JOURNAL OF VIROLOGY, Apr. 2006, p. 3285–3292 Vol. 80, No. 7
0022-538X/06/$08.00+0 doi:10.1128/JVI.80.7.3285–3292.2006

Selection of Novel Vesicular Stomatitis Virus Glycoprotein Variants from a Peptide Insertion Library for Enhanced Purification of Retroviral and Lentiviral Vectors

Julie H. Yu and David V. Schaffer*

Department of Chemical Engineering and the Helen Wills Neuroscience Institute, University of California, Berkeley, Berkeley, California

Received 22 November 2005/Accepted 19 January 2006

The introduction of new features or functions that are not present in an original protein is a significant challenge in protein engineering. For example, modifications to vesicular stomatitis virus glycoprotein (VSV-G), which is commonly used to pseudotype retroviral and lentiviral vectors for gene delivery, have been hindered by a lack of structural knowledge of the protein. We have developed a transposon-based approach that randomly incorporates designed polypeptides throughout a protein to generate saturated insertion libraries and a subsequent high-throughput selection process in mammalian cells that enables the identification of optimal insertion sites for a novel designed functionality. This method was applied to VSV-G in order to construct a comprehensive library of mutants whose combined members have a His_6 tag inserted at likely every site in the original protein sequence. Selecting the library via iterative retroviral infections of mammalian cells led to the identification of several VSV-G-His_6 variants that were able to package high-titer viral vectors and could be purified by Ni-nitrilotriacetic acid affinity chromatography. Column purification of vectors reduced protein and DNA impurities more than 5,000-fold and 14,000-fold, respectively, from the viral supernatant. This substantially improved purity elicited a weaker immune response in the brain, without altering the infectivity or tropism from wild-type VSV-G-pseudotyped vectors. This work applies a powerful new tool for protein engineering to construct novel viral envelope variants that can greatly improve the safety and use of retroviral and lentiviral vectors for clinical gene therapy. Furthermore, this approach of library generation and selection can readily be extended to other challenges in protein engineering.

The envelope protein of retroviral and lentiviral vectors dictates many of their properties, including extracellular stability and cellular tropism, from the time of production to their internalization by a target cell. Desirable improvements to these properties, such as enhanced purification or tissue-specific gene delivery, require engineering the envelope protein to incorporate a new feature without compromising its ability to mediate cellular entry. The vesicular stomatitis virus glycoprotein (VSV-G) is widely used to pseudotype retroviral and lentiviral vectors for gene delivery due to its broad tropism and enhanced stability relative to native envelope proteins (12, 48). In addition, VSV-G is a model protein used to study trafficking in the secretory pathway (34) and its parent virus shows promise for use as an adjuvant expression vector in vaccines (35). Novel functions have been engineered into various envelope proteins through the introduction of new sequences at known binding domains or known tolerated insertion sites (7, 45). Modifications to the extracellular domain of VSV-G, however, have been hindered by a lack of structural knowledge of the protein and limited identification of permissible insertion sites (14, 28, 37), making it an excellent candidate for applying a novel library mutagenesis method to improve its functionality.

Library generation and selection approaches have been broadly successful for engineering or enhancing features of a target protein in the absence of detailed structural knowledge (49). In particular, directed evolution has yielded impressive results in enzyme and antibody engineering through iterative, incremental improvements in protein function (4, 41, 51). However, methods relying on point mutation or recombination of similar DNA sequences typically cannot introduce completely new functions. Fusing or inserting a peptide or domain into a target protein may introduce novel capabilities, but identifying optimal fusion locations in the absence of structural information is challenging. Therefore, novel protein libraries with polypeptides inserted at random locations may offer a high-throughput means of enhancing protein function. Techniques based on endonucleases or viral integrase, which can generate pools of insertion mutants more efficiently than classical linker scanning, have been used for DNA footprinting and bacterial protein engineering (15, 29, 31, 39). However, these methods can produce variable or biased insertions, and library generation efficiencies are often too low to apply to larger genes. Transposon-based insertional mutagenesis has recently emerged as an efficient means of studying the features of viral genomes (1, 5, 22, 25, 44). We hypothesized that a transposon-based approach for saturation insertional mutagenesis, coupled with a high-throughput viral-based library selection process, could rapidly identify optimal sites within VSV-G that could functionally incorporate a novel peptide sequence.

While VSV-G-pseudotyped vectors are commonly concentrated by ultracentrifugation for research applications (8), the vectors should be further purified for clinical use to eliminate cellular contaminants, which can generate an immune re-

* Corresponding author. Mailing address: Department of Chemical Engineering, University of California, Berkeley, 201 Gilman Hall, Berkeley, CA 94720-1462. Phone: (510) 643-5963. Fax: (510) 642-4778. E-mail: schaffer@cchem.berkeley.edu.

sponse in a patient as well as reduce transduction efficiency (2, 10, 26, 43). Immobilized metal affinity chromatography (IMAC) has been used for several viral vectors to significantly improve the purity of a viral preparation (18, 20, 47). We sought to utilize the high affinity of nickel nitrilotriacetic acid (Ni-NTA), an immobilized form of nickel that is used for protein purification (33), to purify VSV-G-pseudotyped retroviral and lentiviral vectors by incorporating a His_6 tag into the Indiana strain VSV-G protein.

Specifically, we developed a transposon-based method to construct a saturated random insertion library of VSV-G mutants whose members have a single His_6 tag inserted at most likely every site in the protein. The resulting library was selected via iterative rounds of retroviral infection of mammalian cells and led to the development of novel VSV-G-His_6 variants, which were able to pseudotype high-titer retroviral and lentiviral vectors that could be purified by Ni-NTA chromatography for use in vitro and in vivo. This novel random insertion approach can readily be applied to address other challenges in viral vector design and, more broadly, to engineer other mammalian proteins.

MATERIALS AND METHODS

Cell lines. HEK 293T and 293 cells were cultured in Iscove's modified Dulbecco's medium (IMDM) supplemented with 10% fetal bovine serum at 37°C and 5% CO_2, except when packaging vectors at 30°C (described below). The 293-*gag-pol* cell line was constructed by cotransfection of pCMV *gag-pol*, a plasmid that expresses Moloney murine leukemia virus *gag-pol* from the cytomegalovirus promoter, and a plasmid that expresses the neomycin resistance gene into 293 cells. Single colonies were expanded and tested for their ability to package infectious retrovirus upon transfection of pCLPIT GFP, a retroviral construct (19) that expresses enhanced green fluorescent protein (eGFP; BD Clontech, Palo Alto, CA), and pcDNA3 IVS VSV-G, a plasmid based on pcDNA3 that expresses the human β-globin intron and wild-type (WT) VSV-G from the cytomegalovirus promoter.

Construction of the pCLPIT VSV-G-His_6 library and clonal helper plasmids. The kanamycin resistance (Kan^r) gene was randomly inserted into a plasmid containing *vsv-g* (Indiana strain) by using a mutation generation system kit (Finnzymes, Espoo, Finland). This plasmid library was digested to excise *vsv-g*-Kan^r fragments, which were subsequently cloned into pCLPIT, allowing for transgene expression to be regulated with doxycycline (13). The His_6 insert was constructed by annealing the following oligonucleotides: 5′-AGTCGGGCCCACCACCACCATCATCATGGGGCCCAGTC-3′ and 5′-GACTGGGCCCCATGATGATGGTGGTGGTGGGCCCGACT-3′, where the region encoding the His_6 is underlined. The pCLPIT VSV-G-Kan^r library was digested with NotI before ligation to His_6 inserts digested with *Psp*OMI. The ligation product was digested with NotI before transformation to eliminate backbone religations. The total library size was estimated by colony counting of a dilution of each transformation. The plasmid library and individual clones were digested with BstXI, which cleaves once in the insert and once in the backbone, to confirm insertion number and diversity.

Individual VSV-G-His_6 sequences were constructed by using splicing by overlap extension PCR, cloned into pcDNA3 IVS, and verified by sequencing analysis (17). Oligonucleotides were designed to include the sequence for the His_6 insert and a short portion of the *vsv-g* sequence neighboring the desired site of insertion.

Viral vector production. Vectors were packaged by calcium phosphate transfection of 293T cells in 10-cm plates. For the VSV-G-His_6 library, 10 μg pCLPIT VSV-G-His_6, 6 μg pCMV *gag-pol*, and 4 μg pcDNA3 IVS VSV-G were first transfected in the presence of doxycyline to suppress the expression of VSV-G-His_6 proteins. For the clonal analysis, 4 to 8 μg individual pcDNA3 IVS VSV-G-His_6 constructs were transfected with 10 μg pCLPIT GFP and 6 μg pCMV *gag-pol* to create retroviral vectors or 3.5-μg VSV-G-His_6 constructs were transfected with 10 μg pHIV CS TRIP CG (a lentiviral construct based on pHIV CS CG [30] that contains the central polypurine tract), 5 μg pMDLg/pRRE (11), and 1.5 μg pRSV Rev for lentiviral vectors. Culture medium was changed after 12 h, and 36 h later, viral supernatant was collected and concentrated by ultracentrifugation in an SW41 rotor (Beckman Coulter, Fullerton, CA) at 50,000 × *g* for 1.5 h at 4°C before resuspension in phosphate-buffered saline (PBS) (pH 7.0). To package clones at 30°C, cells were transfected as described above and incubated at 30°C 12 h after transfection. Production of viral supernatant for use in vivo was performed as described above. For ultracentrifugation enrichment, supernatant was first concentrated through a 20%-sucrose-in-PBS cushion by ultracentrifugation at 50,000 × g for 1.5 h at 4°C. Pellets were resuspended in 10 ml PBS and ultracentrifuged again prior to resuspension in PBS. For column-purified vectors, the column eluate was diluted into 8 ml PBS, concentrated as described above, and resuspended in fresh PBS to remove any imidazole.

To determine the titers of CLPIT VSV-G-His_6 and CLPIT VSV-G stocks, serial dilutions of concentrated virus were used to infect 293T cells with 8 μg/ml polybrene. After 24 h, cells were washed and cultured in the presence of 1 μg/ml puromycin for an additional 48 h. Proliferating cells were counted by using the WST-1 assay (Roche, Indianapolis, IN), and the percentage of puromycin-resistant cells was calculated by comparison to control cells. To determine the titers of eGFP-expressing vectors, 293T cells were infected with at least three different volumes of vector supernatant or concentrate with polybrene for 24 h. Cells were assayed for eGFP expression by flow cytometry 48 h after infection. In both assays, multiplicities of infections (MOIs) were first estimated by assuming a Poisson distribution for infection. Titers were then calculated by linear regression of samples for which the MOIs were less than 1.

Immunofluorescence detection of VSV-G. pCLPIT VSV-G-His_6 or pCLPIT VSV-G plasmids were transfected into 293T cells. Sixteen hours after transfection, cells were washed, fixed, and blocked before incubation with mouse anti-VSV-G antibody P5D4 (1:1,000; Sigma, St. Louis, MO), which recognizes the C terminus, in the presence of 0.3% Triton X-100 to detect intracellular expression or in the presence of the I1 antibody (1:100; gift from Douglas Lyles) (27) without Triton X-100 to detect surface expression. Cells were washed and incubated with donkey anti-mouse AlexaFluor 488 (1:250; Molecular Probes, Eugene, OR) secondary antibody and were counterstained with TO-PRO-3 (1:2,000; Molecular Probes) before imaging by fluorescence confocal microscopy (Leica Microsystems, Wetzlar, Germany). Recognition with the conformation-specific I1 antibody confirms that proteins express a correctly folded epitope. Equivalent results were seen by using the I14 antibody (data not shown) (27).

Western blot detection of Ni-NTA binding. Concentrated vectors were lysed in radioimmunoprecipitation assay buffer and immunoprecipitated by using the P5D4 antibody. Proteins were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis and transferred to a nitrocellulose blot. The blot was blocked in Tris-buffered saline with 1 mg/ml lysozyme (Sigma) and incubated in 1 μM Ni-NTA-biotin (gift from Ravi Kane), 10 mM imidazole, and 1 μg/ml streptavidin-horseradish peroxidase (Amersham Biosciences, Piscataway, NJ). Bands were detected by ECL detection assay (Amersham Biosciences).

Library selection. Vectors containing CLPIT VSV-G-His_6 library genomes pseudotyped with WT VSV-G were used to infect 293-*gag-pol* cells at an MOI of <0.1. Cells were selected by using 1 μg/ml puromycin and propagated in the presence of 100 ng/ml doxycycline to prevent continuous production of virus. To rescue virus, infected 293-*gag-pol* cells were grown to confluence without doxycycline, and 5 mM sodium butyrate was added 2 days before viral harvest (21). Harvested virus (round 1 of selection) was then used to infect at least 10^6 naïve 293-*gag-pol* cells at MOIs of <0.1, and cells were propagated as described above. This process was repeated for each successive round of selection. To select for Ni-NTA binding, vectors were purified with Ni-NTA (described below) before infection of naïve cells. To identify selected sequences, cellular genomic DNA or viral genomic RNA was isolated by using the QIAGEN genomic tip 500/G or QIAamp viral RNA kit (QIAGEN, Palo Alto, CA), respectively. VSV-G-His_6 sequences were amplified by PCR and inserted into a plasmid before sequencing.

Ni-NTA purification of viral vectors. A total of 500 μl of 50% Ni-NTA agarose (QIAGEN) was rinsed with PBS (pH 7.0) and incubated with 300 to 600 μl of concentrated viral stocks with gentle agitation at 4°C for 1 h. The mixture of virus and beads was loaded onto a plastic column (Kontes, Vineland, NJ) before washing with 3 ml of 50 mM imidazole in PBS and elution with 1.5 to 2 ml of 250 mM imidazole in PBS.

Equivalent volumes of each column fraction from a representative purification procedure were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis. Proteins were detected by using the SilverQuest kit (Invitrogen, Carlsbad, CA). The IMDM and viral supernatant samples were diluted 10-fold to prevent oversaturation of the silver stain signal. Protein concentrations in stocks representing a 20-fold concentration of viral supernatant by conventional or column purification were quantified by using the bicinchoninic acid assay (Pierce Biotechnology, Rockford, IL). DNA concentrations in those stocks were quantified by spectrofluorometry after incubation with SYBR green double-stranded DNA dye (Molecular Probes).

FIG. 1. VSV-G-His_6 library design. (A) Structure of the pCLPIT VSV-G-His_6 vector, which expresses the VSV-G-His_6 library from a tetracycline regulatable promoter and puromycin resistance from the viral long-terminal repeat (LTR). IRES, internal ribosome entry site; tTA, tetrecycline-controlled transactivator; TRE, tetracycline response element. (B) Peptide and DNA sequences for the His_6 insert after insertion. X_1 and X_2 will depend on the five-host nucleotides (N) duplicated during insertion. The digested insert sequence is underlined and in bold. (C) The pCLPIT VSV-G-His_6 plasmid library and clones were cut once in the His_6 insertion and once in pCLPIT. Successful insertions into *vsv-g* yield fragments of 1.6 to 3.2 kb in size. Lanes: 1, pCLPIT VSV-G; 2, pCLPIT VSV-G-His_6 library; 3 through 12, randomly selected pCLPIT VSV-G-His_6 clones.

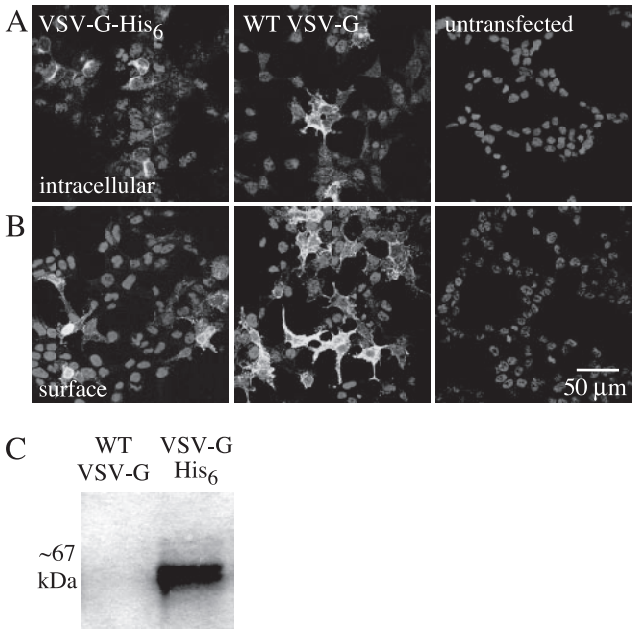

FIG. 2. Expression of library proteins. (A and B) Immunostaining of cells transfected with pCLPIT VSV-G-His_6 or pCLPIT VSV-G to detect intracellular (A) and surface (B) expression of VSV-G (white, 63× objective). Cells are counterstained with TO-PRO-3 (gray). (C) Western blot detection of VSV-G-His_6 library proteins binding to Ni-NTA.

Animal injections and expression analysis. Animal protocols were approved by the UCB Animal Care and Use Committee in accordance with NIH guidelines. Anesthetized adult female Fischer-344 rats were injected with either lentiviral vectors pseudotyped with G-19LH_6 ($n = 3$ animals) or G-24LH_6 ($n = 3$) that had been purified on a Ni-NTA column or vectors pseudotyped with G-24LH_6 ($n = 3$) or WT VSV-G ($n = 2$) that were purified by ultracentrifugation only. Animals received 3 μl of high-titer vector preparations (8×10^8 to 1.2×10^9 IU/ml) into the striatum by stereotaxic injection (coordinates from bregma: AP, +0.2; ML, ±3.5; DV, −4.5 from dura with nose bar at +3 mm). After 2 weeks, brains were harvested and coronal sections (40 μm) were taken as we have previously described (24). Primary antibodies included rabbit anti-GFP (1:2,000 dilution; Molecular Probes), guinea pig anti-glial fibrillary acidic protein (GFAP) (1:1,000; Advanced Immunochemical, Long Beach, CA), mouse anti-NeuN, mouse anti-OX8, and mouse anti-ED1 (1:100; Chemicon, Temecula, CA). Corresponding AlexaFluor 488-, 546-, or 633-conjugated secondaries (1:250; Molecular Probes) were used, and some sections were counterstained with TO-PRO-3 (1:2,000) before imaging by confocal microscopy. The area of eGFP expression in 22 to 26 evenly spaced sections from each animal was measured, and total volume of eGFP expression in each sample was estimated by using modified stereological methods.

RESULTS

Construction of the VSV-G-His_6 library. We first constructed a library of *vsv-g*-Kan^r insertion mutants by using MuA transposase and a modified transposon insert carrying a kanamycin resistance gene (Kan^r) (16). The resulting *vsv-g*-Kan^r gene library was cloned into the retroviral vector construct pCLPIT, which also expresses a puromycin resistance gene (19), and Kan^r was then replaced with a His_6 tag sequence to create pCLPIT VSV-G-His_6 (Fig. 1A). The sequence of the 13-amino-acid (aa) insert is dependent on the five neighboring host nucleotides that are duplicated during the transposition reaction (16) (Fig. 1B). Restriction digest analyses of the pooled plasmid library and randomly selected clones confirmed that insertions occurred at a diverse number of sites and that each clone had a single insertion (Fig. 1C). After accounting for insertions into noncoding regions, the library was estimated to contain over 2.4×10^4 independent insertions into the 1.6-kb VSV-G cDNA. We constructed a similar control vector, pCLPIT VSV-G, which expresses WT VSV-G.

Immunostaining of cells transfected with pCLPIT VSV-G-His_6 or pCLPIT VSV-G plasmids revealed VSV-G expression intracellularly and on the cell surface in both populations, indicating that members of the *vsv-g-his*$_6$ library can express VSV-G and that at least some of these variants are trafficked to the cell surface (Fig. 2A and B). Western blot analysis of VSV-G-His_6-pseudotyped retroviral vectors by using a Ni-NTA probe demonstrated that His_6-containing VSV-G proteins could be incorporated into retroviral particles (Fig. 2C).

Selection of the VSV-G-His_6 library by using retroviruses. To select insertion mutants that retained the ability to mediate cellular infection, the viral library was serially passaged on 293 cells that were stably expressing retroviral *gag-pol* (Fig. 3A). Viral titers were determined at each round of infection and rescue (Fig. 3B). Importantly, sequencing a sample of the library after three selection rounds showed that at least three sites in the signal peptide and three in the extracellular domain allowed the 13-aa insertion. Insertions in the signal peptide likely yield proteins identical to the WT in their mature form and were thus not pursued. Three other sites before amino acid positions 19, 24, and 26 (with the start codon as position 1), however, have not been previously identified. Next, to isolate insertion mutants that presented functionally active peptides, we repeated the selection protocol but also purified the viral library on a Ni-NTA column before infecting cells. This procedure was repeated until the majority of the viral library

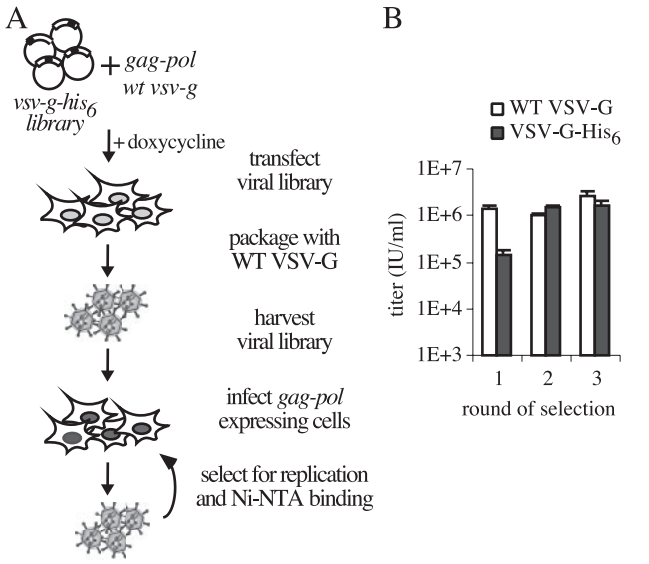

FIG. 3. VSV-G-His_6 library selection. (A) Schematic of library selection by using retroviral infection of cells. (B) Viral titers for each round of selection for replication. Error bars represent the standard error of the linear regression used to determine titers.

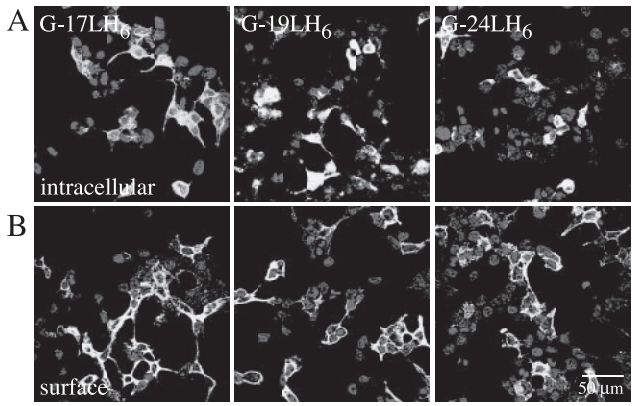

FIG. 4. Immunofluorescence detection of VSV-G-His_6 clones. Detection of (A) intracellular and (B) surface expression of VSV-G (white) from individual VSV-G-His_6 clones (63× objective). Cells are counterstained with TO-PRO-3 (gray).

loaded was present in the final eluate. An analysis of these sequences revealed that a single clone containing a His_6 insertion in frame in site 19 was dominant within the population.

Purification of VSV-G-His_6-pseudotyped vectors by Ni-NTA chromatography. To assess the ability of VSV-G-His_6 variants to pseudotype retroviral and lentiviral vectors, we inserted a His_6 tag into all three novel sites that were identified from the selection as well as in site 17, the N terminus of the mature protein, and in two previously identified sites, the temperature-sensitive site 25 (14) and site 191 (37). Furthermore, to explore the effects of different amino acid linkers flanking the His_6 tag, we constructed multiple variants at two of the sites (Table 1). All variants were inserted into a helper plasmid for vector production, and plasmid transfections demonstrated that the new mutants expressed VSV-G proteins that could be trafficked to the surface of cells (Fig. 4).

All of the novel VSV-G-His_6 helper constructs successfully packaged retroviral and lentiviral vectors expressing eGFP, and many had titers comparable to those of vectors packaged with WT VSV-G (Fig. 5A). Clone G-25LH_6, with an insert at the temperature-sensitive site 25, did not yield infectious retroviruses when packaged at 37°C but was able to produce retroviral vectors at 30°C as previously described (14). Typical titers of viral supernatant for these vectors using WT VSV-G are 3×10^6 to 5×10^6 IU/ml for retrovirus and 1×10^7 to 3×10^7 IU/ml for lentivirus. Importantly, most variants bound to a Ni-NTA column at levels substantially higher than that of WT VSV-G (Fig. 5B). Two of the variants, G-19LH_6 and G-24LH_6, were then used to optimize parameters such as binding volume, incubation time, buffer pH, ratio of virus to Ni-NTA resin, and wash and elution conditions. By using an optimized procedure, the total virus recovery exceeded 50% (Fig. 5C).

Column purification dramatically reduced protein contaminants relative to purification by ultracentrifugation through a sucrose cushion, a conventional method for enriching vectors (3) (Fig. 5D). Protein and DNA concentrations in the column-purified preparations were below the detection limit of their respective assays, so estimates are conservatively based on the minimum values of 25 μg/ml and 10 ng/ml, respectively. For vector stocks that were 20-fold more concentrated than was the original viral supernatant, the ultracentrifugation-enriched virus had only a 5-fold and a 2.7-fold reduction in protein and DNA concentrations relative to the crude viral supernatant. In stark contrast, the column-purified virus had at least 250-fold

TABLE 1. VSV-G-His_6 clone sequences[a,b]

Name	Amino acid sequence
G-17_H_6	MKCLLYLAFLFIGVNCHHHHHHKFTIVFPHNQKGN. . .
G-17L3H_6	MKCLLYLAFLFIGVNCHHHHHHGGSKFTIVFPHNQKGN. . .
G-17L6H_6	MKCLLYLAFLFIGVNCHHHHHHGGSGGSKFTIVFPHNQKGN. . .
G-19LH_6	MKCLLYLAFLFIGVNCKFSGGHHHHHHGGSTIVFPHNQKGN. . .
G-24_H_6	MKCLLYLAFLFIGVNCKFTIVFPHHHHHHHNQKGN. . .
G-24LH_6	MKCLLYLAFLFIGVNCKFTIVFPSGGHHHHHHGGSHNQKGN. . .
G-25LH_6	MKCLLYLAFLFIGVNCKFTIVFPHSGGHHHHHHGGSNQKGN. . .
G-26LH_6	MKCLLYLAFLFIGVNCKFTIVFPHNSGGHHHHHHGGSQKGN. . .
G-191LH_6	. . .DYKVKSGGHHHHHHGGSGLCDSN. . .
WT VSV-G	^{1}MKCLLYLAFL^{11}FIGVNCKFTIV^{21}FPHNQKGN. . .^{186}DYKVK^{191}GLCDSN. . .

[a] Each clone is designated by the site of insertion and presence of a linker peptide (L). Residue positions are provided in the sequence for WT VSV-G.
[b] Newly introduced amino acids are underlined.

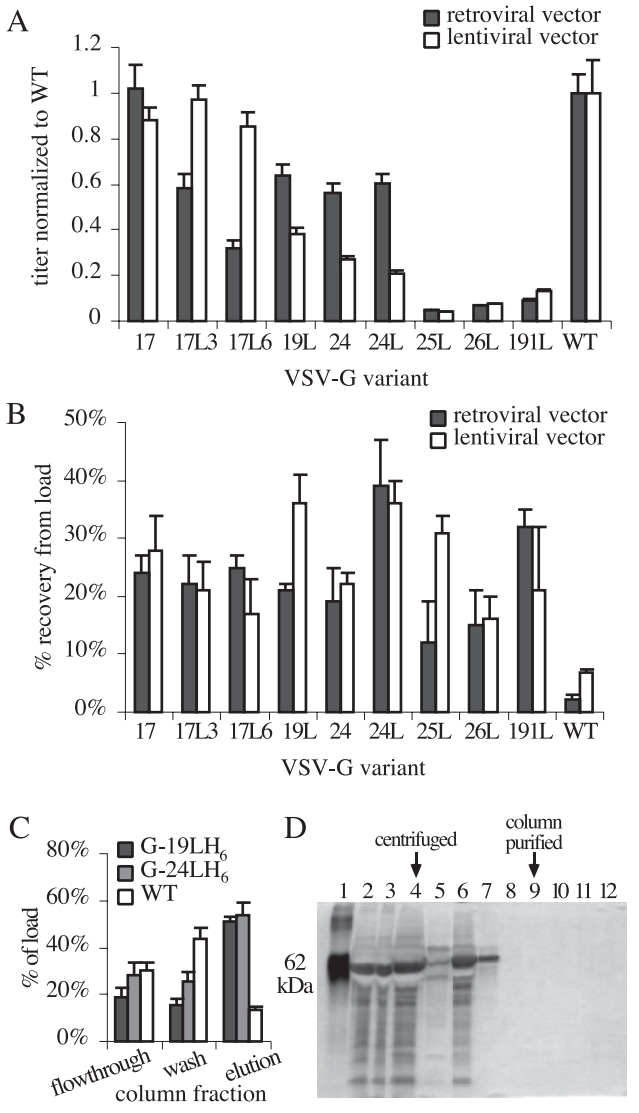

FIG. 5. Column purification of VSV-G-His_6-pseudotyped retroviral and lentiviral vectors. (A) Representative titers of retroviral and lentiviral vectors expressing eGFP pseudotyped with VSV-G-His_6 variants. Results for the G-$25LH_6$-pseudotyped retroviral vector reflect packaging at 30°C. All other vectors were produced at 37°C. Error bars represent the standard error of the linear regression used to determine titers. (B) Recovery of vectors pseudotyped with VSV-G-His_6 variants after Ni-NTA purification. Error bars represent the standard error of the linear regression used to determine titers. (C) Optimized purification profile of G-$19LH_6$- and G-$24LH_6$- pseudotyped lentiviral vectors. Error bars represent the standard error of the linear regression used to determine titers. (D) Silver staining of column fractions. Lanes: 1, marker; 2, IMDM with 10% fetal bovine serum (1:10 dilution); 3, vector supernatant (1:10 dilution); 4, ultracentrifuged virus; 5, column flowthrough; 6 to 8, successive washes; 9 to 12, successive eluates.

and 700-fold reduced protein and DNA concentrations, resulting in an overall reduction in protein and DNA contamination by over 5,000-fold and 14,000-fold, respectively.

Expression of VSV-G-His_6-pseudotyped vectors in vivo. To evaluate the performance of column-purified VSV-G-His_6-pseudotyped vectors in vivo, column and conventionally purified lentiviral vectors expressing eGFP were injected into the striatum of adult rats. Two weeks after injection, eGFP expression was observed in every animal, and the cellular tropism of the VSV-G-His_6-pseudotyped vectors was the same as WT VSV-G-pseudotyped lentiviral vectors, with preferential infection of NeuN-positive neurons and modest colocalization with $GFAP^+$ astrocytes (32) (Fig. 6A and B). There was no statistically significant difference in the infection spread in the anterior-posterior axis, as determined by the number of $eGFP^+$ sections, or the overall volume of spread between any of the vector preparations (Fig. 6C and D).

Reduction of immune response by using column-purified vectors. A safety concern for the use of gene delivery vectors is the potential generation of a patient immune response (2, 43). $CD8^+$ T-cell and macrophage activation and infiltration within the region of transduced tissue were considerably different between injections using column-purified versus conventionally purified virus. Every injection site, including a PBS control, revealed a low-level infiltration of immune cells directly around the needle track, a common result of needle insertion (data not shown). Within high-eGFP-expressing regions that were distal from the injection sites, however, the animals that were injected with the column-purified vectors had substantially fewer immune cells than did those that were injected with conventionally purified preparations, independently of whether the latter vectors were pseudotyped with WT VSV-G or a VSV-G-His_6 variant (Fig. 6E).

DISCUSSION

We have constructed new VSV-G variants by applying a novel, high-throughput protein engineering system that employs saturation insertion mutagenesis before rapid selection in mammalian cells. Modifications to viral proteins require not only a consideration of the complex interactions within an individual folded protein but also the intersubunit interactions that are necessary for intricate assemblies into trimeric envelopes or multimeric capsids. Consequently, polypeptides may behave differently when inserted at the same site due to differences in sequence and length. As a result, although information gained through genetic footprinting could be used to construct novel proteins, the process of initially determining sites and then testing for new functionality in each one is laborious and does not guarantee the identification of optimal or even functional sites for insertional modification, particularly when the molecular properties of the functional insert diverge from the scanning insert, as has been seen in studies of viral genomes (22). Combining the highly efficient, random insertion capabilities of transposases with the high-throughput-screening capacities of retroviruses facilitates the direct functional selection of novel, desired properties within mammalian cells.

We used this method to construct a *vsv-g-his*$_6$ gene library that had an average of 15 independent insertion events in each internucleotide position as a result of the primarily sequence-independent transposition reaction (9). Selection for the abilities to pseudotype retroviruses and bind to Ni-NTA was then performed directly in mammalian cell culture. Much as phages are natural "partners" for library selection in bacteria, retro-

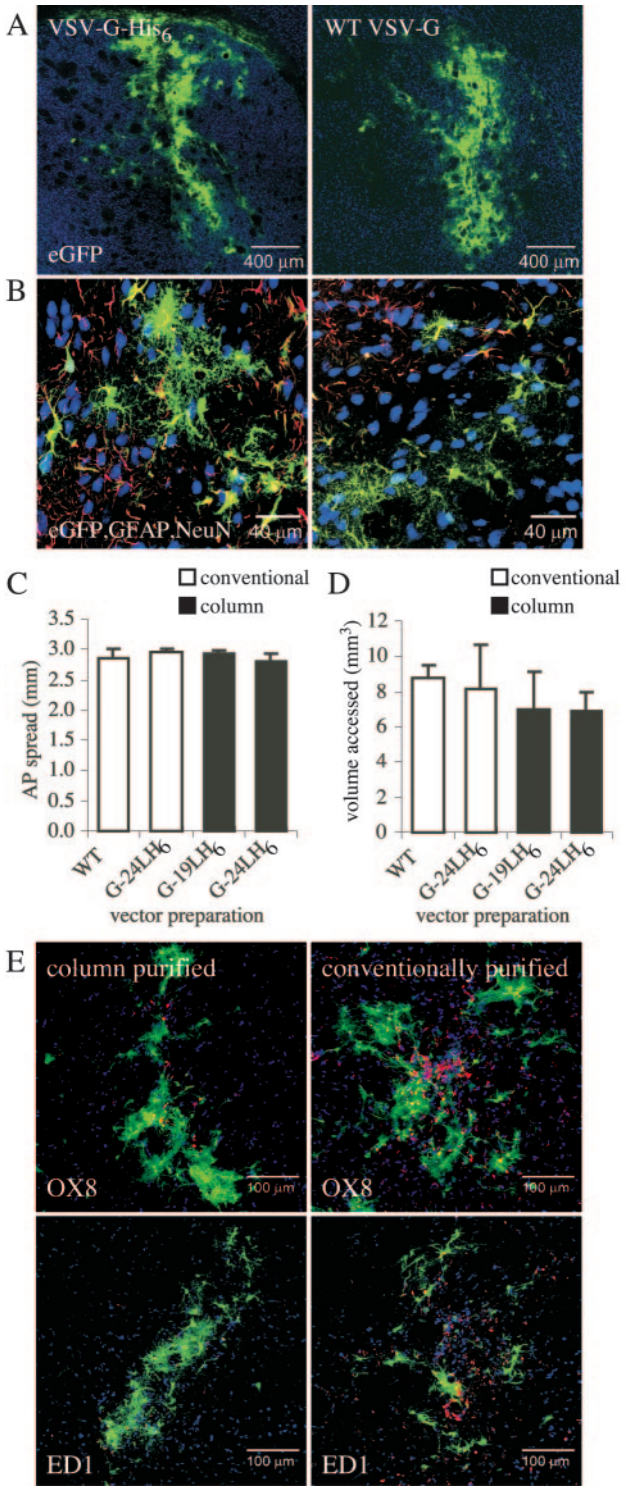

FIG. 6. Behavior of VSV-G-His_6-pseudotyped lentiviral vectors in vivo. (A) Representative images of injections with VSV-G-His_6-pseudotyped vectors display equivalent eGFP expression (green) to vectors pseudotyped with WT VSV-G (10× objective). Cells were counterstained with TO-PRO-3 (blue). (B) Representative images show that VSV-G-His_6 and WT VSV-G-pseudotyped vector tropism are equivalent in the brain (63× objective). Cells were stained with antibodies against NeuN (blue) and GFAP (red) to identify mature neurons and astrocytes, respectively. (C) Vector spread through the brain for each preparation based on the number of $eGFP^+$ sections. The P value was >0.3 by analysis of variance (ANOVA). Error bars represent the standard error of the mean of each preparation. AP spread, anterior-posterior axis spread. (D) Overall volume was assessed by each vector preparation based on eGFP expression in 22 to 26 sections per animal. The P value was >0.4 by ANOVA. Error bars represent the standard error of the mean of each preparation. AP spread, anterior-posterior axis spread. (E) Reduction in immune response by using column-purified vectors. Immunostaining of $CD8^+$ T cells (red, OX8) and macrophages (red, ED1) from animals that were injected with column-purified or conventionally purified viral stocks (10× objective). Cells are counterstained with TO-PRO-3 (blue). Images are representative areas of high eGFP expression that were at least 200 μm away from the site of injection and the corpus callosum to avoid bias introduced by enhanced transport in these areas.

viruses are natural platforms for mammalian library selection due to the iterative transfer of information between the viral and host cell genomes (6). Furthermore, by selecting in the context of mammalian cells, we were able to analyze variants in their natural environment, avoiding the use of bacteria where proteins (particularly transmembrane ones) may not fold or function correctly. A low multiplicity of infection ensured that the vast majority of the new cell library had only one copy of the *vsv-g-his*$_6$ library integrated into their genomes. Therefore the genotype encapsulated by a given virion coded for the phenotype of the VSV-G-His_6 variant that was expressed on its surface.

We first enriched the library with infectious variants, as it is likely that the majority of insertions are deleterious to protein function. The titers of VSV-G-His_6 library vectors were initially much lower than those of the vectors pseudotyped with WT VSV-G, but they became comparable to those of the control after a round of low-MOI infection-and-rescue selected a pool of infectious variants with insertions of at least 13 aa. Since there are few sites in VSV-G that are known to tolerate an insertion of even 2 to 3 aa, we recovered the sequences of integrated viral genomes and identified three novel sites that allow productive insertions. While the insert size ensured that all VSV-G codons remain in frame, transposon insertions occur at the nucleotide level and sequences inserted out of the desired frame express peptides other than the His_6. Therefore, the initial selection identified sites that permitted general polypeptide inserts whose sequences differed depending on the frame of insertion. After selecting the library for Ni-NTA binding, however, we isolated a dominant clone with a His_6 sequence in frame at position 19 that exhibited the property we sought to engineer: the ability to produce infectious retroviral and lentiviral vectors that bind to Ni-NTA.

To assess whether the identified site 19 was the sole productive site for a His_6 tag or whether it had singly emerged as a result of viral population dynamics and other functional mutants were possible, we rationally inserted a His_6 sequence flanked by linker peptides into all three novel sites. Since aa 19 is near the beginning of the protein, we also inserted a His_6 tag into position 17, the N terminus of the mature protein. Analysis of these variants revealed that insertions at every site identified by our screen successfully pseudotyped infectious virions that could bind to Ni-NTA. Additionally, we were able to use these results to predict a successful insertion at the N terminus of the mature protein, which has not been previously

reported. Therefore, while selection in retroviruses successfully identified an optimal site, it may be useful to consider sequences that are present at earlier rounds of evolution to identify multiple permissive sites. Interestingly, the inclusion of flanking linkers at aa 24 yielded titers comparable to those of vectors with a smaller, less disruptive insert but resulted in more efficient recovery by Ni-NTA purification. The larger insert is closer in size to the sequence used to identify the site as permissive and illustrates the benefit of directly selecting the insertion library for the ability to accommodate the final desired peptide rather than an initial generic linker, such as a restriction site. The sites identified by the replication selection were able to produce infectious virions with several distinct polypeptide inserts of the same size, indicating that they may tolerate the insertion of variable sequences. Since Ni-NTA binding requires only one His_6 tag, the other novel sites identified here could be useful for exploring the addition of new surface-associated features, such as targeting ligands, that are 12 to 13 aa long.

Two sites in the extracellular domain of VSV-G have been previously found to tolerate a small peptide insertion. VSV-G proteins with an 18-aa insertion before position 25 developed a temperature-sensitive mutation in intracellular trafficking, and titers of retroviral vectors packaged with this mutant at the permissive temperature were 1 to 2 orders of magnitude lower than those of the control (14). Our results with a His_6 tag rationally inserted in this site, G-25LH$_6$, agreed that retroviral vectors had to be packaged at 30°C in order to generate functional virus. Interestingly, the G-25LH$_6$ variant was able to successfully pseudotype lentiviral vectors when packaged at 37°C, though titers were low. It is important to note that two of the novel sites that are identified in this work flank this rationally determined site. However, while insertions at aa 25 had a temperature-sensitive phenotype, the new insertions at position 24 or 26 created variants that could be properly trafficked at 37°C. This result underscores the power of using library screening to identify protein enhancements that are difficult to predict by rational design. Finally, the site at position 191 has been previously shown to tolerate an insertion of up to 16 aa and still package infectious vesicular stomatitis virus (37), though a titer was not reported. This study confirms that variants with an insertion at site 191 can be used to pseudotype retroviral and lentiviral vectors, albeit at significantly reduced titers.

Analysis of injected animals demonstrated that viral infection and tropism in the striatum were similar to WT VSV-G-pseudotyped vectors. WT tropism was also confirmed in the eye when vectors were delivered by subretinal injection (data not shown). We therefore believe that the VSV-G-His_6 variants are interchangeable with WT VSV-G in numerous applications.

The ability to produce highly purified gene delivery vectors is imperative to avoiding an immune response in a patient. While affinity purification sequences have been successfully engineered into surface proteins of other viruses (18, 20, 47, 50), the introduction of a purification tag into VSV-G has not yet been reported. Anion exchange and size exclusion chromatography have been employed to increase the purity of VSV-G-pseudotyped vectors as well as to provide a scalable method to improve the economics of the therapy (23, 36, 38, 40, 42, 46). However, these methods rely on properties not unique to the virus and can result in the retention of DNA and large protein contaminants that must be removed through additional processing. By contrast, Ni-NTA purification of VSV-G-His_6-pseudotyped vectors by using a simple gravity column dramatically reduced the levels of protein and DNA contaminants in a single step, as seen with IMAC of other viral vectors (20, 47). We observed a direct benefit of this purification for in vivo use, as animals injected with column-purified vectors consistently had lower degrees of immune cell infiltration than did those injected with either His_6-tagged or WT VSV-G vectors purified by ultracentrifugation alone.

Using Ni-NTA purification, we were able to recover a greater amount of virus than has been previously reported for anion exchange or size exclusion chromatography, whose harsh elution conditions or need for further concentration can compromise yields (36, 42). Because IMAC is limited by only the binding capacity of the column resin, it is possible that titers high enough for clinical applications could be eluted directly from the column and dialyzed against a storage buffer. Dialysis of column-purified viral stocks resulted in little or no vector loss (data not shown), consistent with other findings (47). Therefore, affinity purification by Ni-NTA chromatography can provide a simple, economical means for producing highly purified, safe vector preparations.

This study introduces novel sites in VSV-G that were used to generate new His_6-tagged VSV-G variants, which could be highly purified by Ni-NTA chromatography and elicited a reduced immune response in vivo. The use of a transposon-based method to create a random peptide insertion library did not necessitate the structural knowledge required for successful rational design; however, it could yield subsequent insights into the structure-function properties of the target protein. If a desired functionality can be incorporated into a defined polypeptide sequence, these libraries can be selected for the functional insertion of that specific feature. Furthermore, by using iterative retroviral infections, we have demonstrated a convenient approach to select mammalian protein libraries. The combination of these two techniques presents a powerful method of engineering not only viral vectors for gene delivery but also virtually any protein for investigation and application in mammalian cells.

ACKNOWLEDGMENTS

We thank Joshua Leonard for constructing the lentiviral expression vector and Kenneth Greenberg for evaluating the performance of the variants in the eye. The I1 and I14 antibodies were kindly provided by Douglas Lyles (Wake Forest University). The Ni-NTA-biotin conjugate was a gift from Ravi Kane (Rensselaer Polytechnic Institute).

This work was supported by the Whitaker Foundation Graduate Fellowship to J.H.Y. and an NSF CAREER Award and NIH EB003007 to D.V.S.

REFERENCES

1. **Auerbach, M. R., C. Shu, A. Kaplan, and I. R. Singh.** 2003. Functional characterization of a portion of the Moloney murine leukemia virus *gag* gene by genetic footprinting. Proc. Natl. Acad. Sci. USA **100:**11678–11683.
2. **Baekelandt, V., A. Claeys, K. Eggermont, E. Lauwers, B. De Strooper, B. Nuttin, and Z. Debyser.** 2002. Characterization of lentiviral vector-mediated gene transfer in adult mouse brain. Hum. Gene Ther. **13:**841–853.
3. **Baekelandt, V., K. Eggermont, M. Michiels, B. Nuttin, and Z. Debyser.** 2003. Optimized lentiviral vector production and purification procedure prevents immune response after transduction of mouse brain. Gene Ther. **10:**1933–1940.

4. **Boder, E. T., and K. D. Wittrup.** 1997. Yeast surface display for screening combinatorial polypeptide libraries. Nat. Biotechnol. **15:**553–557.
5. **Brune, W., C. Menard, U. Hobom, S. Odenbreit, M. Messerle, and U. H. Koszinowski.** 1999. Rapid identification of essential and nonessential herpesvirus genes by direct transposon mutagenesis. Nat. Biotechnol. **17:**360–364.
6. **Bupp, K., and M. J. Roth.** 2002. Altering retroviral tropism using a random-display envelope library. Mol. Ther. **5:**329–335.
7. **Bupp, K., A. Sarangi, and M. J. Roth.** 2005. Probing sequence variation in the receptor-targeting domain of feline leukemia virus envelope proteins with peptide display libraries. J. Virol. **79:**1463–1469.
8. **Burns, J. C., T. Friedmann, W. Driever, M. Burrascano, and J. K. Yee.** 1993. Vesicular stomatitis virus G glycoprotein pseudotyped retroviral vectors: concentration to very high titer and efficient gene transfer into mammalian and nonmammalian cells. Proc. Natl. Acad. Sci. USA **90:**8033–8037.
9. **Butterfield, Y. S., M. A. Marra, J. K. Asano, S. Y. Chan, R. Guin, M. I. Krzywinski, S. S. Lee, K. W. MacDonald, C. A. Mathewson, T. E. Olson, P. K. Pandoh, A. L. Prabhu, A. Schnerch, U. Skalska, D. E. Smailus, J. M. Stott, M. I. Tsai, G. S. Yang, S. D. Zuyderduyn, J. E. Schein, and S. J. Jones.** 2002. An efficient strategy for large-scale high-throughput transposon-mediated sequencing of cDNA clones. Nucleic Acids Res. **30:**2460–2468.
10. **Chen, J., L. Reeves, N. Sanburn, J. Croop, D. A. Williams, and K. Cornetta.** 2001. Packaging cell line DNA contamination of vector supernatants: implication for laboratory and clinical research. Virology **282:**186–197.
11. **Dull, T., R. Zufferey, M. Kelly, R. J. Mandel, M. Nguyen, D. Trono, and L. Naldini.** 1998. A third-generation lentivirus vector with a conditional packaging system. J. Virol. **72:**8463–8471.
12. **Emi, N., T. Friedmann, and J. K. Yee.** 1991. Pseudotype formation of murine leukemia virus with the G protein of vesicular stomatitis virus. J. Virol. **65:**1202–1207.
13. **Gossen, M., and H. Bujard.** 1992. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. Proc. Natl. Acad. Sci. USA **89:**5547–5551.
14. **Guibinga, G. H., F. L. Hall, E. M. Gordon, E. Ruoslahti, and T. Friedmann.** 2004. Ligand-modified vesicular stomatitis virus glycoprotein displays a temperature-sensitive intracellular trafficking and virus assembly phenotype. Mol. Ther. **9:**76–84.
15. **Guntas, G., and M. Ostermeier.** 2004. Creation of an allosteric enzyme by domain insertion. J. Mol. Biol. **336:**263–273.
16. **Haapa, S., S. Taira, E. Heikkinen, and H. Savilahti.** 1999. An efficient and accurate integration of mini-Mu transposons in vitro: a general methodology for functional genetic analysis and molecular biology applications. Nucleic Acids Res. **27:**2777–2784.
17. **Ho, S. N., H. D. Hunt, R. M. Horton, J. K. Pullen, and L. R. Pease.** 1989. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. Gene **77:**51–59.
18. **Hu, Y. C., C. T. Tsai, Y. C. Chung, J. T. Lu, and J. T. A. Hsu.** 2003. Generation of chimeric baculovirus with histidine-tags displayed on the envelope and its purification using immobilized metal affinity chromatography. Enzyme Microb. Technol. **33:**445–452.
19. **Ignowski, J. M., and D. V. Schaffer.** 2004. Kinetic analysis and modeling of firefly luciferase as a quantitative reporter gene in live mammalian cells. Biotechnol. Bioeng. **86:**827–834.
20. **Jiang, C., J. B. Wechuck, W. F. Goins, D. M. Krisky, D. Wolfe, M. M. Ataai, and J. C. Glorioso.** 2004. Immobilized cobalt affinity chromatography provides a novel, efficient method for herpes simplex virus type 1 gene vector purification. J. Virol. **78:**8994–9006.
21. **Kafri, T., H. van Praag, L. Ouyang, F. H. Gage, and I. M. Verma.** 1999. A packaging cell line for lentivirus vectors. J. Virol. **73:**576–584.
22. **Kretschmer, P. J., F. Jin, C. Chartier, and T. W. Hermiston.** 2005. Development of a transposon-based approach for identifying novel transgene insertion sites within the replicating adenovirus. Mol. Ther. **12:**118–127.
23. **Kuiper, M., R. M. Sanches, J. A. Walford, and N. K. Slater.** 2002. Purification of a functional gene therapy vector derived from Moloney murine leukaemia virus using membrane filtration and ceramic hydroxyapatite chromatography. Biotechnol. Bioeng. **80:**445–453.
24. **Lai, K., B. K. Kaspar, F. H. Gage, and D. V. Schaffer.** 2003. Sonic hedgehog regulates adult neural progenitor proliferation in vitro and in vivo. Nat. Neurosci. **6:**21–27.
25. **Laurent, L. C., M. N. Olsen, R. A. Crowley, H. Savilahti, and P. O. Brown.** 2000. Functional characterization of the human immunodeficiency virus type 1 genome by genetic footprinting. J. Virol. **74:**2760–2769.
26. **Le Doux, J. M., J. R. Morgan, R. G. Snow, and M. L. Yarmush.** 1996. Proteoglycans secreted by packaging cell lines inhibit retrovirus infection. J. Virol. **70:**6468–6473.
27. **Lefrancios, L., and D. S. Lyles.** 1982. The interaction of antibody with the major surface glycoprotein of vesicular stomatitis virus. I. Analysis of neutralizing epitopes with monoclonal antibodies. Virology **121:**157–167.
28. **Li, Y., C. Drone, E. Sat, and H. P. Ghosh.** 1993. Mutational analysis of the vesicular stomatitis virus glycoprotein G for membrane fusion domains. J. Virol. **67:**4070–4077.
29. **Luckow, B., R. Renkawitz, and G. Schutz.** 1987. A new method for constructing linker scanning mutants. Nucleic Acids Res. **15:**417–429.
30. **Miyoshi, H., U. Blomer, M. Takahashi, F. H. Gage, and I. M. Verma.** 1998. Development of a self-inactivating lentivirus vector. J. Virol. **72:**8150–8157.
31. **Murakami, H., T. Hohsaka, and M. Sisido.** 2000. Random insertion and deletion mutagenesis for construction of protein library containing nonnatural amino acids. Nucleic Acids Symp. Ser. **44:**69–70.
32. **Naldini, L., U. Blomer, F. H. Gage, D. Trono, and I. M. Verma.** 1996. Efficient transfer, integration, and sustained long-term expression of the transgene in adult rat brains injected with a lentiviral vector. Proc. Natl. Acad. Sci. USA **93:**11382–11388.
33. **Porath, J., J. Carlsson, I. Olsson, and G. Belfrage.** 1975. Metal chelate affinity chromatography, a new approach to protein fractionation. Nature **258:**598–599.
34. **Presley, J. F., N. B. Cole, T. A. Schroer, K. Hirschberg, K. J. Zaal, and J. Lippincott-Schwartz.** 1997. ER-to-Golgi transport visualized in living cells. Nature **389:**81–85.
35. **Rose, N. F., P. A. Marx, A. Luckay, D. F. Nixon, W. J. Moretto, S. M. Donahoe, D. Montefiori, A. Roberts, L. Buonocore, and J. K. Rose.** 2001. An effective AIDS vaccine based on live attenuated vesicular stomatitis virus recombinants. Cell **106:**539–549.
36. **Scherr, M., K. Battmer, M. Eder, S. Schule, H. Hohenberg, A. Ganser, M. Grez, and U. Blomer.** 2002. Efficient gene transfer into the CNS by lentiviral vectors purified by anion exchange chromatography. Gene Ther. **9:**1708–1714.
37. **Schlehuber, L. D., and J. K. Rose.** 2004. Prediction and identification of a permissive epitope insertion site in the vesicular stomatitis virus glycoprotein. J. Virol. **78:**5079–5087.
38. **Segura, M. M., A. Kamen, P. Trudel, and A. Garnier.** 2005. A novel purification strategy for retrovirus gene therapy vectors using heparin affinity chromatography. Biotechnol. Bioeng. **90:**391–404.
39. **Singh, I. R., R. A. Crowley, and P. O. Brown.** 1997. High-resolution functional mapping of a cloned gene by genetic footprinting. Proc. Natl. Acad. Sci. USA **94:**1304–1309.
40. **Slepushkin, V., N. Chang, R. Cohen, Y. Gan, B. Jiang, E. Deausen, D. Berlinger, G. Binder, K. Andre, L. Humeau, and B. Dropulic.** 2003. Large-scale purification of a lentiviral vector by size exclusion chromatography or Mustang Q ion exchange capsule. Bioprocess. J. **2:**89–95.
41. **Stemmer, W. P.** 1994. DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution. Proc. Natl. Acad. Sci. USA **91:**10747–10751.
42. **Transfiguracion, J., D. E. Jaalouk, K. Ghani, J. Galipeau, and A. Kamen.** 2003. Size-exclusion chromatography purification of high-titer vesicular stomatitis virus G glycoprotein-pseudotyped retrovectors for cell and gene therapy applications. Hum. Gene Ther. **14:**1139–1153.
43. **Tuschong, L., S. L. Soenen, R. M. Blaese, F. Candotti, and L. M. Muul.** 2002. Immune response to fetal calf serum by two adenosine deaminase-deficient patients after T cell gene therapy. Hum. Gene Ther. **13:**1605–1610.
44. **Vilen, H., J. M. Aalto, A. Kassinen, L. Paulin, and H. Savilahti.** 2003. A direct transposon insertion tool for modification and functional analysis of viral genomes. J. Virol. **77:**123–134.
45. **Wu, B. W., J. Lu, T. K. Gallaher, W. F. Anderson, and P. M. Cannon.** 2000. Identification of regions in the Moloney murine leukemia virus SU protein that tolerate the insertion of an integrin-binding peptide. Virology **269:**7–17.
46. **Yamada, K., D. M. McCarty, V. J. Madden, and C. E. Walsh.** 2003. Lentivirus vector purification using anion exchange HPLC leads to improved gene transfer. BioTechniques **34:**1074–1078, 1080.
47. **Ye, K., S. Jin, M. M. Ataai, J. S. Schultz, and J. Ibeh.** 2004. Tagging retrovirus vectors with a metal binding peptide and one-step purification by immobilized metal affinity chromatography. J. Virol. **78:**9820–9827.
48. **Yee, J. K., T. Friedmann, and J. C. Burns.** 1994. Generation of high-titer pseudotyped retroviral vectors with very broad host range. Methods Cell Biol. **43:**99–112.
49. **Yuan, L., I. Kurek, J. English, and R. Keenan.** 2005. Laboratory-directed protein evolution. Microbiol. Mol. Biol. Rev. **69:**373–392.
50. **Zhang, H. G., J. Xie, I. Dmitriev, E. Kashentseva, D. T. Curiel, H. C. Hsu, and J. D. Mountz.** 2002. Addition of six-His-tagged peptide to the C terminus of adeno-associated virus VP3 does not affect viral tropism or production. J. Virol. **76:**12023–12031.
51. **Zhao, H., and F. H. Arnold.** 1997. Functional and nonfunctional mutations distinguished by random recombination of homologous genes. Proc. Natl. Acad. Sci. USA **94:**7997–8000.