pMINERVA: A Donor-Acceptor System for the *in vivo* Recombineering of scFv into IgG Molecules.

Batonick M[1], Kiss MM[1], Fuller EP[1], Magadan CM[1], Holland EG[1], Zhao Q[1], Wang D[1], Kay BK[2], Weiner MP[1]

[1] AxioMx, Inc., 688 E. Main St., Branford, CT 06405

[2] University of Illinois at Chicago, 845 West Taylor Street Chicago, IL 60607

*To whom correspondence should be addressed:

Melissa Batonick, Ph.D.
AxioMx, Inc.
688 E. Main St.
Branford, CT 06405
Office: (203)208-1918
mbatonick@axiomxinc.com

ABSTRACT

Phage display is the most widely used method for selecting binding molecules from recombinant antibody libraries. However, validation of the phage antibodies often requires early production of the cognate full-length immunoglobulin G (IgG). The conversion of phage library outputs to a full immunoglobulin via standard subcloning is time-consuming and limits the number of clones that can be evaluated. We have developed a novel system to convert scFvs from a phage display vector directly into IgGs without any *in vitro* subcloning steps. This new vector system, named pMINERVA, makes clever use of site-specific bacteriophage integrases that are expressed in *E. coli* and intron splicing that occurs within mammalian cells. Using this system, a phage display vector contains both bacterial and mammalian regulatory regions that support antibody expression in *E. coli* and mammalian cells. A single-chain variable fragment (scFv) antibody is expressed on the surface of bacteriophage M13 as a genetic fusion to the gpIII coat protein. The scFv is converted to an IgG that can be expressed in mammalian cells by transducing a second *E. coli* strain. In that strain, the phiC31 recombinase fuses the heavy chain constant domain from an acceptor plasmid to the heavy chain variable domain and introduces controlling elements upstream of the light chain variable domain. Splicing in mammalian cells removes a synthetic intron containing the M13 gpIII gene to produce the fusion of the light chain variable domain to the constant domain. We show that phage displaying a scFv and recombinant IgGs generated using this system are expressed at wild-type levels and retain normal function. *Use of the pMINERVA completely eliminates the labor-intensive subcloning and DNA sequence confirmation steps currently needed to convert a scFv into a functional IgG Ab.*

1. INTRODUCTION

There has been tremendous progress in DNA sequencing technology to decrease cost and increase capacity. Yet technologies for proteomics have not advanced much in the last 30 to 40 years. Gene expression does not always correlate with protein abundance levels. And post-translational modifications (especially important in cell signaling) *cannot* be studied at the nucleic acid level. Improved methods for deriving high quality affinity reagents are needed to keep up with the 'explosion' of genomics information. Currently, less than half of around 6,000 routinely-used commercial antibodies (Abs) recognize only their specified targets (1, 2), and according to a publication in Nature Methods, over $350 million in biomedical research is wasted yearly in the United States alone due to poorly characterized Abs (3). To overcome this problem, there is a need for all Abs to be defined by their sequences and made recombinant, such that researchers will be able to use the same binding reagents under the same conditions (2-5). However, while a number of approaches for generating recombinant affinity reagents exist (6-9), the high costs and low throughput of current technologies represent significant roadblocks to the development of a comprehensive and broadly available resource of renewable affinity reagents.

Phage display is the most widely used method for selecting binding molecules from recombinant antibody libraries (10-13). In this technique, large numbers ($>10^{10}$) of single-chain variable fragment (scFv) or Fab antibodies are displayed on the surface of filamentous phage, and specific binders are selected by enrichment of binding phage during cycles of biopanning and propagation. High affinity phage Abs have been selected against numerous cellular proteins and small molecules (8). However, validation of the phage antibodies often requires early production of the cognate full-length immunoglobulin G (IgG). Phage library outputs can be sub-cloned to generate a full immunoglobulin, but this step is time-consuming and limits the number of clones that can be evaluated.

We have developed a system for the facile subcloning of phage display scFvs into IgG molecules *in vivo*, and for little cost. The system takes advantage of two genetic principles, recombination in *E.coli* (14) and splicing in mammalian cells (15, 16). A phage display vector contains both bacterial and mammalian regulatory regions that support antibody expression in bacteria and mammalian systems. The scFv is expressed as a fusion to the bacteriophage M13 gp3 gene in bacteria and converted to an IgG that can be expressed in mammalian cells by transducing the phagemid into a second *E.coli* F^+ strain. In that strain, the phiC31 serine integrase is used to fuse the heavy chain constant domain (C_H) from an acceptor plasmid to the heavy chain variable domain (V_H) and to introduce controlling elements upstream of the light chain variable domain (V_L) (Fig. 1). Positive selection for the recombination events is built into the system (Fig. 1). To generate the light-chain V_L-C_L fusion, mammalian splice sites flank the M13 gIII gene, allowing the V_L to be fused to the light chain constant domain (C_L) in mammalian cells (Fig. 1). Using the designed vector system, a single shuttle vector can be employed for phage library construction, phage display screening, and IgG antibody production in mammalian cells. Here, we demonstrate the utility of the features of this 'pMINERVA' subcloning system and proof-of-concept functional testing.

2. MATERIALS AND METHODS

2.1 Bacterial strains and vectors. The TG1 *E. coli* strain (F' (traD36 proAB+ lacIq lacZΔM15) supE thi-1 Δ(lac-proAB) Δ(mcrB-hsdSM)5, (rK-mK-) was purchased from Lucigen (Middleton, WI). The template phagemid, pAX1565, is a derivative of the phagemid, pAP-III_6 (17) with a single-chain variable fragment antibody (scFv) fused to coat protein III of bacteriophage M13. The scFv in pAX1565 is based on the monoclonal antibody, Herceptin (DrugBank # DB00072), and contains a $(Gly_4Ser)_3$ linker between the V_H and V_L domains.

*2.2 Molecular Biology***.** Standard cloning methods were used for cloning, sub-cloning, DNA extraction, protein purification, protein and DNA quantitation. Required mutagenesis was done using Agilent's site directed mutagenesis kits (18). Synthetic genes were constructed at GeneArt (Life Technologies, Carlsbad, CA). Plasmid pCDF-1b was purchased from Novagen (EMD Millipore, Billerica, MA). Restriction enzymes, ligases and polymerases were purchased from (New England Biolabs, Ipswich, MA) and used according to the manufacturer's recommendations. Electro-competent cells were purchased from Lucigen (Middletown, WI). CHO Free style and HEK293 Free Style cells were acquired from Life Technologies (Carlsbad, CA). Mammalian cell growth media was purchased from Life Technologies.

2.3 Construction of pDonor and pAcceptor plasmids. The phagemid, $pAPIII_6$ (17) was used as the template vector for the pDonor construct. The full-length M13 gpIII gene flanked by splice sites was synthesized by GeneArt (Life Technologies, Carlsbad, CA) and cloned into $pAPIII_6$ vector using *Sal* I and *Xho* I restriction sites. The constant region of the kappa light chain (C_L) and the SV40 late polyA sequence was synthesized by GeneArt and cloned into the above vector using the *Xho* I restriction site. Recombinants were sequence confirmed for directionality. A synthetic scFv based on the anti-Her2 antibody, Herceptin, with the 36 bp *attP* sequence for phiC31 as a linker between the variable heavy (V_H) and variable light (V_L) genes was cloned between the *Hind* III and *Sal* I sites in the vector. A *Mlu* I site was introduced by site-directed mutagenesis (QuikChange, Agilent Technologies, Santa Clara, CA) upstream of the phoA promoter in the vector. The synthetic promoters and leader peptides (Pro^{Cat} or Pro^{Splice}; See figure 5) were cloned between the *Mlu* I site and *Hind* III sites, replacing the phoA promoter upstream of the scFv gene. The origin and references for all genes and regions included in the pDonor vector are described in Table 1 (14, 17, 19-21).

To engineer the pAcceptor vector: The phiC31 gene (22) was synthesized by GeneArt (Life Technologies) and cloned into pCDF-1b (EMD Millipore, Billerica, MA) using *Not* I and *Avr* II restriction sites. The heavy chain acceptor construct consisting of the human EF1a promoter, IL-2 signal sequence, 36 bp *attB* sequence for phiC31 in frame with the signal sequence, human IgG1 constant domain, a ribosome binding site and spacer followed by the chloramphenicol gene (CAM), the BGH polyA sequence, and the T7 transcription terminator, was synthesized by GeneArt (Life Technologies). This 3.3 Kb DNA region was cloned into the pCDF-1b/phiC31 vector using the *Acc65* I restriction site to generate pAcceptor. Recombinants were sequence confirmed. The origin and references for all genes and regions included in the pAcceptor vector are described in Table 1.

2.4 Construction of IgG expression test vector. A single vector system for IgG expression consisting of the EF1a promoter upstream of the light chain (variable light (V_L) and constant light (C_L) regions) and a second mammalian promoter upstream of the heavy chain (variable heavy (V_H) and constant heavy (C_H) regions) was used to test individual components of the pMINERVA system. For the addition of the *attL* sequence between the signal sequence and V_L, the *attL* duplex was cloned into the IgG expression vector using the *Not* I restriction enzyme (NEB) and sequenced to screen for directionality. For the addition of the attR sequence between V_H and C_H, site-directed mutagenesis was performed to introduce the 36 base pair *attR* sequence following the manufacturer's protocol (QuikChange, Agilent Technologies).

To introduce the M13 gpIII sequence flanked by splice sites between V_H and C_H of the IgG expression test vector, the gene was synthesized by GeneArt (Life Technologies) and cloned into the IgG expression test vector using *BsiW* I restriction sites. Recombinants were screened for directionality by sequencing.

To exchange the mammalian promoter driving the light chain of the IgG with Pro^{cat} or Pro^{splice}, the promoter sequences were synthesized by GeneArt (Life Technologies) and cloned into the IgG expression test vector using *Mlu* I and *Not* I restriction sites. To exchange the mammalian promoter driving the heavy chain of the IgG with Pro^{cat}, the Pro^{cat} sequence was synthesized by GeneArt (Life Technologies) and cloned into the IgG expression test vector using *EcoR* V and *Nde* I restriction sites.

To exchange V_H and V_L regions in the IgG vector, the V_H gene was PCR amplified from a commercially available human IgG heavy chain expression vector (Invivogen, San Diego, CA) and cloned into the IgG expression test vector using *Nde* I and *BsiW* I restriction enzymes (NEB) and T4 DNA ligase (NEB). The entire light chain (V_L and C_L) and a piece of the SV40 late polyA sequence were PCR amplified from a commercially available human IgG light chain expression vector (Invivogen) and cloned into the IgG expression test vector containing the correct V_H using *Not* I and *Hpa* I restriction sites. Recombinants were sequence confirmed.

*2.5 DNA transformation***.** Transformation reactions were set up as follows: 0.5 µl ligation product was mixed with 50 µl of TG1 electrocompetent cells (Lucigen) and added to a 0.1 cm gap cuvette. DNA was electroporated into bacterial cells using a Gene Pulser (Bio-Rad, at the following settings: 1.6 kV, 200 ohms, 25 µF). One ml of recovery media was added and the electroporated cells were transferred to 14 ml culture tubes and shaken at 37 °C. After 1 hr, one hundred µl of each of the dilutions was plated onto LB plates containing ampicillin (100 µg/ml), and the plates were incubated overnight at 37°C. Chemically competent NEB5α cells (NEB) were transformed according to the manufacturer's protocol. Briefly, 50 µL of cells were incubated with 1 µL of DNA for 30 min on ice, heat shocked at 42ºC for 30 sec and 250 µL of recovery media was added. Cells were incubated while shaking for 1 hr at 37ºC and 100 µL was plated on LB plates with 100 µg/mL of ampicillin.

*2.6 Monoclonal phage ELISA***.** For vectors containing the Pro^{Splice} promoter, a single colony of transformed TG1 cells was picked off of a TYE/Amp/Glucose plate into a 2 ml starter culture of 2YT supplemented with ampicillin (100 µg/ml) and 1% glucose. Single colonies containing Pro^{cat} vectors were picked off of LB plates containing ampicillin (100 µg/ml) into a 2 ml starter culture of LB supplemented with ampicillin (100 µg/ml). Cultures were incubated at 37°C for 2-3 hours with shaking, diluted into 50 ml of their respective media, and further incubated at 37°C with shaking until an absorbance at 600 nm of 0.4 was reached. Once the cultures reached mid-log phase, 10 ml were transferred into individual 15 ml conical tubes, and 5 µl of KM13 helper phage was added (23). After a 30 min incubation at 37°C without shaking, tubes were centrifuged at 2000 x g for 10 minutes and the supernatants were aspirated. Pellets were resuspended in 50 ml of appropriate media (Pro^{Splice}: 2YT supplemented with ampicillin (100 µg/ml), kanamycin (50 µg/ml) and 0.1% glucose; Pro^{Cat}: LB supplemented with ampicillin (100 µg/ml) and kanamycin (50 µg/ml)) and incubated at 30°C with shaking overnight. After overnight incubation, the cultures were centrifuged at 15,000 x g for 10 minutes. Supernatants were transferred to an ELISA plate pre-coated with the specific target antigen.

For coating of the ELISA plates, 100 µl/well of antigen diluted to 2.5 µg/ml in 1X PBS was added to maxisorp ELISA plates (ThermoScientific, NUNC) and incubated overnight at 4°C. Wells were coated with either specific antigen or with non-relevant protein to test for non-specific binding. The wells were washed three times with PBS (250 µl/well) and blocked for 1 hr with 2% nonfat dry milk in PBS (MPBS). The wells were washed for a total of three times with PBS. 100 µl of undiluted supernatant containing phage was added to the ELISA plate and incubated for 1 hr at room temperature. Wells were washed three times with PBS containing 0.01% Tween (PBS-T) followed by a 1 hr room temperature incubation with anti-M13 monoclonal antibody conjugated to Horseradish Peroxidase (HRP; GE Healthcare, Piscataway,

NJ). The ELISA was developed by adding 100 µl TMB Ultra (Pierce, cat# 34029) to each well, and the reactions were stopped with 50 µl 2M H_2SO_4. Plates were read at 450 nm in a standard plate reader and fold over background (FOB) was calculated by dividing the OD450 of a well containing phage on specific antigen by the signal of the same phage on non-relevant protein.

*2.7 Production and purification of soluble IgG antibodies***.** IgG expression vectors were transiently transfected into mammalian HEK293 (Life Technologies; 293-F; cat# R790-07) suspension cells under sterile conditions in a cell culture ventilation hood. On the day prior to transfection, the HEK293 suspension cells were diluted back to $0.7x10^6$ cells/ml in 30 ml total volume using Freestyle 293 Expression Medium (Life Technologies). On the day of transfection, 24 µg of sterile IgG expression vector DNA was diluted in 3 ml Freestyle 293 Expression Medium and vortexed for 10 seconds. To the diluted DNA, 24 µl of FectoPRO transfection reagent (Polyplus-transfection, New York, NY) was added and vortexed for 10 sec. The DNA-transfection reagent mixture was incubated at room temperature for 10 min. The DNA-transfection reagent mixture was added to the 30 ml culture of HEK293 suspension cells that were seeded the day prior. Post-transfection the HEK293 cells were incubated at 37˚C in 8% CO_2 with shaking at 130 rpm for 72 hours, at which point 15 ml of fresh Freestyle 293 Expression Medium was added. The cultures continued incubating at 37˚C and 8% CO_2, shaking at 130 rpm for an additional 48-72 hours. The transfected cells were centrifuged at 1,000 x g for 10 min and the IgG-containing supernatants were transferred to new, sterile tubes. The IgG antibodies were purified from the clarified supernatants with 0.5 ml Protein A bead slurry. The Protein A beads were washed three times with PBS and incubated with 45 mL of IgG supernatant for one hour while rocking at 4˚C. The Protein A bead + IgG mixture was poured into an empty 2 ml chromatography column and loaded via gravity flow. The columns were washed twice with 3 mL of PBS supplemented with 1 mM PMSF and eluted with 1.5 mL of 0.1 M glycine at pH 3.0 and neutralized with 60 µl of 1 M Tris buffer at pH 9.0. The eluate was analyzed by SDS-PAGE under reducing and non-reducing conditions. Final purified IgG was quantitated by Coommassie Plus assay kit (ThermoFisher, cat# 23236).

*2.8 Human IgG ELISA on protein***.** For biotinylated antigens, ELISA plates were coated with 100 µL per well of neutravidin at 10 µg/mL and incubated overnight at 4ºC. The wells were washed three times with PBS (250 µl/well) and blocked for 1 hr at room temperature with 3% bovine serum albumin (BSA) in PBS. The wells were incubated with 1 µg/mL of biotinylated peptide (either specific target or non-relevant protein) for 1 hr at room temperature after which they were washed three times with PBS (250 µL/well) and blocked again for 1 h with 3% BSA in PBS. After washing the wells with PBS 3x, the purified IgG antibodies were added at 1 µg/ml in 3% BSA/PBS and incubated for 1 hr at room temperature. The wells were washed four times with PBS + 0.01% Tween (PBS-T) and incubated with anti-human-IgG-HRP (ThermoFisher, cat# AHI0404) diluted 1:5,000 in block for 1 hr at room temperature. The wells were washed three times with PBS-T and the HRP-conjugated secondary antibody was detected with TMB reagent (100 µL/well) with a 2-3 minute incubation. The reaction was stopped with 2M H_2SO_4 (50 µL/well), absorbance was read at 450 nm and FOB was calculated by dividing the OD450 of a well containing IgG on specific antigen by the signal of the same IgG on non-relevant protein.

2.9 Human IgG ELISA on mammalian cells. HEK293 cells stably expressing a target antigen and HEK293 cells not expressing the antigen were added to V-bottom ELISA plates (Phenix, cat# MPG-651101) at a concentration of $2.3x10^6$ cell/mL (0.15 mL per well). Cells were blocked with 3% BSA for 30 min at room temperature after which they were centrifuged at 500x g for 4 min and the supernatant was removed. Purified IgG was added to the cells at 1 µg/mL (100 µL per well) and incubated while shaking for 1 hr at room temperature. Cells were washed 2x with

PBS, centrifuging the plate in between each wash at 500 x g for 4 min. Cells were incubated while shaking with anti-human-HRP at 1:5,000 in block for 1 hr at room temperature, Cells were washed 2x with PBS and transferred to a pre-blocked V-bottom plate and washed once more. 100 μL of luminescence reagent (Piece, cat #37069) was added to the cells/well and transferred to a white bottom ELISA plate (NUNC, cat# 436110). After a 2 min incubation the plates were read with a luminescence detector and FOB was calculated by dividing the relative luminescence unit (RLU) of a well containing IgG on HEK293 cells expressing the target antigen by the RLU of the same IgG on HEK293 cells not expressing the target antigen.

*2.10 Testing phiC31 recombinase function***.** Competent TG1 cells containing the spectinomycin-resistant pAcceptor plasmid were prepared according to the New England Biolabs on-line protocol (https://www.neb.com/protocols/2012/06/21/making-your-own-electrocompetent-cells). The cells were transformed with the pDonor or a mock-recombined control plasmid following standard electroporation procedure (described above). Transformants were plated on LB media containing ampicillin (100 µg/ml) or chloramphenicol (10 µg/ml). The ratio of ampicillin to chloramphenicol-resistant transformants was determined.

3. RESULTS

*3.1 Cloning of phiC31 phage integrase and in vivo testing of functionality using a promotorless CamR gene polycistronic message***.** Phage-encoded serine integrases mediate directionally regulated site-specific recombination between short *attP* and *attB* DNA sites without host factor requirements (14). The phiC31 serine integrase can be used to induce recombination between two plasmids in *E. coli* (22). We designed a system that uses phiC31 to fuse the heavy chain variable domain (V_H) on our phagemid vector (pDonor) to the heavy chain constant domain (C_H) on a second plasmid (pAcceptor; Fig. 1). For this approach to be feasible, the linker between the variable heavy (V_H) and variable light (V_L) domains of the scFv must contain a 36 bp phiC31 integrase recognition site (*attP* or *attB*) that is able to function as both a peptide linker in the scFv and a functional substrate for phiC31 integrase.

We replaced the wild-type $(Gly_4Ser)_3$ linker between the V_H and V_L of a model scFv (anti-Her2) in our phagemid vector with the phiC31 integrase site, *attP* or *attB*. The phage-scFv were produced in *E.coli* and tested for function in a phage ELISA against the target antigen, Her2, or a non-relevant control protein. As shown in figure 2, phage containing the *attP* linker bound selectively to Her2 with comparable activity to phage containing the wild-type linker (Fig. 2a). Phage with the *attB* site as a linker in the scFv was unable to bind to its specific antigen (data not shown).

Successful recombination between the *attP* site on the phagemid and the *attB* site on the acceptor plasmid generates a 36 base pair recombined phiC31 site (*attR*) at the hinge junction of the variable heavy domain and constant heavy domain of the immunoglobulin (Fig. 1). In addition, a 36 base pair recombined phiC31 site (*attL*) is produced between the leader peptide and variable domain of the light chain (V_L) (Fig. 2b). Tested IgGs containing the *attR* or *attL* sequences at these junctions, respectively, express at wild-type levels and recognize their respective target antigens (Fig. 2c, and Fig. 2d).

To test the function of phiC31 in our system, the gene was cloned into a plasmid (derived from pCDF-1b; Novagen) under the control of a constitutive *E. coli* promoter (pCam; pACYC184) and transformed into the *E. coli* strain, TG1. A promoterless chloramphenicol resistance (*cam*R) gene was placed 3' of the phiC31 integrase *att*B site on the same plasmid (pAcceptor). Successful integration of the phagemid (pDonor) into the pAcceptor plasmid places an *E. coli*

promotor upstream of the pAcceptor *cam*R gene and produces a bicistronic gene-pair composed of the immunoglobulin heavy chain and *cam*R genes (Fig. 3a). The cells that are harboring this resulting co-integrated plasmid thereby become chloramphenicol resistant. The phiC31-mediated integration efficiency was tested in *E. coli* by transforming a pAcceptor TG1 strain with the pDonor plasmid and measuring the percent of ampicillin-resistant transformants converting to *cam*R. Conversion would only occur as a result of successful co-integration. The ratio of chloramphenicol-resistant (*cam*R) transformants to ampicillin-resistant (*amp*R) transformants was compared with the ratio of *cam*R to *amp*R transformants using an already co-integrated plasmid in TG1 cells. Our results (Fig. 3b) demonstrate that the phiC31 integrase is able to site-specifically recombine the pDonor vector into a pAcceptor vector at >90% efficiency. No chloramphenicol-resistant colonies were obtained when the pDonor lacked the *attP* site (data not shown).

*3.2 Design and testing of synthetic intron containing M13 gpIII***.** While phiC31-mediated recombination is used to join the V_H to the C_H domain in our pMINERVA system, the light chain variable domain (V_L) domain in the scFv also needs to be fused to the light chain constant domain (C_L) (Fig. 1). In eukaryotic cells, large sequences can be efficiently removed from mRNA by the RNA splicing machinery. We designed a synthetic intron containing the M13 gpIII gene flanked by consensus 5' and 3' splice site sequences. In *E. coli*, the scFv is expressed on the surface of bacteriophage M13 as a genetic fusion to the gpIII gene. In mammalian cells, the intron containing the gpIII gene is excised, creating a fusion of the V_L to the C_L (Fig. 1). In the absence of proper splicing, the C_L will be out of frame with the V_L, and functional IgG cannot be produced. As shown in figure 4a, introduction of splice sites between the scFv and gpIII gene does not interfere with phage production or function. In HEK-293 cells, IgG expression from the intron-containing vector is comparable to expression from the original construct containing no intron, indicating that splicing is efficient (Fig. 4b, top). Both IgGs tested are functional and show comparable activity in an ELISA against the target antigen (Fig. 4b, bottom).

3.3 Design and testing of dual-functional promoters: Pro^{splice} and Pro^{cat} expression in mammalian cells and E. coli. To enable scFv production in bacteria and IgG production in mammalian cells, a dual-functional promoter that supports both *E. coli* and mammalian expression is required. We tested two different promoter systems here: a promoter intron-splicing system (Pro^{splice}) and a catenated polyfunctional promoter system (Pro^{cat}) (Fig. 5). In the Pro^{splice} system, the EF1a promoter is followed by the mammalian IgG heavy chain secretion signal sequence that contains an intron. The lac promoter/operator and bacterial signal peptide is contained within the mammalian intron, and the bacterial signal sequence overlaps the splice acceptor site. In *E. coli*, transcription from the bacterial lac promoter/operator within the mammalian intron results in expression of the scFv in the bacterial periplasm. In mammalian cells, splicing removes the bacterial regulatory sequences located in the intron, generating the mammalian signal sequence.

In the catenated promoter system (Pro^{cat}) developed by the Weiner team (24) the mammalian promoter is followed by a polyhedron promoter (for expression in insect cells) and bacterial promoter. The ATGs are removed from the downstream polyhedron and bacterial promoter such that the first ATG fMet start site for bacterial, insect and mammalian expression is identical. In this case, the same signal sequence is used for all three organisms. The Pro^{cat} developed here uses the EF1a promoter for mammalian expression, the phoA promoter for *E. coli* expression, and the IL-2a signal peptide for secretion in both systems. Phage displaying the model scFv is produced from both promoter systems at equivalent levels (Fig. 5c). In addition, the IgG yield from the dual promoter systems in HEK-293 cells is equivalent to the yield using the EF1a promoter alone (Fig. 5d and Table 2).

By combining together all of these elements, the pMINERVA system completely eliminates the labor-intensive subcloning and DNA sequence confirmation steps currently needed to convert a scFv into a functional IgG antibody.

4. DISCUSSION

We are building a platform that takes advantage of fundamental molecular biology principles to enable the rapid identification of specific Abs. Our screening platform is being developed with single-chain variable fragment Abs (scFvs) but the technology is applicable to other scaffolds, including both Fabs (4) and yeast display libraries (12, 25-27). Here, we describe a vector system that enables high-throughput conversion of Ab fragments to full length immunoglobulin G (IgG) molecules that can be directly validated in standard immunoassays.

Our system relies on a phage integrase, phiC31, to induce site-specific recombination between a donor and acceptor plasmid in *E. coli* to generate a fusion of the antibody heavy chain variable domain (V_H) to the heavy chain constant domain (C_H) and to introduce regulatory elements upstream of the variable light chain gene (V_L). Recombinases such as Cre and FLP are commonly used for genome engineering in multiple organisms (28-30). However, the efficiency of integration with these enzymes is low because the enzymes perform both integration and excision with the same target sites. Some phage integrases, including phiC31, perform only the integration reaction and require accessory factors for the reverse reaction (22). We have shown here that phiC31 is greater than 90% efficient at inducing recombination between two plasmids in *E. coli*. There are many additional known serine- and tyrosine-integrases, such as BxB1, lambda, etc. that also mediate unidirectional, site-specific recombination and could be evaluated in this system (31).

The phiC31 integrase has a 36 base pair (i.e., equivalent to 12 amino acids) recognition sequence. Our system requires incorporation of these integrase attachment sites in the linker of the scFv, and subsequently, at the hinge junction between the heavy chain variable and constant domains and between the signal sequence and light chain variable domain. We show that these 12 amino acid sequences do not interfere with phage production or IgG expression and function of the specific recombinant antibodies tested in this study. If needed, attP or attB mutations that would alter the resulting linker peptide to a more flexible linker sequence could be tested for their ability to support efficient integration. Additionally, in order to generate a “seamless” IgG, it may be possible to incorporate splice sites flanking the integrase attachment sites in the vector. We successfully used mammalian splicing in this study to excise the M13 gpIII gene and fuse the light chain variable domain (V_L) to the light chain constant domain (C_L). Intron splicing has further been shown to enhance the expression of recombinant proteins in mammalian cells (32).

We describe two dual expression promoter systems here that both worked well for *E. coli* and mammalian expression. One system utilizes mammalian splicing to excise bacterial expression elements, and the second system uses catenated promoters for mammalian, insect cell, and *E. coli* expression. The catenated promoter system requires the same signal peptide to work in both *E. coli* and mammalian cells. We show here that the IL-2 signal sequence is able to function as a secretion signal in bacteria. Other signal peptides that work in both mammalian and bacterial systems have been reported (15) and can be tested for improved expression in our system. We utilized the human EF1a promoter to build our dual function promoter system, but other promoters that enable high-level protein expression in mammalian cells, including the CMV promoter, can be tested. Our catenated promoter system (24) includes the polyhedron

promoter such that insect cells could be used as an alternative to HEK-293 cells for high-throughput IgG production.

Our current system is modeled on a human IgG1 antibody. However, additional pAcceptor plasmids can be constructed to extend the utility of the system. Acceptor constructs containing other human immunoglobulin isoforms (including IgG2, IgG3, IgG4, IgM) can be generated, as well as acceptors containing the constant domains for mouse or rabbit antibodies. Different types of 3' fusions could also be generated, including enzyme fusions (33-37), protein purification tags (38-41), labels (42-44), and fusion-tags that can direct proteins to the endosome (45-54). We believe that this will be the ideal vector system for producing recombinant antibodies.

To capitalize on the investment and availability of sequenced genomes, it will be essential to devise methods for rapidly and inexpensively generating affinity reagents at a proteome-scale. Combining the pMINERVA system with our previous improvements, AXM mutagenesis (55), Phage ESCape (56), and phage libraries using entirely pre-defined complementarity determining regions (CDR) (57) we present a platform for the rapid generation of recombinant monoclonal Abs in as few as three weeks and at a cost comparable to that of polyclonal antibodies.

Acknowledgements

We would like to thank Chris McLeod and Janet Verney for their useful discussions and help over the last number of years. We would also like to thank Andrew Bradbury for his valuable discussions and prior work. This work was supported by Small Business Innovative Research (SBIR) grants from the National Institutes of Health [1R43GM1122204-01A1].

REFERENCES

1. **Prinz F, Schlange T, Asadullah K.** 2011. Believe it or not: how much can we rely on published data on potential drug targets? Nat Rev Drug Discov **10:**712.
2. **Bradbury A, Pluckthun A.** 2015. Reproducibility: Standardize antibodies used in research. Nature **518:**27-29.
3. **Bradbury AM, Pluckthun A.** 2015. Antibodies: validate recombinants once. Nature **520:**295.
4. **Miersch S, Li Z, Hanna R, McLaughlin ME, Hornsby M, Matsuguchi T, Paduch M, Saaf A, Wells J, Koide S, Kossiakoff A, Sidhu SS.** 2015. Scalable high throughput selection from phage-displayed synthetic antibody libraries. J Vis Exp doi:10.3791/51492**:**51492.
5. **Taussig MJ, Schmidt R, Cook EA, Stoevesandt O.** 2013. Development of proteome-wide binding reagents for research and diagnostics. Proteomics Clin Appl **7:**756-766.
6. **Dubel S, Stoevesandt O, Taussig MJ, Hust M.** 2010. Generating recombinant antibodies to the complete human proteome. Trends Biotechnol **28:**333-339.
7. **Finlay WJ, Bloom L, Cunningham O.** 2011. Phage display: a powerful technology for the generation of high specificity affinity reagents from alternative immune sources. Methods Mol Biol **681:**87-101.
8. **Hoogenboom HR, de Bruine AP, Hufton SE, Hoet RM, Arends JW, Roovers RC.** 1998. Antibody phage display technology and its applications. Immunotechnology **4:**1-20.
9. **Stoevesandt O, Taussig MJ.** 2012. European and international collaboration in affinity proteomics. N Biotechnol **29:**511-514.
10. **Lee CV, Sidhu SS, Fuh G.** 2004. Bivalent antibody phage display mimics natural immunoglobulin. J Immunol Methods **284:**119-132.
11. **Marks JD, Hoogenboom HR, Bonnert TP, McCafferty J, Griffiths AD, Winter G.** 1991. By-passing immunization. Human antibodies from V-gene libraries displayed on phage. J Mol Biol **222:**581-597.
12. **Schofield DJ, Pope AR, Clementel V, Buckell J, Chapple S, Clarke KF, Conquer JS, Crofts AM, Crowther SR, Dyson MR, Flack G, Griffin GJ, Hooks Y, Howat WJ, Kolb-Kokocinski A, Kunze S, Martin CD, Maslen GL, Mitchell JN, O'Sullivan M, Perera RL, Roake W, Shadbolt SP, Vincent KJ, Warford A, Wilson WE, Xie J, Young JL, McCafferty J.** 2007. Application of phage display to high throughput antibody generation and characterization. Genome Biol **8:**R254.
13. **Sheets MD, Amersdorfer P, Finnern R, Sargent P, Lindquist E, Schier R, Hemingsen G, Wong C, Gerhart JC, Marks JD.** 1998. Efficient construction of a large nonimmune phage antibody library: the production of high-affinity human single-chain antibodies to protein antigens. Proc Natl Acad Sci U S A **95:**6157-6162.
14. **Groth AC, Calos MP.** 2004. Phage integrases: biology and applications. J Mol Biol **335:**667-678.
15. **Tesar D, Hotzel I.** 2013. A dual host vector for Fab phage display and expression of native IgG in mammalian cells. Protein Eng Des Sel **26:**655-662.
16. **Quinlan BD, Gardner MR, Joshi VR, Chiang JJ, Farzan M.** 2013. Direct expression and validation of phage-selected peptide variants in mammalian cells. J Biol Chem **288:**18803-18810.
17. **Haidaris CG, Malone J, Sherrill LA, Bliss JM, Gaspari AA, Insel RA, Sullivan MA.** 2001. Recombinant human antibody single chain variable fragments reactive with Candida albicans surface antigens. J Immunol Methods **257:**185-202.

18. **Weiner MP, Costa GL, Schoettlin W, Cline J, Mathur E, Bauer JC.** 1994. Site-directed mutagenesis of double-stranded DNA by the polymerase chain reaction. Gene **151:**119-123.
19. **Saito M, Takano T, Nishimura T, Kohara M, Tsukiyama-Kohara K.** 2015. 3beta-hydroxysterol delta24-reductase on the surface of hepatitis C virus-related hepatocellular carcinoma cells can be a target for molecular targeting therapy. PLoS One **10:**e0124197.
20. **Holt LJ, Enever C, de Wildt RM, Tomlinson IM.** 2000. The use of recombinant antibodies in proteomics. Curr Opin Biotechnol **11:**445-449.
21. **Matz RL, Erickson B, Vaidyanathan S, Kukowska-Latallo JF, Baker JR, Jr., Orr BG, Banaszak Holl MM.** 2013. Polyplex exposure inhibits cell cycle, increases inflammatory response, and can cause protein expression without cell division. Mol Pharm **10:**1306-1317.
22. **Thorpe HM, Smith MC.** 1998. In vitro site-specific integration of bacteriophage DNA catalyzed by a recombinase of the resolvase/invertase family. Proc Natl Acad Sci U S A **95:**5505-5510.
23. **Kristensen P, Winter G.** 1998. Proteolytic selection for protein folding using filamentous bacteriophages. Fold Des **3:**321-328.
24. **Kadwell SH, Bisi, J. E., Brown, C.E., Cogswell, J.P., Colapietro, A.-M., Luther, M.A., Neill, S.D., Richards, C.A., Weiner, M.P.** 2001. The Admid System: Generation of Recombinant Adenoviruses by Tn7-Mediated Transposition in *E. coli.* Cloning and Expression Vectors for Gene Function Analysis. *Edited by* Weiner and Lu (ed).
25. **Ferrara F, D'Angelo S, Gaiotto T, Naranjo L, Tian H, Graslund S, Dobrovetsky E, Hraber P, Lund-Johansen F, Saragozza S, Sblattero D, Kiss C, Bradbury AR.** 2015. Recombinant renewable polyclonal antibodies. MAbs **7:**32-41.
26. **Sheehan J, Marasco WA.** 2015. Phage and Yeast Display. Microbiol Spectr **3:**AID-0028-2014.
27. **Zhao L, Qu L, Zhou J, Sun Z, Zou H, Chen YY, Marks JD, Zhou Y.** 2014. High throughput identification of monoclonal antibodies to membrane bound and secreted proteins using yeast and phage display. PLoS One **9:**e111339.
28. **Sauer B.** 1994. Site-specific recombination: developments and applications. Curr Opin Biotechnol **5:**521-527.
29. **Diaz V, Rojo F, Martinez AC, Alonso JC, Bernad A.** 1999. The prokaryotic beta-recombinase catalyzes site-specific recombination in mammalian cells. J Biol Chem **274:**6634-6640.
30. **O'Gorman S, Fox DT, Wahl GM.** 1991. Recombinase-mediated gene activation and site-specific integration in mammalian cells. Science **251:**1351-1355.
31. **Fogg PC, Colloms S, Rosser S, Stark M, Smith MC.** 2014. New applications for phage integrases. J Mol Biol **426:**2703-2716.
32. **Chang CY, Hong WW, Chong P, Wu SC.** 2006. Influence of intron and exon splicing enhancers on mammalian cell expression of a truncated spike protein of SARS-CoV and its implication for subunit vaccine development. Vaccine **24:**1132-1141.
33. **Benink HA, Urh M.** 2015. HaloTag technology for specific and covalent labeling of fusion proteins. Methods Mol Biol **1266:**119-128.
34. **Blackstock D, Chen W.** 2014. Halo-tag mediated self-labeling of fluorescent proteins to molecular beacons for nucleic acid detection. Chem Commun (Camb) **50:**13735-13738.
35. **Duong-Ly KC, Gabelli SB.** 2015. Affinity Purification of a Recombinant Protein Expressed as a Fusion with the Maltose-Binding Protein (MBP) Tag. Methods Enzymol **559:**17-26.
36. **Maeda K, Poletto M, Chiapparino A, Gavin AC.** 2014. A generic protocol for the purification and characterization of water-soluble complexes of affinity-tagged proteins and lipids. Nat Protoc **9:**2256-2266.

37. **Wille T, Barlag B, Jakovljevic V, Hensel M, Sourjik V, Gerlach RG.** 2015. A gateway-based system for fast evaluation of protein-protein interactions in bacteria. PLoS One **10:**e0123646.
38. **Bornhorst JA, Falke JJ.** 2000. Purification of proteins using polyhistidine affinity tags. Methods Enzymol **326:**245-254.
39. **Kimple ME, Brill AL, Pasker RL.** 2013. Overview of affinity tags for protein purification. Curr Protoc Protein Sci **73:**Unit 9 9.
40. **Pina AS, Batalha IL, Roque AC.** 2014. Affinity tags in protein purification and peptide enrichment: an overview. Methods Mol Biol **1129:**147-168.
41. **Wood DW.** 2014. New trends and affinity tag designs for recombinant protein purification. Curr Opin Struct Biol **26:**54-61.
42. **Phillips GJ.** 2001. Green fluorescent protein--a bright idea for the study of bacterial protein localization. FEMS Microbiol Lett **204:**9-18.
43. **Medina-Kauwe LK, Chen X.** 2002. Using GFP--ligand fusions to measure receptor-mediated endocytosis in living cells. Vitam Horm **65:**81-95.
44. **March JC, Rao G, Bentley WE.** 2003. Biotechnological applications of green fluorescent protein. Appl Microbiol Biotechnol **62:**303-315.
45. **Iwamoto M, Bjorklund T, Lundberg C, Kirik D, Wandless TJ.** 2010. A general chemical method to regulate protein stability in the mammalian central nervous system. Chem Biol **17:**981-988.
46. **Miyazaki Y, Imoto H, Chen LC, Wandless TJ.** 2012. Destabilizing domains derived from the human estrogen receptor. J Am Chem Soc **134:**3942-3945.
47. **Banaszynski LA, Chen LC, Maynard-Smith LA, Ooi AG, Wandless TJ.** 2006. A rapid, reversible, and tunable method to regulate protein function in living cells using synthetic small molecules. Cell **126:**995-1004.
48. **Grimley JS, Chen DA, Banaszynski LA, Wandless TJ.** 2008. Synthesis and analysis of stabilizing ligands for FKBP-derived destabilizing domains. Bioorg Med Chem Lett **18:**759-761.
49. **Chu BW, Banaszynski LA, Chen LC, Wandless TJ.** 2008. Recent progress with FKBP-derived destabilizing domains. Bioorg Med Chem Lett **18:**5941-5944.
50. **Hagan EL, Banaszynski LA, Chen LC, Maynard-Smith LA, Wandless TJ.** 2009. Regulating protein stability in mammalian cells using small molecules. Cold Spring Harb Protoc **2009:**pdb prot5172.
51. **Maynard-Smith LA, Chen LC, Banaszynski LA, Ooi AG, Wandless TJ.** 2007. A directed approach for engineering conditional protein stability using biologically silent small molecules. J Biol Chem **282:**24866-24872.
52. **Cheeseman LP, Harry EF, McAinsh AD, Prior IA, Royle SJ.** 2013. Specific removal of TACC3-ch-TOG-clathrin at metaphase deregulates kinetochore fiber tension. J Cell Sci **126:**2102-2113.
53. **Hirst J, Borner GH, Antrobus R, Peden AA, Hodson NA, Sahlender DA, Robinson MS.** 2012. Distinct and overlapping roles for AP-1 and GGAs revealed by the "knocksideways" system. Curr Biol **22:**1711-1716.
54. **Robinson MS, Hirst J.** 2013. Rapid inactivation of proteins by knocksideways. Curr Protoc Cell Biol **61:**15 20 11-17.
55. **Holland EG, Buhr DL, Acca FE, Alderman D, Bovat K, Busygina V, Kay BK, Weiner MP, Kiss MM.** 2013. AXM mutagenesis: an efficient means for the production of libraries for directed evolution of proteins. J Immunol Methods **394:**55-61.
56. **Kiss MM, Babineau EG, Bonatsakis M, Buhr DL, Maksymiuk GM, Wang D, Alderman D, Gelperin DM, Weiner MP.** 2011. Phage ESCape: an emulsion-based approach for the selection of recombinant phage display antibodies. J Immunol Methods **367:**17-26.

57. **Batonick M, Holland EG, Busygina V, Alderman D, Kay BK, Weiner MP, Kiss MM.** 2015. Platform for high-throughput antibody selection using synthetically-designed antibody libraries. N Biotechnol doi:10.1016/j.nbt.2015.11.005.
58. **Groth AC, Olivares EC, Thyagarajan B, Calos MP.** 2000. A phage integrase directs efficient site-specific integration in human cells. Proc Natl Acad Sci U S A **97**:5995-6000.

Figure legends

Table 1. Description of genetic elements used in the pDonor and pAcceptor system. The source of each element used in the pMINERVA system is indicated. Also given is whether the genetic element was synthesized or cloned from an existing plasmid.

Table 2. Yield analysis of IgG expression test vector. Eleven scFv V_H and V_L sequences were cloned into a single IgG expression test vector using the Pro^{cat} promoter to drive the heavy chain. The IgG vectors were transfected into HEK293 cells and the secreted IgG proteins were purified from the supernatant six days post-transfection. The amount of final purified IgG was determined and yields were calculated based on transfecting 100mL cell culture volumes.

Fig. 1. Description of the pMINERVA transformer system. The scFv Abs encoded on the pDonor vector as M13gpIII-fusions can be screened in a phage display biopanning procedure to identify a phagemid encoding a scFv with the desired biophysical properties. This phagemid is transduced into an *E. coli* strain expressing phiC31 integrase and harboring an IgG acceptor vector. The product of the recombination event introduces a polyadenylation signal site adjacent to the 3' end of the C_H gene. Furthermore, the recombination event introduces both a mammalian promoter and functional protein initiation site 5' to the V_L gene. Of special note, the linker between the V_H and V_L domains of the scFv is composed of a phiC31 36-bp *att*P site that is able to function as both a: i) peptide linker between the heavy and light variable domains, and ii) 36-bp functional substrate for phiC31 integrase. [P_{mam}, mammalian promoter; 5'ss and 3'ss, splice signals; V_L, variable section of the light chain; V_H, variable section of the heavy chain; gp3, phage M13 gene3 product; $P_{E.coli}$, *E.coli* promoter; C_H or F_c, constant region of the heavy chain; *att*B, *att*P, substrates for an integrase gene; *att*R and *att*L, products of an integrase gene; polyA, polyadenylation sequence; Cam^S, CamR, chloramphenicol resistance gene without and with a promoter, respectively; Pro^{splice}, Pro^{cat}, dual-function promoter-types (see text and Fig 5 for details); RBS, ribosome binding site].

Fig. 2. Expression and function of scFv and IgGs containing phiC31 integrase sites. (A) The same scFv with two different linker sequences, WT $(Gly_4Ser)_3$ or the phiC31 *att*P site in reading frame 2, was produced. Both phage-scFvs were tested in an ELISA against the purified target protein or a non-relevant antigen control. Anti-M13 antibody that is conjugated to Horseradish peroxidase (HRP) was used for phage detection and the enzyme linked immunosorbent assay (ELISA) was developed with Ultra Tetramethylbenzidine (TMB) reagent. The fold over background (FOB), which is the signal against target over the signal against non-relevant control, is shown for each phage tested. Error bars represent the standard deviation of phage binding tested in triplicate (B) IgG molecule modeled with attL and attR. The *attL* (thick blue loop) and *attR* (thick red loop) peptides are inserted schematically in a typical human IgG1 molecule (PDB ID: 1hzh) shown as ribbons (heavy chain: white, light chain: cyan). (C) The same IgG, with either no linker or the recombined phiC31 integrase *attL* site between the IL2 signal sequence (ss) and V_L, was produced. In parallel, expression of the wild-type IgG was compared to expression of the same IgG with the recombined phiC31 integrase site *attR* site between V_H and C_H. A coommassie stained SDS-PAGE gel is shown. (D) An IgG with attL between the IL2ss and the V_L (top graph) and an IgG with attR between V_H and C_H (bottom

graph) were tested for binding to both a specific and a non-relevant target antigen in a cell ELISA. The binding of both molecules was compared to the binding of a wildtype IgG. Both ELISAs used anti-human-HRP to detect the IgGs and the ELISA was developed with Ultra TMB reagent. The ODs at 450nm are shown.

Fig. 3. Positive selection of phiC31 integrase activity. (A) Two plasmids pDonor and pAcceptor, each having a needed component in *trans* for a functional *cam*R gene were constructed. The *att*P site (underlined) in pDonor was flanked upstream by the *E. coli* 5' controlling elements and. The *att*B site (lower case DNA sequence) in pAcceptor was placed 5' of the promoter-less *cam*R gene. Successful recombination between the *att*P and *att*B sites on the two plasmids in the presence of the phiC31 integrase (blue dashed line) generates the co-integrant (pMINERVA; bottom sequence) and fuses an *E. coli* promoter in front of the bi-cistronic heavy chain-Cam^R message. (B) Competent TG1 cells containing pAcceptor were transformed with pDonor or a control mock-recombined vector and grown on plates containing ampicillin or chloramphenicol. The ratio of colonies on the ampicillin plates to the chloramphenicol plates was calculated.

Fig. 4. Expression and function of scFv and IgGs containing splice sites flanking the M13 gpIII gene. (A) Phage-scFv fusions with and without splice sites flanking the M13 gpIII gene in pDonor were produced from *E.coli* and tested for functionality in a phage ELISA. The phage-scFv were tested for binding to purified target antigen and a non-relevant control protein. The fold over background (FOB) is shown for both and error bars represent the standard deviation of phage binding tested in triplicate (B) The same IgG, with either no linker or the intron containing the M13 gpIII gene between the V_H and C_H was produced. An ochre stop codon placed 3' of the M13 gpIII gene prevents full length light chain protein expression from non-spliced mRNAs. Upper arrow indicates the band corresponding to the heavy chain and the lower arrow corresponds to the light chain gene product in the SDS-PAGE. IgG molecules were tested for functionality in an ELISA using purified antigen. FOB (fold-over-background) signal is shown for both.

Fig. 5. Dual expression promoter systems. (A) Promoter type Pro^{cat}. In the catenated promoter system, the ATGs downstream of the mammalian EF1a promoter (P^{EF1a}) are removed from the downstream polyhedron (not shown) and PhoA bacterial promoter. The Kozak sequence and *E.coli* ribosomal binding site (RBS) are designed such that the first ATG *f*Met start site for either bacterial, insect (not shown) or mammalian expression is identical. In this case, the same signal sequence (SigPep) is used for all three organisms. (B) Promoter type Pro^{splice}. In this dual function promoter, the LacPO and bacterial signal peptide ($SigPep^{E.c.}$) sequence are contained within the mammalian intron. The bacterial signal peptide sequence overlaps the 3' splice site (3'ss). In *E. coli*, transcription from the bacterial promoter within the mammalian intron results in the expression of the scFv in the bacterial periplasm fused to the M13 gpIII protein in an amber-suppressing strain of *E.coli* (for example, TG1). In mammalian cells, intron splicing of the mRNA at the 5' (5'ss) and 3'ss removes the bacterial LacPO regulatory sequences located within the intron. Intron splicing generates the mammalian signal sequence. (C) Phage-scFv production from Pro^{cat} and Pro^{splice} promoters. Phage were tested for binding to purified target antigen and to a non-relevant control antigen in a phage ELISA. FOB (fold-over-background) signal is shown and error bars represent the standard deviation of

phage binding tested in triplicate (D) Expression of the wild type (wt), Pro^{splice} and Pro^{cat} promoters in HEK293 mammalian cells. IgG purification from HEK293 cells where the IgG heavy chain gene was under the expression control of either an EF1A promoter alone, Pro^{splice} or Pro^{cat}.