conserved heavy-chain CH1 and light-chain CL domains from the extensively characterized, anti-lysozyme antibody D1.3 [18] ahead of the respective IRES elements. CH1 and CL domains do not confer antigen specificity to the antibody, and are interchangeable, even between species [6,19]. These provided the scaffold to express cloned VH and VL domains as Fab fragments. Six histidine residues were genetically fused to the C-terminus of CH1 for purification purposes [8]. Finally, coding sequences for VH and VL domains, inclusive of signal peptide sequences, were cloned from our hybridoma cells [17] using established protocols and inserted in-frame into vectors with matching C regions. A schematic of the two resulting expression constructs, one for each chain of the Fab fragment, is shown in Fig. 1a.

After co-transfection of the two plasmids into HEK293 cells, together with that for a selectable marker for antibiotic resistance, stable integrants were selected. As expected, the resulting antibiotic-resistant colonies displayed a wide spectrum of fluorescence levels, due to the presence of both RFP and GFP. First, then, we wanted to determine if the cells with both red and green fluorescence also actually secreted antibody fragments. Second, we were interested in verifying the extent to which fluorescence intensity correlated with Fab expression. In other words, we wanted to test the predictive value of our fluorescent markers for the production of desired protein products.

Using a fluorescence-activated cell sorter (FACS), we classified cells according to their two-color fluorescence profiles into a $4 \times 4$ grid of $-, +, ++$ and +++ for each color. The vast majority of cells after antibiotic selection were non-fluorescent, or only green, and populations fell off progressively with intensity for each color (Fig. 1b). Pools of cells with pertinent characteristics were selected and analyzed. Cells from these selected pools were visualized by confocal microscopy for fluorescence from RFP, GFP and a nuclear stain, and Fabs secreted from these different pools were purified by metal affinity chromatography and observed by SDS–PAGE. Results from such experiments are shown in Fig. 1b for 2A11, a Fab fragment from an anti-5HT2c MAb [17]. We find that while negative cells (RFP/C0/GFP/C0) showed no detectable antibody fragments, Fab expression was always associated with fluorescence. Moreover, robust Fab expression could only be detected in double-positive cells, in amounts that were correlated to the levels of GFP and RFP fluorescence.

Each production cell line was cloned from a single RFP+++/GFP+++ progenitor cell and chosen from a few such cultures to assure that maximal fluorescence was maintained (Fig. 2a). For these experiments, the RFP+++/GFP+++ grid was typically set to include only 0.1% of the cells. FACS analysis of a production cell line for Fab 2A11, after having been maintained for several months in continuous culture is shown in Fig. 1d. Cells were grown in adherent monolayers, and for Fab 2A11 the yield was quantified to be above 30 mg of purified protein per liter of medium from the endpoint of a time course experiment (Fig. 3) and from subsequent scale-ups (data not shown). Comparable yields have also been obtained from each of six other Fab productions attempted with this system, three reactive to 5HT2c, and three to a different GPCR. Relative yields were judged from the relative intensities of Coomassie blue-stained bands on gels. Three of these, all from antibodies against 5HT2c, are shown in Fig. 2. Ultimate production levels (Fig. 2b) seem to correlate with maximal fluorescence across these proteins (Fig. 2a) much as Fab production correlates with fluorescence for a given protein (Fig. 1c).

In contrast to our 100% success rate so far from two-color selection in HEK293 cells, alternative methods have proved far less reliable for us. We attempted Fab production by digestion of the intact IgG with papain, but only one of four different antibody-secreting hybridoma cell lines tried could be adapted to grow in serum-free medium and scaled-up for Fab production. We also attempted to

---

**Fig. 2.** Fluorescence levels, yields and folding of expressed Fabs from production cell lines. (a) Cells secreting a specific Fab fragment, 1C4, 2A11, 5D8 or 1A12, as indicated on the top of each corresponding column, were visualized by confocal microscopy for fluorescence from RFP (top row), GFP (middle row) and TOTO-3 (bottom row). (b) The different Fabs were run, after purification, on an SDS–PAGE denaturing gel after treatment with loading dye with (+) and without (−)β-mercaptoethanol. The identity of each Fab is shown above the corresponding lane.
express recombinant Fabs in *E. coli* following established procedures (reviewed in [20]), but achieved yields in the micrograms per liter range for the two antibodies tested. Reported expression levels of Fabs in *E. coli* [20,21] seem highly variable, while in the yeast *Pichia pastoris*, proper Fab assembly seems to be the major rate limiting factor [22]. In both these hosts, high-density cell growth in appropriate fermenters appears to have been essential for very high yields [20,22]. Indeed, if required, yields from our cell lines could most likely be improved dramatically by optimizing cell growth conditions, as in current practice for industrial production of antibodies in mammalian cells [6].

The light and heavy chains of the purified Fab appeared to be disulfide-bond linked, in agreement with what is typical for a correctly folded antibody (Fig. 2b). Furthermore, functionality of the expressed Fab was assessed by immunohistochemistry on antigen-expressing cells, showing a staining profile essentially identical to that observed with the intact IgG antibody derived from hybridomas (Fig. 4a). Finally, in preparation for crystallization trials, the Fab-5HT2c complex was successfully reconstituted with purified components as shown by size-exclusion chromatography (Fig. 4b). All these data are consistent with proper assembly and function of the expressed Fab.

**Discussion**

Both the tight correlation that we observe between production levels of functional antibody fragments and two-color fluorescence of transfected HEK293 cells (Fig. 1c) and the high reliability of the procedure (seven successes in seven trials) provide a compelling validation of this selection method. Our application to monoclonal antibody production itself has broad potential. Monoclonal antibodies are important tools for structural studies on membrane proteins [23]. Indeed, Fabs described here are now being used in our crystallization of the 5HT2c receptor. Monoclonal antibodies are also gaining an increasingly relevant role in clinical applications [6,19]. This two-color system should have other applications as well, however, as we have already shown that our previous GFP system works for other secreted proteins and for membrane proteins such as GPCRs [7]. Moreover, given the availability of several different fluorescent proteins [24], one could anticipate being able to

![Fig. 3. Time course of Fab expression.](image1)

18 x 10^6 cells expressing Fab 2A11 were equally divided into 6, 10 cm-diameter dishes, in 10 mL of medium. At each time point, the medium from a plate was removed, cells counted, and the Fab purified. 1/10th of each purified sample was run on an SDS–PAGE gel, stained with Coomassie blue. The length of each incubation period is indicated above the matching lane, in days, and the corresponding number of cells, in millions, below. The yield of Fab resulting from a 12 day incubation period was quantified by absorbance reading at 280 nM.

![Fig. 4. Functionality of the expressed Fab.](image2)

(a) Immunocytochemistry using recombinant Fab 2A11 (rFab) and intact antibody (IgG) on cells, before and after induction of 5HT2c expression. (b) Overlay of elution profiles from size-exclusion chromatography, monitored by absorbance at 280 nM, of purified 5HT2c (blue trace), purified Fab (green trace) and reconstituted 5HT2c-Fab complex (red trace).
extend this methodology to studies of multi-component protein complexes.

In contrast to time-consuming multi-step conventional amplification systems for generating stably transfected mammalian cell lines, this fluorescence-based approach is simple and rapid. Fluorescence surveillance is also convenient for monitoring and maintaining the continued health of an established cell line. We hope that these procedures may aid in finally bridging the gap between multi-gene expression in mammalian cells for functional studies, which is commonplace and routine, and the use of mammalian cells for biophysical studies, which has had very limited popularity even as the structural community is increasingly interested in macromolecular assemblies.

Acknowledgments

We are grateful to Richard Axel for discussions. We thank Susan Brenner-Morton for advice and help with cloning of the antibody fragments and Yonghua Sun for excellent technical assistance. This work was supported in part by NIH Grants GM68671 and GM75026.

References


