

The Trp Cage Motif as a Scaffold for the Display of a Randomized Peptide Library on Bacteriophage T7*

Received for publication, November 20, 2006, and in revised form, January 26, 2007. Published, JBC Papers in Press, January 30, 2007, DOI 10.1074/jbc.M610722200

Richard E. Herman, Douglas Badders, Mark Fuller, Ekaterina G. Makienko, Michael E. Houston, Jr., Steven C. Quay, and Paul H. Johnson¹

From the Natestch Pharmaceutical Company, Inc., Bothell, Washington 98021

Phage libraries displaying linear or disulfide-constrained peptides often yield weak binders, upon screening against a target, and must be optimized to improve affinity. The disadvantages of libraries based on larger complex proteins, such as single chain antibodies, have stimulated interest in the development of smaller nonimmunoglobulin protein scaffolds. A promising candidate is the Trp cage motif, a 20-residue C-terminal sequence of exendin-4. Amino acid substitution within the Trp cage resulted in a 20-mer peptide recognized as an ultrafast cooperative folding miniprotein, with ideal characteristics for the discovery of small structured nonimmunoglobulin motifs having a stable tertiary structure. Although we were unable to display the Trp cage on M13 phage, successful display was achieved using the lytic T7 phage. Interestingly, mutations were observed at a frequency dependent on display valency. A Trp cage library designed with randomized amino acids at seven solvent-exposed positions was developed from 1.6×10^9 primary clones in T7Select10-3b. DNA sequencing of 109 library clones revealed 38% mutants and 16% truncations by TAG codons at randomized positions. Amino acid frequencies were largely within expected bounds and DIVAA analysis revealed that the library had an average diversity of 0.67. Utility of the library was demonstrated by identification of HPQ containing Trp cage miniproteins, which bound streptavidin, and AAAD-PYAQWLQSMGPHSGRPPPR, which bound to human bronchial epithelial cells. A high complexity library based on the Trp cage miniprotein has demonstrated potential for identifying novel cell and protein binding peptides that could be used for the delivery of therapeutic molecules or as target-specific therapeutic agents.

Since the development of phage display technology (1), it has been widely used for the screening of random combinatorial peptide libraries in drug discovery initiatives (2). When peptides are displayed as unconstrained linear molecules, they can adopt numerous conformations, very few of which may represent stable conformations. Panning with these unconstrained peptides against targets of interest usually leads to the isolation of peptides with low binding affinity. Reducing the conformational freedom of the displayed peptides is necessary to

decrease entropy and increase the binding affinity (3, 4). Constrained libraries have often achieved a reduction in conformational freedom by incorporating a pair of cysteine residues at both ends of linear peptides, resulting in disulfide bridges that form peptide loops yielding cyclic peptide libraries reported to increase binding affinity (5) and stability in biological fluids (6).

The use of defined protein scaffolds for the generation of random libraries has the potential to increase binding affinity further by providing a more rigid conformation (7). Examples of folded protein scaffolds that have been used for peptide display include the zinc finger motif, which has been used to construct a degenerate library for the isolation of peptides that bind to DNA and RNA (8, 9), and Kunitz domains, which have been used as a scaffold for libraries designed for the discovery of improved serine protease inhibitors (10).

We have evaluated another protein that has potential to serve as a scaffold for the generation of a random peptide library, exendin-4, a 39-amino acid peptide isolated from the saliva of the Gila monster, *Heloderma suspectum*. NMR studies of exendin-4 (11) showed that residues 21–38 of the C terminus form a stable tertiary structure designated a Trp cage motif. The Trp cage is a hydrophobic cluster with Trp-25 buried in a central location, where it is shielded from solvent exposure. The residues that form the cage around Trp-25 include multiple proline residues (Pro-31, and prolines 36–38) where the proline rings are located on both faces of the indole ring of Trp-25 and the Phe side chain completes the hydrophobic cluster. Further studies aimed at developing a miniprotein (a stable self-folding peptide with a minimum number of amino acids) showed that the C-terminal 20 amino acids of exendin-4 (RLFIEWLKNGGPSSGAPPPS) are adequate for forming the Trp cage. In addition, amino acid substitution experiments resulted in optimization of Trp cage folding (estimated by NMR chemical shift deviations, the presence of long range nuclear Overhauser effect, the absence of cis/trans isomerization around proline residues, and NH exchange protection) from 40 to >95% in water at physiological pH. Analysis of the conformation of the optimized Trp cage in the presence of 30% trifluoroethanol, a known helix-inducing solvent, resulted in a minimal increase in helical content. Moreover, this construct is fairly stable with a T_m of 43.5 °C (12, 13). The optimized Trp cage miniprotein TC5b (NLYIQWLKDGGPSSGRPPPS) has been extensively studied, because it is an ultrafast, cooperatively folding system (14). Within the optimized Trp cage sequence, there are several positions where the substitution of amino acids does not compromise folding of the miniprotein. Some of those positions are solvent-exposed, making them

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

¹ To whom correspondence should be addressed: Natestch Pharmaceutical Company, 3830 Monte Villa Pkwy., Bothell, WA 98021. E-mail: pjohnson@natestch.com.

ideal for the display of random amino acids in a highly diverse library.

This report describes the construction and characterization of a peptide library of high complexity based on the 20-amino acid Trp cage miniprotein TC5b with seven randomized positions displayed on bacteriophage T7. We demonstrate the library's potential utility by the identification of a peptide containing the HPQ motif that binds to streptavidin and a peptide that binds to human bronchial epithelial cells.

EXPERIMENTAL PROCEDURES

Chemicals, Media, and Reagents—General chemicals were purchased from Sigma. Components for bacterial growth media were purchased from Fisher. T7Select® cloning kits, vector arms, packaging extracts, S-Tag™ antibody, and T7 tail fiber monoclonal antibody were purchased from Novagen (EMD Biosciences, San Diego, CA). Restriction enzymes EcoRI and HindIII and streptavidin were purchased from New England Biolabs (Beverly, MA). Mammalian cell media, Klenow fragment of DNA polymerase I, MagicMark™ XP Western Protein Standards, and reagents for SDS-PAGE electrophoresis were purchased from Invitrogen. Goat anti-rabbit IgG horseradish peroxidase and goat anti-mouse IgG horseradish peroxidase antibodies were purchased from Santa Cruz Biotechnology, Inc. (Santa Cruz, CA). SuperSignal West Femto Maximum Sensitivity Substrate was purchased from Pierce.

Bacteria, Bacteriophages, and Cell Lines—*Escherichia coli* strains BL21 [F^- , *ompT*, *hsdS_B* ($r_B^- m_B^-$), *gal*, *dcm*] and BLT5403 [F^- , *ompT*, *hsdS_B* ($r_B^- m_B^-$), *gal*, *dcm* pAR5403 (*Amp^R*)] and bacteriophages T7Select® 415-1b, 10-3b, 1-1b and S-Tag™ were obtained from Novagen (EMD Biosciences). *E. coli* C600 ATCC 23724 (F^- , *supE44*, *lacY1*, *thr-1*, *leuB6*, *mcrA*, *thi-1*, *rfbD1*, *fluA21*, *lambda*⁻) was obtained from the American Type Culture Collection (Manassas, VA), and V517 was obtained from Larry L. McKay (University of Minnesota). Human bronchial epithelial cell line 16HBE14o— was obtained from D. C. Gruenert (University of California, San Francisco, CA).

Growth and Purification of Bacteriophage—T7Select 415-1b was propagated in *E. coli* strain BL21 using M9LB liquid medium. T7Select 10-3b and 1-1b were propagated in BLT5403 using M9LB supplemented with carbenicillin (50 µg/ml) at 37 °C. T7 phage was enumerated using the plaque assay method on LB² solid medium supplemented with carbenicillin (50 µg/ml) as needed. Phage lysates were prepared following the instructions in the T7Select System manual (Novagen, San Diego, CA), which typically resulted in titers of 10¹⁰–10¹¹ pfu/ml. Phage were purified and concentrated from centrifuged lysates (12,000 × *g* for 15 min at 4 °C) by the addition of one-sixth volume of 20% polyethylene glycol 8000, 2.5 M NaCl (PEG/NaCl) and incubating on ice for at least 60 min or overnight at 4 °C. Precipitated phage were pelleted by centrifugation as

above. The phage pellet was resuspended in TBS (50 mM Tris, pH 7.5, 150 mM NaCl) and reprecipitated with one-sixth volume of PEG/NaCl on ice for 15–60 min, centrifuged, resuspended in TBS, and stored at 4 °C.

Preparation of Phage DNA for Sequencing—Well-isolated phage plaques were punched out of solid media and resuspended in 100 µl of TE (10 mM Tris, pH 8.0, 1 mM EDTA). Host bacteria were grown to a density of 2 × 10⁸ cfu/ml, and 2.5 µl of the phage suspension was added to 1 ml of the bacterial culture in a 17 × 100-mm tube followed by incubation at 37 °C for 3 h with shaking. Then 1 µl of Nuclease mix containing RNase A and DNase I (Promega, Madison, WI) was added, and incubation was continued for 15 min at 37 °C without shaking. Cell debris was removed by centrifugation in a microcentrifuge (16,000 × *g*, 10 min). A 900-µl volume of the supernatant was recovered and transferred to a new 1.5-ml tube containing 200 µl of PEG/NaCl and placed at 4 °C overnight. Precipitated phage were collected by centrifugation (16,000 × *g*, 10 min), resuspended in 100 µl of TE, and disrupted by extraction with an equal volume of phenol/chloroform (1:1) and then centrifuged in a microcentrifuge (16,000 × *g*, 1 min). The aqueous phase was recovered, extracted with an equal volume of chloroform, and centrifuged for 4 min. The aqueous phase was recovered again, and one-tenth volume of 3 M sodium acetate was added. Alternatively, the phage were disrupted by resuspension in 100 µl of iodide buffer (10 mM Tris, pH 8, 1 mM EDTA, 4 M NaI). Phage DNA was precipitated by the addition of 250 µl of ethanol with incubation for 10 min at room temperature followed by centrifugation (16,000 × *g*, 10 min). The DNA pellet was washed with 70% ethanol, dried briefly under vacuum, and resuspended in 20–30 µl of TE.

DNA Sequencing/Oligonucleotide Synthesis—T7 DNA was submitted to Retrogen (San Diego, CA) for automated DNA sequencing by an ABI 3730 DNA analyzer using the primer T7-125 (5'-TGCGTGACTTGGCTCTGGAG-3'). A protocol model was constructed for sequence analysis using Teranode software (Teranode Corp., Seattle, WA), which also served as a sequence data base. Oligonucleotides were synthesized by Retrogen, using standard phosphoramidite chemistry.

Insertion of DNA Sequences Encoding TC3b or TC5b into T7Select Vectors—Oligonucleotide pairs were designed to encode the Trp cage TC3b and TC5b miniproteins (Table 1). Each oligonucleotide, 5'-phosphorylated (5'p), was resuspended in TE to 200 pmol/µl before annealing. Oligonucleotide TC3b (+) (5'p-AATTTATTTATCGAATGGCTCAA-AAATGGTGGTCTTCCAGTGGTGCTCCTCCCCCTT-CCTAA-3') was annealed with TC3b (–) (5'p-AGCTTTA-GGAAGGGGAGGAGCACCAGTGAAGGACCACCA-TTTTGTAGCCATTCGATAAATA-3'), and TC5b (+) (5'p-AATTCGGCAGCTGCGAATTTGTATATTCAGTG-GCTTAAGGATGGTGGTCTTTCGTCGGGGCGGCCTC-CGCCAAGTTAAA-3') was annealed with TC5b (–) (5'p-AGCTTTTAACTTGGCGGAGGCCGCCCGACGAAGG-ACCACCATCCTTAAGCCACTGAATATACAAATTCG-CAGCTGCCG-3') by combining 25 µl of each of the appropriate oligonucleotides in annealing buffer (10 mM Tris, pH 7.5, 100 mM NaCl, 1 mM EDTA), the mixtures were heated to 95 °C in a heat block for 5 min and allowed to cool at room

² The abbreviations used are: LB, Luria-Bertani; pfu, plaque-forming units; cfu, colony-forming units; BSA, bovine serum albumin; DPBS, Dulbecco's phosphate-buffered saline; TBS, Tris-buffered saline; PEG, polyethylene glycol; Fmoc, N-(9-fluorenyl)methoxycarbonyl; HPLC, high pressure liquid chromatography; ELISA, enzyme-linked immunosorbent assay.

temperature for 2 h in the block. The annealed oligonucleotides were ligated to the arms of T7Select 415-1b, 10-3b, or 1-1b and packaged *in vitro* using the T7 Select Cloning kit according to the manufacturer's directions. Packaged phage were amplified by growth in liquid media with the appropriate host. Oligonucleotide insertion was confirmed by DNA sequencing.

Peptide Synthesis—Peptides were synthesized by solid phase Fmoc chemistry on Tentagel amide resin using a Rainin Symphony synthesizer. Coupling steps were performed using 5 eq of 2-(6-chloro-1*H*-benzotriazole-1-yl)-1,1,3,3-tetramethylaminium hexafluorophosphate and Fmoc amino acid (NovaBiochem, San Diego, CA) with an excess of *N*-methylmorpholine for 40 min. Fmoc removal was accomplished by treating the peptide resin with 20% piperidine (Aldrich) in dimethylformamide (Burdick and Jackson, Morristown, NJ) for two 10-min cycles. Upon completion of the entire peptide, the Fmoc group was removed with piperidine and washed extensively with dimethylformamide. The peptides were cleaved from the resin by the addition of 10 ml of trifluoroacetic acid (Aldrich) containing 2.5% water and 2.5% triisopropylsilane (Aldrich) followed by gentle agitation at room temperature for 2 h. The resulting crude peptide was collected by trituration with ether followed by filtration. The crude product was dissolved in Millipore water and lyophilized to dryness. The crude peptide was taken up in 15 ml of water containing 0.05% trifluoroacetic acid and 3 ml of acetic acid and loaded onto a Zorbax RX-C8 reversed-phase column (22-mm inner diameter \times 250 mm, 5- μ m particle size) through a 5-ml injection loop at a flow rate of 5 ml/min. The purification was accomplished by running a linear AB gradient of 0.1% B/min for 180 min, where solvent A is 0.05% trifluoroacetic acid in water and solvent B is 0.05% trifluoroacetic acid in acetonitrile. The purified peptides were analyzed by reversed-phase HPLC and electrospray mass spectrometry. Peptides were purified to greater than 95% purity as determined by reversed-phase HPLC.

Production of Antibody AB167—The TC5b miniprotein was synthesized with the addition of a cysteine residue at the N terminus to allow conjugation of keyhole limpet hemocyanin and designated PN0522 (NH₂-CNLYIQWLKDGPPSSGRP-PPS-amide). The keyhole limpet hemocyanin-PN0522 conjugate was used to produce protein G-purified antibodies from New Zealand White rabbits by Orbigen (San Diego, CA).

Phage ELISAs—For the detection of Trp cage TC5b, purified phage were diluted in Dulbecco's phosphate-buffered saline (DPBS), distributed 100 μ l/well to a 96-well MaxisorpTM plate (Nalge Nunc International, Rochester, NY), and allowed to bind for 1 h. All incubations were done at room temperature with slow orbital shaking. After phage binding, the wells were washed three times with 400 μ l of DPBS containing 0.1% Tween 20 and 1% BSA (wash buffer) and then blocked with 200 μ l of wash buffer for 1 h. A 100- μ l volume of primary antibody (AB167 diluted 1:500 in wash buffer) was applied and allowed to bind for 1 h and then washed five times as above. Secondary antibody (goat anti-rabbit horseradish peroxidase diluted 1:1000 in wash buffer) was applied in 100 μ l, incubated for 1 h, washed, and developed with OptEIA reagents (BD Biosciences) according to the manufacturer's directions. Means were compared using Student's *t* test.

Phage binding to 16HBE14o– cells was analyzed using confluent cells grown with Dulbecco's modified Eagle's medium/F-12 complete medium in 96-well collagen-coated plates (BD Biosciences) at 37 °C in 5% CO₂. The cells were washed with DPBS, and phage diluted in DPBS containing 1% BSA (DPBS-BSA) were applied in 100 μ l. Bound phage were detected using primary antibody (mouse IgG T7 tail fiber monoclonal antibody diluted 1:2000 in DPBS-BSA) and secondary antibody (goat anti-mouse IgG horseradish peroxidase diluted 1:5000 in DPBS-BSA) followed by detection with OptEIA reagents as above.

Western Analysis of Phage Displaying TC5b—Purified phage were denatured by boiling in Tris/glycine SDS sample buffer (2 \times) with 1% β -mercaptoethanol. Samples containing 10¹¹ pfu/ml in 10 μ l were loaded into wells of a 1-mm 4–20% Tris/glycine gel and electrophoresed for 1.5 h at 150 V (constant voltage) in Tris/glycine SDS running buffer. The separated proteins were transferred to a polyvinylidene difluoride membrane for 1.5 h at room temperature and blocked with 3% nonfat dry milk in TBS containing 0.05% Tween 20 (TBS-T) for 1 h. The membrane was probed with antibody AB167 diluted 1:2500 for 2 h at room temperature, followed by three 5-min washes with TBS-T. Bound antibody was detected by incubation with goat anti-rabbit IgG horseradish peroxidase diluted 1:5000 for 1 h at room temperature followed by four 10-min washes and development with SuperSignal West Femto maximum sensitivity substrate according to the manufacturer's instructions. MagicMarkTM XP Western protein standards were used for size references.

Construction of the Trp Cage Library—The Trp cage library was constructed in the T7Select 10-3b vector using oligonucleotide T-3 (–) (5'-CATGTTCAAGCTTGTAMNNGGGGGAGGACGACCAGAMNNAGGACCMNNMNMNN-TAACCACTGMNNGTAMNNATCTGCTGCTGCCGAATTCGGGGCATGTG-3') as a template (where M represents equal molar amounts of A and C, and N is equal molar amounts of A, C, G, and T) and T-3 (+) (5'-CACATGCCCCGAATTCGGCA-3') as a primer. The template and primer were annealed by combining 310 pmol of T-3 (–) and 820 pmol of T-3 (+) in annealing buffer (10 mM Tris, pH 7.5, 100 mM NaCl, 1 mM EDTA) to a final volume of 50 μ l, the mixture was heated to 95 °C in a heat block for 5 min, and the block was allowed to cool at room temperature for 2 h. The primer was extended by combining 310 pmol of annealed template-primer, dNTP mixture (New England Biolabs, Beverly, MA), REact 2 buffer (Invitrogen), and Klenow fragment (10 units) in a total volume of 250 μ l for 10 min at 37 °C. The reaction was stopped by incubation at 65 °C for 15 min followed by phenol/chloroform extraction and ethanol precipitation. The extended duplexes were double digested with EcoRI and HindIII, and the cleavage products were subjected to electrophoresis in a 20% TBE (89 mM Tris, 89 mM boric acid, 2 mM EDTA) gel. The desired band was excised and extracted by shaking overnight in 100 mM sodium acetate, pH 4.5, 1 mM EDTA, 0.1% SDS at room temperature. Gel fragments were removed by centrifugation, and the DNA was purified by phenol/chloroform extraction and ethanol precipitation before being resuspended in TE. The extracted inserts were ligated into arms of T7Select[®]10-3b and packaged *in vitro* using

the T7Select® cloning kit according to the manufacturer's directions, and primary clones were enumerated. Ligation and packaging was repeated until greater than 10^9 primary clones were generated and amplified. The amplified clones were pooled, treated with Nuclease mix at 37 °C for 15 min and centrifuged ($8000 \times g$, 10 min) to remove cell debris. Phage were concentrated by adding solid PEG 8000 to 10% (w/v), incubation overnight at 4 °C, centrifugation ($10,400 \times g$, 10 min, 4 °C), and resuspension in 1 M NaCl, 10 mM Tris, pH 8, 1 mM EDTA. For further purification, the concentrated phage were layered on top of a CsCl density block gradient consisting of 1.5 ml of 62%, 3.0 ml of 41%, 3.0 ml of 31%, and 3.0 ml of 20% CsCl in TE and centrifuged at 28,000 rpm ($150,000 \times g$) in a Beckman SW28.1 rotor at 25 °C for 98 min. The phage band was recovered by side puncture and dialyzed against three changes of 20 mM Tris, pH 8.0, 100 mM NaCl, 6 mM $MgSO_4$.

Construction of a Suppressor Host Strain for T7Select—Plasmid pAR5403 was isolated from BLT5403 using a QIAprep spin miniprep kit (Qiagen, Valencia, CA) following the manufacturer's instructions. *E. coli* strain C600 ATCC 23724 was made competent for transformation by inoculating LB broth with a 1% inoculum of fresh overnight culture and incubating at 37 °C with shaking until the cell density reached $A_{600} = 0.3$. Cells were pelleted by centrifugation ($480 \times g$, 10 min), resuspended in ice-cold 0.1 M $CaCl_2$, recentrifuged, and resuspended in one-fiftieth volume of ice-cold 0.1 M $CaCl_2$. Transformation was accomplished by combining 200 μ l of competent cells with 2 μ l of pAR5403 (200 ng/ μ l), incubating on ice for 30 min, heat shocking at 42 °C for 90 s, and then placing on ice for 2 min. Transformants were selected by inoculating 100 μ l of heat-shocked cells into 3 ml of molten top agar, which was then spread onto LB solid medium containing 50 μ g/ml carbenicillin and incubated at 37 °C overnight. Plasmid DNA was prepared from transformants by QIAprep Spin Miniprep and examined after electrophoresis (65 V for 2 h) through a 10×15 -cm 0.6% SeaKem LE (Cambrex, Rockland, ME)-agarose gel in TBE running buffer followed by staining with ethidium bromide (0.5 μ g/ml).

Stability of Primary Clones—BLT5403 grown to $A_{600} = 0.85$ (1×10^8 cfu/ml) was split into three aliquots and infected at a multiplicity of infection of 0.02 with a total of 4×10^7 pfu of either mutant, nonmutant, or both mixed in a 1:1 ratio. The infected cultures were incubated with shaking at 37 °C for 3 h or until lysis was complete and progeny were enumerated. Progeny phage plaques were randomly selected and suspended in 100 μ l of TE to elute phage. Lysates were prepared by infecting 1-ml cultures of midlog BLT5403 with 10 μ l of each eluate and concentrated with PEG/NaCl, and phage DNA was extracted for sequencing.

One-step Growth Experiment—Based on the one-step growth experiment described by Edgar (15), BLT5403 was grown in LB broth at 37 °C to a density of 2×10^8 cfu/ml and centrifuged, and the cell pellet was resuspended in half the original volume with LB broth and enumerated by the spread plate method to determine input cells. 0.5 ml of resuspended cells at 4×10^8 cfu/ml were placed in a 17×100 -mm tube and incubated at 30 °C for 5 min, and then phage infection was initiated by adding 0.5 ml of the appropriate phage stock adjusted to 4×10^9

pfu/ml to provide a multiplicity of infection of 10, and incubation was continued. At 5 min postinfection, 2 μ g of T7 tail fiber monoclonal antibody was added to neutralize unadsorbed phage. At 10 min postinfection, the cell and phage mixture was diluted 10^4 -fold into LB broth contained in a 250-ml growth flask, and incubation was continued with vigorous shaking. At 15 min postinfection, 1.0 ml was removed from the growth flask, mixed with 1 ml of LB broth containing 2–3 drops of $CHCl_3$, and vortexed. The unadsorbed phage were enumerated. At 20 min postinfection, 1.0 ml was removed from the growth flask for determination of infected cells (total pfu – unadsorbed pfu). At 90 min postinfection, 2–3 drops of $CHCl_3$ were added to the growth flask, and progeny phage were enumerated.

Streptavidin Panning—A 96-well Maxisorp™ plate was coated with 200 ng/well streptavidin (2 ng/ μ l) for at least 1 h and blocked with 1% BSA in DPBS with 0.1% Tween 20 (DPBS-BT) for 1 h and then rinsed three times with 200 μ l of DPBS-BT. The crude Trp cage library (*i.e.* before CsCl purification) was panned against streptavidin-coated wells by applying 1.2×10^9 pfu that were allowed to bind for 1 h at room temperature followed by six washes with 200 μ l of DPBS-BT. Bound phage were eluted with 1% SDS for 10 min and amplified with BLT5403. The lysate was centrifuged, and 10 μ l of the phage supernatant was used for the next round of panning. After three rounds of panning, well isolated plaques of eluted phage were selected for DNA sequencing.

Panning against Human Bronchial Epithelial Cells—Human bronchial epithelial cells (16HBE14o–) at low passage were seeded into 24-well collagen-coated plates (BD Biosciences) and allowed to grow at 37 °C in 5% CO_2 to confluence in Dulbecco's modified Eagle's medium/F-12 complete medium containing 10% fetal bovine serum. To prepare the cells for panning, the medium was removed and replaced with Dulbecco's modified Eagle's medium/F-12 with no serum (F12i) and incubated at 37 °C for 1.5 h. Then the F12i was removed and replaced with 500 μ l/well of 1% BSA in DPBS and incubated at 37 °C for 60 min. Just before panning, the DPBS was removed, and the cells were rinsed with 500 μ l of TBS. A 50- μ l aliquot of the naive T7Select Trp cage library was combined with 160 μ l of TBS and added to the cells followed by incubation at 4 °C for 60 min with gentle rocking. The unbound portion of the phage library was removed by 25 washes with 500 μ l of TBS. Bound phage were eluted by adding 200 μ l of a BLT5403 culture grown to $A_{600} = 1.0$ in M9LB followed by incubating at room temperature for 15 min. Eluted phage were added to 20 ml of the BLT5403 culture and amplified by incubation at 37 °C for 2–3 h until cell lysis was observed. The amplified phage was concentrated with PEG/NaCl and resuspended in 200 μ l of DPBS. Subsequent rounds of panning were performed using 20 μ l of concentrated phage from the previous round. In the second round, unbound phage were removed by washing with TBS containing 0.1% Tween, and in the third round TBS with 0.3% Tween was used. After three rounds of panning, well-isolated plaques of eluted phage were selected for DNA sequencing.

RESULTS

Display of Trp Cage Miniproteins on Bacteriophage T7 Is Associated with a Valency-dependent Increase in Peptide Mutation—The T7Select system, which is designed to display peptides with a free C terminus on the surface of bacteriophage T7 by fusion to the phage capsid protein (gp10B), was used to display Trp cage miniproteins (Table 1) that were based on sequences from Neidigh (13). Initial work designed to test the feasibility of displaying the Trp cage miniprotein was performed with Trp cage version TC3b, and later work focused on the further optimized version TC5b because it has better folding characteristics. Three T7Select vectors were used, each one varying by the average number of capsid fusion proteins displayed per phage: 415-1b (average copy number = 415), 10-3b (average copy number = 5–15), and 1-1b (average copy number = 0.1–1). In order to determine if Trp cage TC3b could be displayed on bacteriophage T7, oligonucleotides TC3b (+) and TC3b (–) were annealed and inserted into the three T7Select vectors. After *in vitro* packaging and amplification, several well-isolated plaques were selected and analyzed by DNA sequencing. The deduced amino acid sequences revealed a high frequency of mutants (7 of 8 clones) when TC3b was expressed in the 415-1b vector and a low frequency of mutants in the 10-3b and 1-1b vectors (1 of 10 and 1 of 9 clones, respectively). A comparison of the expected TC3b peptide sequence and representative mutants (Table 2) from expression in 415-1b (mutants 415-1b A1, A2, and Ai5) revealed single base changes in the mutant sequences that resulted in nonsense or missense

mutations that prevented expression of the Trp codon TGG by truncating the peptide upstream or changing the codon so that an amino acid substitution occurred. Mutants that resulted from expression of TC3b in 10-3b and 1-1b contained deletions that resulted in frame shifts, which altered the amino acid sequence, including changes to the amino acids that are required for the Trp cage structure. Since expression of a high copy number of Trp cage protein fusions in vector 415-1b led to a high percentage of mutants, the medium valency vector 10-3b was utilized for the remaining library work.

Trp Cage Miniproteins Are Displayed on T7 through Fusion to gp10B—To show that the Trp cage was displayed as a result of fusion to the capsid protein (gp10B) of T7Select10-3b, proteins from phage lysates were separated by SDS-PAGE, and binding of Trp cage-specific antibody AB167 to lysate proteins was detected by Western blot analysis. As shown in Fig. 1, the antibody bound to a protein in the lane containing the lysate from cells infected with phage expressing the TC5b insert. This band was not observed in the lanes containing lysates from cells infected by phage with no insert or phage displaying the linear 15-amino acid S-TagTM or in lysates from uninfected host bacteria. The size of the positive band is consistent with the predicted size of 39,640 daltons (376 amino acids) for the Trp cage-gp10B phage capsid protein fusion. As expected, Western analysis using an antibody against the S-TagTM instead of the Trp cage resulted in detection of a band only in the S-TagTM lane, which had the same relative mobility as the Trp cage-gp10B phage capsid protein fusion (data not shown).

Display of Trp Cage Miniproteins Reduces the Yield of Progeny Phage—One-step growth experiments were conducted in *E. coli* BLT5403 to determine if display of the Trp cage affected the production of progeny phage as measured by the average burst size. Each experiment compared three phage: T7Select10-3b with no displayed peptide, T7Select10-3b dis-

TABLE 1
Amino acid sequence of Exendin-4 and Trp cage miniproteins

Exendin-4 (residues 20–39)	RLFIEWLKNKGPGSSGAPPPS
TC3b	NLFIEWLKNKGPGSSGAPPPS
TC5b	NLYIQWLKDGPGSSGRPPPS

TABLE 2
Amino acid and DNA sequences from mutants resulting from insertion of TC3b into T7Select 415-1b, 10-3b, and 1-1b

TC3b	N L F I E W L K N G G P S S G A P P P S stop aat tta ttt atc gaa tgg ctc aaa aat ggt ggt cct tcc agt ggt gct cct ccc cct tcc taa
415-1b A1	N stop aat tga ttt atc gaa cgg ctc aaa aat ggt ggt cct tcc agt ggt gct cct ccc cct tcc taa
415-1b A2	N L F I E stop aat tta ttt atc gaa tag ctc aaa aag ggt ggt cct tcc agt ggt gct cct ccc cct tcc taa
415-1b Ai5	N L F I E S aat tta ttt atc gaa tgc ct n tnt aaa g ggt ggt cct tcc
10-3b	N L F I E G S K M V V L P R G A P P P S stop aat tta ttt atc gaa_ggc tca aaa atg gtg gtc ctt cca cgt ggt gct cct ccc cct tcc taa Note: mutations are deletion of a 't' at _ and insertion of a 'c' at c.
1-1b	N L F I E W L K N G G L P V V L L P L P aat tta ttt atc gaa tgg ctc aaa aat ggt ggt_ctt cca gtg gtg ctc ctc ccc ctt cct aa Note: mutation is a deletion of a 'c' at _.

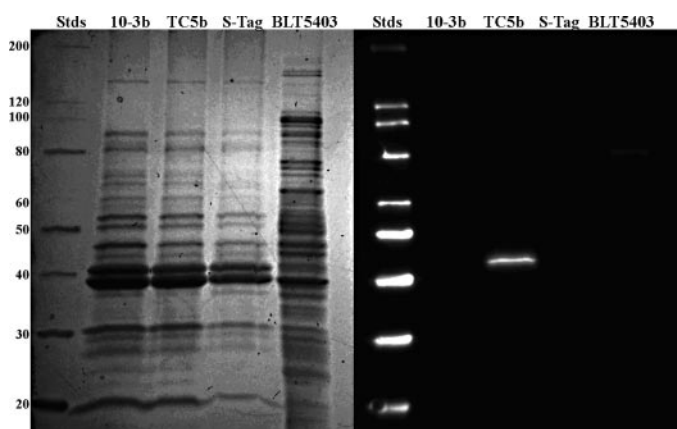


FIGURE 1. Western blot analysis of Trp cage TC5b displayed on T7Select10-3b. Proteins in lysates from phage-infected cells were separated by SDS-PAGE as shown in the Coomassie-stained gel on the left. The lysates were prepared from cells that were infected with T7 displaying either no peptide (10-3b), Trp cage TC5b (TC5b), or the S-TagTM peptide (S-Tag). The lane labeled BLT5403 contained lysate from uninfected cells. The Western transfer on the right shows that only the lysate from infection by phage displaying Trp cage TC5b contained a protein that bound the Trp cage-specific antibody AB167. MagicMarkTM XP Western protein standards (Stds) were loaded in the first lane and are visible in the gel and the Western blot with sizes indicated in kDa.

TABLE 3

Burst size analysis from one-step growth experiments (n = 3)

Burst sizes were compared by Student's *t* test: 10-3b > 10-3b/TC5b ($p = 0.001$), 10-3b > 10-3b/S-Tag ($p < 0.05$) and 10-3b/S-Tag > 10-3b/TC5b ($p < 0.05$).

Phage tested	10-3b	10-3b/ S-Tag	10-3b/ TC5b
Average infected cells (cfu/ml)	3.56×10^8	3.88×10^8	4.20×10^8
Average burst size \pm S.D.	216.9 \pm 20.2	157.1 \pm 23.3	104.4 \pm 11.1

playing the 15-amino acid S-TagTM, and T7Select10-3b displaying Trp cage TC5b. As shown in Table 3, the results from three independent one-step growth experiments indicated that although the average number of infected cells was similar for all phage tested, the burst size (a ratio of progeny phage to infected cells) of phage displaying Trp cage TC5b was lower than phage with no displayed peptide by greater than 50% ($p = 0.001$). Phage displaying the linear 15-amino acid peptide S-TagTM gave an intermediate value reflecting a 30% reduction ($p < 0.05$). These results indicate that displaying the Trp cage, even at low valency, reduces the average yield of progeny phage produced per infected cell greater than a similar size linear peptide.

The Diversity of the T7 Trp Cage Library Is Similar to That of M13 Ph.D.-C7CTM—A library displaying the Trp cage TC5b miniprotein (Table 1) with seven randomized positions on bacteriophage T7Select10-3b was produced by the primer extension method. The primer T-3 (+) and template T-3 (–) oligonucleotides were designed to produce an insert in T7 such that the coding strand would include codons commonly utilized by *E. coli* and phage T7 for the invariable amino acids and the degenerate codon NNK at the seven randomized positions where N represents G/A/T/C and K represents G/T. Utilizing the NNK codon reduced the number of possible amino acid codons at the randomized positions from 64 to 32 to provide a more uniform distribution of amino acids and eliminate all of the stop codons except TAG (amber). Expression of the insert

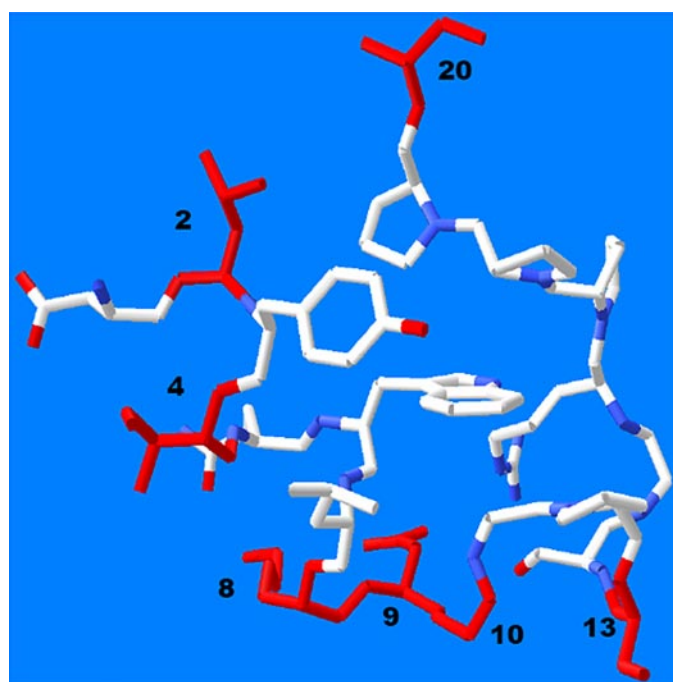


FIGURE 2. NMR structure of Trp cage TC5b (13) (NCBI accession number 1LY_A). Based on analysis of the Trp cage sequence that indicated which amino acids must be conserved for Trp cage folding, seven numbered positions (highlighted in red) that were nonconserved and solvent-exposed were chosen for randomization in the library Trp cage structure. Two corresponding sequences below the molecular structure represent the TC5b sequence (top) and the library sequence with variable positions, marked X (bottom), and an indication of the α -helix and a 3_{10} helix regions.

was designed to result in fusion of the 23-amino acid miniproteins AAADXXQWLXXXGPXSGRPPPX (where X indicates a randomized position) to the phage capsid protein. Due to their high helical propensity, three alanine residues were encoded at the N terminus to initiate and propagate the N-terminal helix and to stabilize the protein (16), followed by aspartic acid, substituted for the Asm-1 of TC5b, to act as a helix-capping residue. The positions chosen for randomization are not required for folding (13) and are located on the solvent-exposed portion of the folded protein as determined from the NMR structure (Fig. 2), where they are available to interact with potential binding sites on target molecules.

A total of 1.6×10^9 packaged primary phage clones were obtained in order to provide a library complexity approaching the number of possible combinations ($20^7 = 1.28 \times 10^9$). To determine the expected proportion of the library expressing the Trp cage as designed, DNA from 100 primary phage clones was sequenced. Among the 100 clones, 64% contained the expected sequence (of which 8% contained the stop codon TAG in one of the randomized positions), and 36% contained frameshift mutations within the Trp cage sequence.

Since a high percentage of the primary clones contained mutations in the Trp cage sequence, we were concerned that

TABLE 4

Mutant DNA sequence and resulting peptide sequence from the two major types of mutants observed in the Trp cage T-3 library

An underscore indicates a deletion, and a substitution is underlined.

T-3	aat tcg gca gca gca gat nnk tac nnk cag tgg tta nnk nnk nnk ggt cct nnk tct ggt cgt cctccc ccc ccc taa NSA AADXYXQWLXXXGPXSGRPPPX*
Type I	aat tcg _cag cag cag atc cgt act tgc agt ggt taa NSQQQIRTC SG*
Type II	aat tcg gca gca gca gat agg tac cat cag tgg tta agg ctt tct ggt cct ccg tct ggt cgt <u>ccc</u> tcc ccc ccg gta aca agc ttg cgg ccg cac tcg agt aac tag NSA AADRYHQWLRLSGPPSGRPSPPVTSLRPHSSN*

further mutations during library amplification might occur. The genetic stability of nonmutant and mutant primary phage clones was examined by selecting one of each type and infecting BLT5403 with each phage separately or mixed in a 1:1 ratio. When progeny from the infections were examined by DNA sequencing, the mutant and nonmutant phage only produced progeny (10 of 10 for each) with exactly the same DNA sequence as the corresponding parental phage. When 20 progeny plaques were picked and sequenced from the mixed infection, there were nine mutant sequences and 11 nonmutant sequences. In each case, the mutant and nonmutant sequences matched the sequence of the corresponding parental phage. The results from the infections with a single parental phage, either mutant or nonmutant, suggest that the mutant and nonmutant genotypes are stable. Moreover, since the mixed infection produced progeny in a ratio nearly equal to the ratio of the input parental phages, we conclude that there is no apparent growth advantage for the mutant or nonmutant clones.

The primary clones were amplified once before purification by a CsCl step gradient. After amplification, the naive library was characterized by randomly selecting 109 individual clones for DNA sequencing. Three categories of library members were observed. There were 50 clones (46%) with deduced amino acid sequences that contained the expected Trp cage scaffold, including seven randomized positions, and these clones were classified as nonmutants. There were 18 clones (16%) that contained the designed scaffold but had a stop codon in one of the randomized positions, and they were classified as stops. The remaining 41 clones (38%) had amino acid sequences that did not contain the expected sequence, and these clones were classified as mutants. Most of the mutations observed were single base deletions or insertions that resulted in missense or nonsense mutants. The most common mutation observed (Type I) was the deletion of a G, indicated by an underscore in Table 4, at the beginning of the protein fusion where the capsid protein ends and the Trp cage starts, which changed the reading frame of the inserted sequence and resulted in a sequence encoding a truncated protein.

The second most common mutant (Type II) resulted from the insertion of a C, indicated by c in Table 4, in the area between the sixth and seventh randomized positions. These mutants usually resulted in altered reading frames that continued past the designed stop codon (TAA) until an alternate stop

codon was reached downstream. The change in reading frame altered the amino acid sequence in the area of the three proline residues that have been shown to be important for formation and stability of the Trp cage fold.

The frequency of the amino acids in the randomized positions of 50 nonmutant clones is shown in Table 5. Confidence interval bounds for the frequency of amino acids in each of the seven randomized positions were determined using the Clopper-Pearson method for the determination of the exact confidence bounds for a binomially distributed random variable (17) and was based on the use of a 32-codon genetic code. An adjustment was made to the usual 95% confidence level based on the simplest Bonferroni correction for the seven positions ($100 \times (1 - 0.05/7) = 99.3\%$). The resulting bounds are 0–7 for amino acids encoded by only one codon, 0–10 for amino acids encoded by two codons, and 1–13 for amino acids encoded by three codons. The amino acid frequencies all fall within the expected bounds except the arginine at position 5, which was zero but was expected to be at least 1 with a confidence of 99.3%.

In addition, the amino acid sequences deduced from the 50 nonmutant clones were analyzed using the DIVAA program available at the RELIC Web site (relic.bio.anl.gov/) to determine the diversity of the amino acids by position in the library. As shown in Fig. 3, the diversity of the randomized positions ranges from 0.55 to 0.75 with overlapping error bars. For comparison, data available at the Web site were used to generate a plot of the diversity for the M13 Ph.D.-C7CTM library. The Trp cage library has diversity that is comparable with the constrained M13 Ph.D.-C7CTM library at all positions except for position 3, where the Trp cage library exhibited lower diversity. The average diversity for the Trp cage library was 0.67, compared with 0.70 for Ph.D.-C7CTM.

The C600CR Host Cell Line Effectively Suppresses Premature Peptide Termination—Since phage display systems typically encode randomized amino acids using the reduced 32-codon genetic code, which includes one stop codon, TAG, a suppressor host is utilized to prevent premature peptide termination resulting from incorporation of the stop codon. Because a suppressor host is not available for the T7Select system, a host capable of suppressing the TAG stop codon while limiting the number of peptides displayed on the T7 phage capsid was obtained by transforming competent cells of the carbenicillin-

TABLE 5

Amino acid frequency at randomized positions of Trp cage library T-3 based on 50 nonmutant clones

Amino acid (codons)	Randomized positions							Total	Frequency ^a	
	1	2	3	4	5	6	7		T-3	Ph.D.-C7C
Basic										
His (1)	4	3	0	4	4	2	0	17	4.9	6.9
Arg (3)	4	5	4	7	0	2	4	26	7.4	4.3
Lys (1)	0	0	0	2	2	2	2	8	2.3	3.8
Nonpolar										
Ile (1)	0	4	1	1	0	2	0	8	2.3	2.1
Phe (1)	0	2	0	0	0	1	2	5	1.4	2.1
Leu (3)	8	8	4	5	4	5	7	41	11.7	9.6
Trp (1)	1	2	0	0	3	0	1	7	2	1.9
Ala (2)	7	2	4	4	4	2	4	27	7.7	6.5
Met (1)	1	0	0	0	2	3	1	7	2	3.3
Pro (2)	5	5	6	5	8	8	6	43	12.3	10.7
Val (2)	4	0	3	3	2	2	0	14	4	1.9
Polar										
Cys (1)	0	1	1	0	3	2	0	7	2	0
Asn (1)	0	0	2	3	2	0	4	11	3.1	6.4
Gly (2)	2	2	1	1	1	2	1	10	2.9	2.2
Ser (3)	4	5	11	8	2	3	4	37	10.6	8.6
Gln (1)	3	1	1	2	2	1	5	15	4.3	7.1
Tyr (1)	3	2	2	2	0	0	5	14	4	2.4
Thr (2)	3	6	7	2	7	9	2	36	10.3	13.1
Acidic										
Asp (1)	1	1	3	1	2	3	2	13	3.7	4.1
Glu (1)	0	1	0	0	2	1	0	4	1.1	3.1
Basic (5)	8	8	4	13	6	6	6	51	14.6	15
Nonpolar (13)	26	23	18	18	23	23	21	152	43.4	38.1
Polar (14)	15	17	25	18	17	17	21	130	37.1	39.8
Acidic (2)	1	2	3	1	4	4	2	17	4.9	7.2

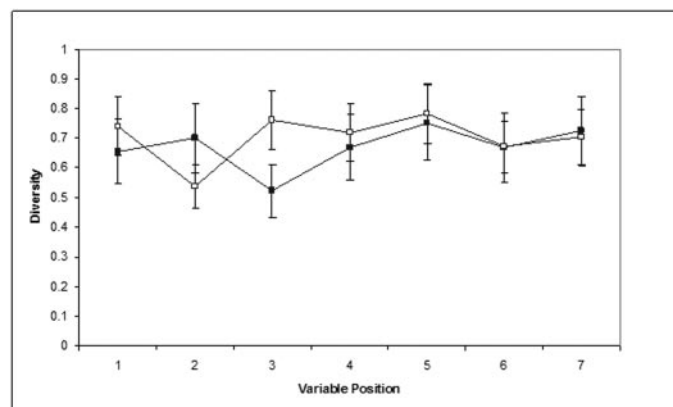
^a Frequency of the Ph.D.-C7CTM library from Ref. 32.

FIGURE 3. **Diversity of the Trp cage library as determined by DIVAA analysis.** Diversity of 50 nonmutant clones from the Trp cage library (■) was compared with the diversity of clones from the M13 Ph.D.-C7CTM library (□) obtained on the RELIC Web site. The error bars represent one S.D.

sensitive (Carb^S) strain C600 (which contains suppressor mutation *supE44* that inserts a glutamine at TAG codons) with purified plasmid pAR5403, which encodes carbenicillin resistance (Carb^R) and provides excess 10A capsid protein to limit the number of displayed peptides. A clone of C600 transformed with pAR5403 was isolated as a Carb^R colony and designated C600CR. Plasmid transformation was confirmed by agarose gel electrophoresis of purified plasmid DNA from cell lysates. As shown in Fig. 4, the pAR5403 plasmid donor strain BLT5403 contained a single plasmid estimated to be 3.0 MDa by comparison with the relative mobility of reference plasmids obtained from *E. coli* V517 (18). The parental recipient strain C600 contained no detectable plasmid DNA, whereas the Carb^R trans-

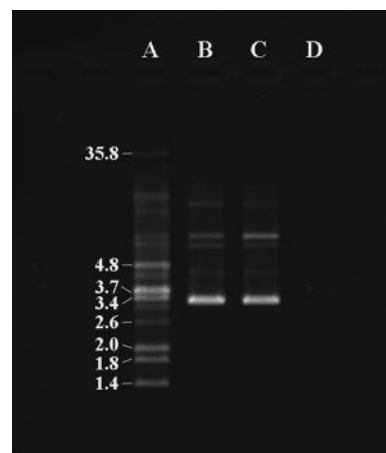


FIGURE 4. **Agarose gel electrophoresis analysis of purified plasmid DNA.** Transformation of plasmid pAR5403 into the *supE44*-expressing host C600 was demonstrated by comparing plasmid DNA extracted from the following *E. coli* strains: V517, a plasmid reference standard (A); BLT5403, plasmid donor strain (B); C600CR, transformant (C); and C600, plasmid-free recipient (D). Sizes of the reference plasmids in MDa are indicated.

formant C600CR contained a plasmid the same size as pAR5403. The presence of gene 10A in the transformed host was confirmed by PCR (data not shown).

Growth of phage containing the insert encoding Trp cage TC5b in the host BLT5403, which limits expression to 5–10 copies of the peptide, produced clones displaying the expected sequence NLYIQWLKDGGPSSGRPPPS with a low frequency of mutations (1 of 12 sequenced clones). When the same phage clones were grown in the host BL21, which does not limit the expression of the peptide, mutants were obtained at high fre-

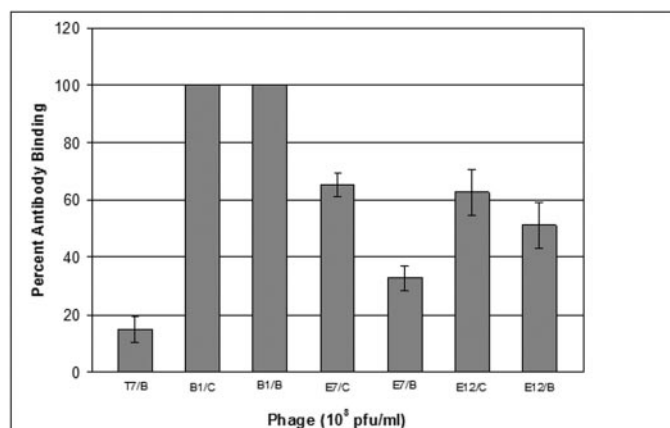


FIGURE 5. Phage ELISA for demonstration of TAG suppression. Binding of antibody AB167 to T7Select10-3b phage grown in the nonsuppressor host BLT5403 (B) was compared with binding to phage grown in the suppressor host C600CR (C). The comparison included phage without a displayed peptide grown in the nonsuppressor host (T7/B) and phage displaying TC5b grown in both hosts (B1/B and B1/C). Binding of the antibody to phage mutants E7 and E12 displaying truncated Trp cages grown in B (E7/B and E12/B, respectively) was related to antibody binding to phage clone B1 displaying the full-length Trp cage on a percentage basis. The phage mutants grown in C600CR (E7/C and E12/C) were likewise related to B1 grown in C600CR. The error bars represent one S.D.

quency (6 of 6 sequenced clones). Some of the mutant clones contained the TAG stop codon within the Trp cage sequence and were expected to result in the display of truncated peptides. Two nonsense mutants, E7 (which contained a TAG stop codon at Tyr-3 of the Trp cage sequence (NL*IQWLKDGGPSSGRPPPS)) and E12 (which contained a TAG stop codon at Trp-6 of the Trp cage sequence (NLYIQ*LKDGGPSSGRPPPS)), were chosen to test the effectiveness of the suppressor host C600CR by ELISA. Each mutant and phage clone B1, which displays TC5b without a stop codon, was grown in the nonsuppressor strain BLT5403 and the suppressor strain C600CR. T7Select10-3b, which does not have a displayed peptide, was grown in BLT5403. Phage displaying Trp cage TC5b were detected using rabbit IgG polyclonal antibody AB167, which was raised against TC5b and demonstrated to bind to phage displaying the TC5b peptide by Western analysis (see above). Since AB167 is a polyclonal antibody, some binding to the truncated peptides of E7 and E12 was expected, with possibly greater binding to E12 because it is 3 residues longer than E7. As shown in Fig. 5, the ELISA indicated that the antibody detected display of the TC5b peptide when B1 was grown in either host (B1/B and B1/C), and there was minimal binding of the antibody to phage T7Select10-3b (T7/B), which has no displayed peptide. The percentage of binding to E12 after growth in BLT5403 (E12/B) was numerically lower than when grown in C600CR (E12/C), but the difference was not significant ($p > 0.1$). Antibody binding to mutant E12 was lower than the binding to B1 regardless of the host. Antibody binding to mutant E7 was significantly higher ($p < 0.01$) when grown in the suppressor C600CR (E7/C) compared with growth in the nonsuppressor BLT5403 (E7/B), and both exhibited lower binding than B1. The results from E7 indicate that C600CR was able to suppress the codon TAG in the T7Select phage. Values lower than B1 were expected, since suppression by *supE44* is not 100% efficient (19, 20).

TABLE 6

Peptide sequences obtained from panning against streptavidin (variable positions are in boldface type, and the expected HPQ motif is underlined)

Clone number	Peptide sequence
255-104-3	AAADY YLQWL <u>HPQ</u> GPHSGRPPPA*
255-104-5	AAAD SYRQWL <u>HPQ</u> GPWSGRPPPA*
255-104-2	AAADLY RQWL <u>ALC</u> GPQSGRPPPT*
255-104-4	AAADHY RQWL <u>LGL</u> GPQSGRPPPC*
255-104-9	AAADKY TQWL <u>SIAG</u> PQSGRPPPQ*
255-104-10	AAADLY NQWL <u>THSG</u> PRSGRPPPN*
255-104-11	AAADAY MQWL <u>RYHG</u> PFSGRPPPT*

TABLE 7

Peptides obtained from panning against human bronchial epithelial cell line 16HBE14o—

Clone number	Peptide sequence (variable positions in boldface type)
Full-length Trp cage	
624-55-20	AAADRY PQWL <u>NGM</u> GPSSGRPPPN*
624-55-21	AAADAY PQWL <u>FTT</u> GPSTSGRPPPL*(5) ^a
624-55-29	AAADPY AQWL <u>QSM</u> GPSSGRPPPR*
624-55-36	AAADMY AQWL <u>DNM</u> GPSSGRPPPY*
Mutant Trp cage	
624-55-24	AAADAY SQWL <u>LQT</u> GPYSGRPPSPRV
624-55-32	AAADTY IQWL <u>KING</u> PRLVLPSPRN
624-55-33	AAADAY GQWL <u>RTS</u> GPLSGRSLPPR
624-55-37	AAADWY QWL <u>PPG</u> GPSSGRPPPHL
Truncated Trp cage	
624-55-22	AAADLY LQWL <u>D</u> *PGPSSGRPPPLL
624-55-26	AAAD*Y TQWL <u>YLQ</u> GPNSGRPPPY*(5) ^a

^a There were five instances of these peptides among the clones sequenced.

The T7 Trp Cage Library Allows Rapid Identification of Peptides Containing the HPQ Motif—Streptavidin has been used as a model receptor system for analysis of linear and constrained phage display libraries yielding the well known His-Pro-Gln (HPQ) motif (5). We used streptavidin as a target to assess the naive Trp cage library for the display of peptides containing the HPQ motif in a limited panning experiment using 1.2×10^9 pfu of the amplified library before final purification by CsCl. After panning the Trp cage library against streptavidin for three rounds, a small number of clones were selected for DNA sequencing. There were three mutants among the 10 clones analyzed, and, as shown in Table 6, two of the seven nonmutant clones contained the expected HPQ motif (AAADY-YLQWLHPQGPHSGRPPPA and AAADSYRQWLHPQGPWSGRPPPA). This result demonstrated that a motif reported to bind to streptavidin after panning with linear and constrained phage libraries can quickly be obtained from the Trp cage library. Within our small sample size, the HPQ motif was only observed at the three randomized positions that are contiguous, which is located at the junction of the α -helix and the 3_{10} helix of the Trp cage fold.

The T7 Trp Cage Library Enables Selection of Cell-binding Peptides—One of the potential uses of the Trp cage library is to rapidly discover peptides with affinity to human cell types for the development of cell-specific targeting ligands. In order to evaluate the library's capability to produce such peptide leads, it was panned against the human bronchial epithelial cell line 16HBE14o—. After three rounds of panning, the peptide sequences displayed by 18 randomly selected clones were determined from DNA sequence analysis. As shown in Table 7, there were four full-length Trp cage sequences, including five

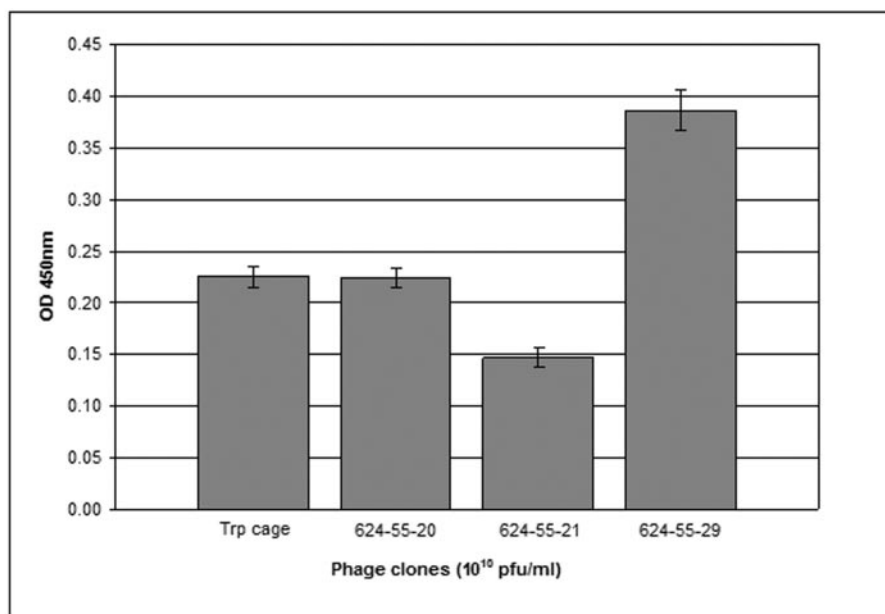


FIGURE 6. **Phage ELISA for detection of binding to unfixed 16HBE14o– cells.** Binding of phage T7Select10-3b displaying Trp cage TC5b (Trp cage) was compared with binding of three phage clones selected from the naive Trp cage library after three rounds of panning against the human bronchial epithelial cell line 16HBE14o–. Results represent the average of triplicate wells, and error bars represent one S.D.

instances of the same sequence, which is represented in the table by clone number 624-55-21. There were also five instances of the truncated peptide represented by clone number 624-55-26. Binding of the displayed peptides to 16HBE14o– cells was evaluated by phage ELISA using unfixed cells. As shown in Fig. 6, binding to 16HBE14o– cells was lower for the negative control phage displaying the TC5b version of the Trp cage (NLYIQWLKDGGPSSGRPPPS) and for two of the phage clones selected at random after panning against the 16HBE14o– cells; phage clone 624-55-29 displaying the peptide AAADPYAQWL**QSMGPHSGRPPPR** (amino acids in the variable positions are indicated in boldface type) demonstrated significantly higher binding to 16HBE14o– cells, compared with the negative control. This result indicated that the Trp cage library can be used to select peptides with cell binding capability.

DISCUSSION

The Trp cage miniprotein has been optimized by amino acid substitution to produce versions with improved folding characteristics, including the versions utilized in this study, TC3b (NLFIEWLKNGGPSSGAPPPS) and TC5b (NLYIQWLKDGGPSSGRPPPS), which fold 38 and >95%, respectively, in water. Amino acid substitution analysis also revealed that some of the amino acids must be conserved to obtain a folded Trp cage, but other positions could be changed without compromising folding characteristics (11, 12). The aim of this study was to construct a phage display library using the Trp cage as a self-folding protein scaffold with substitution of randomized amino acids in some of the nonconserved positions in order to create a library of novel peptides that can be used to select for target-specific binding ligands.

The stable fold adopted by peptides in the Trp cage library should result in tighter binding to target molecules by decreas-

ing the change in conformational entropy upon binding. Gomez and Freire (21) studied the energetics of the binding of pepstatin A to endo-thiapepsin and noted that the main obstacle to binding was the reduction in rotational, translational, and, most importantly, conformational entropy of the peptide. The conformational entropy of a peptide can be attributed to immobilization of the peptide backbone upon binding, burying of side chain upon binding, and a smaller term for solvent-accessible side chains. Since the Trp cage library is a stable folded structure, the peptide backbone is already fixed; thus, the change in conformational entropy is minimal.

Initially, we attempted to display the Trp cage on the gene III protein of bacteriophage M13, but those attempts failed to produce phage containing Trp cage inserts. Lack of success may have been due to

incompatibility of the highly folded Trp cage and the biological constraints of the M13 system, which relies on extrusion of the phage and any displayed proteins through the cell membrane (22). Also, the tail fiber protein pIII, which is fused to displayed peptides in the M13 system, is anchored to the inner membrane of the cell prior to phage assembly (23), and a displayed peptide with a hydrophobic domain, such as the Trp cage, might inhibit phage assembly by altering the interaction of pIII with the inner membrane. The M13 system that we used (Ph.D. system; New England Biolabs) relies on the pIII signal sequence for Sec-dependent protein translocation of displayed peptides (24). Since Sec translocates unfolded proteins, incorporation of signal sequences from the signal recognition pathway has recently been shown to improve translocation and display of folded proteins on filamentous phage (25). Display of the Trp cage library with the M13 system might be achieved by using the signal recognition pathway or an alternate protein translocation pathway.

The T7Select system was chosen next to display the library, because it utilizes a lytic phage maturation process and therefore was expected to avoid potential problems with membrane interactions and/or protein translocation. Although it is a lytic system, display of TC3b on T7 was not compatible with growth of the high valency vector T7Select415-1b, as evidenced by a high frequency of mutations within the insert encoding the Trp cage. The other T7Select vectors 10-3b (displaying 5–10 copies per phage) and 1-1b (displaying 0.1–1 copies per phage) both exhibited a low mutation frequency within the inserted sequence. The one-step growth experiments showed that display of Trp cage TC5b did not affect the number of infected cells producing phage but did affect the number of phage produced per infected cell greater than display of a linear peptide. This suggests that display of the Trp cage interferes in some way

with phage assembly. Western analysis showed that despite the affect of the Trp cage on phage production, display on the capsid protein of T7Select10-3b was achieved.

The design of the Trp cage library was based on version TC5b of the miniprotein with randomized amino acids at seven positions, resulting in display of the sequence AAADXYXQWLXXXGPXSGRPPX as a C-terminal protein fusion on the phage capsid protein (gp10B). The three alanine residues (AAA) were added to the N terminus to increase helical propensity (16). Secondary structure analysis of gp10B using PepTool software (BioTools, Edmonton, Canada) predicted the phage protein to be highly helical, which could also contribute to the folding of the Trp cage N-terminal helix. The N-capping residue Asn of TC5b was replaced with Asp (D). Randomized positions, indicated by an X, were selected because they are not conserved and are located on the solvent-exposed surfaces of the Trp cage NMR structure (Fig. 2), where they are available for ligand binding. As a result of the fold, the randomized positions are in close proximity.

Characterization of the naive library by DNA sequencing revealed that the library consisted of three groups designated nonmutants, stops, and mutants. Nonmutants comprised 46% of the clones, and they contained the Trp cage scaffold as designed with seven randomized amino acids. Stops (16%) contained the stop codon TAG in one of the randomized positions. In order to obtain a more uniform distribution of amino acid frequency and reduce the occurrence of stop codons, the oligonucleotide used to construct the Trp cage library was based on a 32-codon genetic code that only contains one stop codon, TAG. Systems that use the bacteriophage M13 for phage display utilize a host that can suppress the remaining stop codon. However, the bacterial host used for the T7 system does not contain a suppressor for the stop codon TAG, resulting in a truncated Trp cage when this codon appears in one of the variable positions. Among the random clones that were selected for sequencing, 18 contained the TAG stop codon in a randomized position. The stop codon was observed in all of the randomized positions except the last one. Since the frequency of TAG within the 32-codon genetic code is 0.031 and 109 clones with seven random positions were sequenced, the expected observations of TAG occurring would be 23.6. Since we only observed 18 occurrences, there does not appear to be a selective pressure increasing the frequency of stop codons at the randomized positions of the library. The truncated proteins would still be displayed but might not have a Trp cage structure. To reduce the impact of the stop codon, a suppressor host was constructed for use in any further amplifications of the library. Since this host contains the *supE44* suppressor mutation, glutamine residues would be inserted at TAG codons raising the frequency of that amino acid in the library. Mutants were observed at high frequency (38%) and contained primarily single base insertions or deletions. It is likely that the origin of observed mutants comes from the original chemical synthesis of the 103-base oligonucleotide template used to construct the library (26, 27). The frequency of mutation indicates that the Trp cage exerts selec-

tive pressure on the phage even with the display of only 5–10 copies.

The inability to display the Trp cage in the M13 system, the common occurrence of mutation resulting from display of TC3b in the T7Select system, and mutations observed within the T7 Trp cage library all suggest that the Trp cage is a highly active structure that is selected against based on the copy number and modifications resulting from randomized amino acids. Since the Trp cage affects phage growth, it is possible that some of the modified structures that are too active for compatibility with this system might be observable by ribosome display or an alternate *in vitro* display system. After observing a high mutation frequency from display of TC3b, the library was designed so that the stabilizing salt bridge between Asp-9 and Arg-16 of TC5b was sacrificed when Asp-9 was chosen as one of the randomized positions. We reasoned that decreasing the stability of the Trp cage fold by removing the salt bridge would reduce selective pressure against the Trp cage, providing the potential for discovery of a greater number of active molecules from the library by, in effect, reducing the stringency. Among the nonmutant clones that were sequenced from the library, we observed one occurrence of a random insertion that restored Asp-9 that was within the expected frequency range, indicating that the restoration of the salt bridge is neither selected for or against. The stability of any peptide sequences selected from the library by panning experiments could potentially be enhanced by reestablishing the salt bridge as long as restoring Asp-9 does not adversely affect the desired binding characteristics.

The Trp cage library was used successfully to identify an HPQ-containing peptide that binds to streptavidin. Such HPQ-containing peptides were previously identified using M13 phage libraries (5, 28). Crystallographic analysis has demonstrated that the HPQ sequence binds to streptavidin as a Type I β -turn (29). Selection of an HPQ sequence in the Trp cage miniprotein (Table 6) occurred at the three variable positions located between an α -helix and a 3_{10} helix (Fig. 2), a region expected to most easily accommodate a β -turn. Detailed structure-activity studies of the HPQ containing Trp cage miniprotein are in progress and will be reported elsewhere.

Despite the frequency of clones containing mutations and stops, the library still contains enough nonmutant clones to provide substantial complexity. One estimate of library complexity is the number of primary clones. Since we produced 1.6×10^9 primary clones for this library, the estimated complexity from nonmutants is 7.4×10^8 novel peptides. An additional measure of library complexity was provided by analysis of 50 nonmutant clones with the DIVAA program. The Trp cage library had an average diversity of 0.67, which is comparable with the average diversity of 0.70 for the constrained Ph.D.-C7CTM library. Linear and disulfide-constrained libraries constructed in T7 have been reported to have higher diversity than M13 libraries after evaluation by DIVAA (30). A linear library of 12 amino acids in T7 had an average diversity of 0.85 compared with the linear M13 library Ph.D.-12TM, which had an average of 0.68, and a T7 library with 7 amino acids constrained by a disulfide bridge had an average diversity of 0.80. The lower diversity of the Trp cage library is probably due to amino acid

combinations that are not tolerated, as indicated by the occurrence of mutations. The diversity of the Trp cage library was lowest at the third randomized position, which is located near the junction of the α -helix and the 3_{10} helix, suggesting that this region is conserved.

Although immunoglobulins have been relied upon for therapeutic and diagnostic applications where protein scaffolds with high binding affinity are needed, there has been recent interest in the development of scaffolds based on nonimmunoglobulin proteins (31). Characteristics considered ideal for nonimmunoglobulin binding proteins include a small single-chain protein, high thermodynamic stability, absence of disulfide bonds, and no free cysteines. Proteins with these characteristics would be easier to manufacture, more stable, and capable of folding in the intracellular environment, where many drug targets are located, and they would be amenable to adding a single cysteine for the coupling of effector compounds. The Trp cage miniprotein satisfies all of these characteristics, making it a good candidate nonimmunoglobulin scaffold.

An interesting example of a phage display peptide containing a disulfide bond that has some similarities to the Trp cage is the β -hairpin peptide (NLPRCTEGPWGWVCM), which binds to the human IgE receptor (6). Both are stably folded peptides of small size that have a central Gly-Pro motif structure and Trp residues that are involved in folding. In the case of the β -hairpin peptide, the GP motif is part of the sequence (GPWG) that forms the hairpin, whereas in the Trp cage, the GP motif is part of a 3_{10} helix (GGPSSGR). The hairpin structure relies upon the disulfide bond for conformational stability, whereas the Trp cage conformation is disulfide-independent. Unlike the Trp cage, there does not appear to be interaction between Trp and Pro in the β -hairpin peptide; instead, Trp-12 contributes to fold stability by packing up against the disulfide bond. The Trp cage miniprotein contains a Trp sandwiched between prolines, whereas binding of the β -hairpin peptide to the IgE receptor appears to be dependent upon formation of a "proline sandwich" between Pro-9 and two tryptophans in the IgE receptor. Some of the prolines of the Trp cage might also be available for sandwich formation, or peptides from the library containing proline in the randomized positions, such as the one selected for as a binder to human bronchial epithelial cells containing a proline at the first randomized position, could be involved in such formations.

This study demonstrated that a random amino acid library utilizing the Trp cage miniprotein TC5b as a scaffold can be displayed on bacteriophage T7. A library of high complexity and diversity was obtained and was shown to support the rapid identification of peptides containing the HPQ motif when panned against streptavidin. This library was useful for the selection of a peptide that binds to human bronchial epithelial cells and may therefore be valuable for the selection of other cell binding peptides to improve drug delivery technology and for the selection of receptor-binding ligands. Since the Trp cage has been studied as a model for protein folding, this library could also be useful for the selection of novel Trp cage sequences for folding studies.

Acknowledgments—We thank Niels Andersen (University of Washington) and Jianzhu Chen (Massachusetts Institute of Technology) for critical review of the manuscript; Martin Lee and Sonia Souza for expert assistance with statistical analysis; Mohammad Ahmadian for helpful discussions; the Teranode Corporation, Ken Cooper, Jason Bohreer, and Jeremy Thompson for assistance in the development of Teranode software applications for processing sequence data and database construction; and Harry Wang and Ken Farber for peptide synthesis and purification support.

REFERENCES

- Smith, G. P. (1985) *Science* **228**, 1315–1317
- Lowman, H. B. (1997) *Annu. Rev. Biophys. Biomol. Struct.* **26**, 401–424
- Ladner, R. C. (1995) *Trends Biotechnol.* **13**, 426–430
- Uchiyama, F., Tanaka, Y., Minari, Y., and Tokui, N. (2005) *J. Biosci. Bioeng.* **99**, 448–456
- Giebel, L. B., Cass, R. T., Milligan, D. L., Young, D. C., Arze, R., and Johnson, C. R. (1995) *Biochemistry* **34**, 15430–15435
- Nakamura, G. R., Starovasnik, M. A., Reynolds, M. E., and Lowman, H. B. (2001) *Biochemistry* **40**, 9828–9835
- Ladner, R. C., and Ley, A. C. (2001) *Curr. Opin. Biotechnol.* **12**, 406–410
- Rebar, E. J., and Pabo, C. O. (1994) *Science* **263**, 671–673
- Friesen, W. J., and Darby, M. K. (1998) *Nat. Struct. Biol.* **5**, 543–546
- Dennis, M. S., Herzka, A., and Lazarus, R. A. (1995) *J. Biol. Chem.* **270**, 25411–25417
- Neidigh, J. W., Fesinmeyer, R. M., Prickett, K. S., and Andersen, N. H. (2001) *Biochemistry* **40**, 13188–13200
- Andersen, N. H., Barua, B., Fesinmeyer, R. M., and Neidigh, J. (2001) in *Peptides: The Wave of the Future* (Lebl, M., and Houghton, R. A., eds) pp. 406–408, American Peptide Society, Springer, New York
- Neidigh, J. W., Fesinmeyer, R. M., and Andersen, N. H. (2002) *Nat. Struct. Biol.* **9**, 425–430
- Qiu, L., Pabst, S. A., Roitberg, A. E., and Hagen, S. J. (2002) *J. Am. Chem. Soc.* **124**, 12952–12953
- Edgar, R. S. (1963) in *Methodology in Basic Genetics* (Burdette, W. J., ed) pp. 19–36, Holden-Day, Inc., San Francisco
- Lin, J. C., Barua, B., and Andersen, N. H. (2004) *J. Am. Chem. Soc.* **126**, 13679–13684
- Clopper, C. J., and Pearson, E. S. (1934) *Biometrika* **26**, 404–413
- Macrina, F. L., Kopecko, D. J., Jones, K. R., Ayers, D. J., and McCowen, S. M. (1978) *Plasmid* **1**, 417–420
- Miller, J. H., and Albertini, A. M. (1983) *J. Mol. Biol.* **164**, 59–71
- Edelmann, P., Martin, R., and Gallant, (1987) *J. Mol. Gen. Genet.* **207**, 517–518
- Gomez, J., and Freire, E. (1995) *J. Mol. Biol.* **252**, 337–350
- Castagnoli, L., Zucconi, A., Quondam, M., Rossi, M., Vaccaro, P., Panni, S., Paoluzi, S., Santonico, E., Dente, L., and Cesareni, G. (2001) *Comb. Chem. High Throughput Screen.* **4**, 121–133
- Marvin, D. A. (1998) *Curr. Opin. Struct. Biol.* **8**, 150–158
- Rapoza, M. P., and Webster, R. E. (1993) *J. Bacteriol.* **175**, 1856–1859
- Steiner, D., Forrer, P., Stumpp, M. T., and Pluckthun, A. (2006) *Nat. Biotechnol.* **24**, 823–831
- Beaucage, S. L., and Iyer, R. P. (1992) *Tetrahedron* **48**, 2223–2311
- Krotz, A. H., Klopchin, P. G., Walker, K. L., Srivatsa, G. S., Cole, D. L., and Ravikumar, V. T. (1997) *Tetrahedron Lett.* **38**, 3875–3878
- Devlin, J. J., Panganiban, L. C., and Devlin, P. E. (1990) *Science* **249**, 404–406
- Katz, B. A. (1995) *Biochemistry* **34**, 15421–15429
- Krumpe, L. R., Atkinson, A. J., Smythers, G. W., Kandel, A., Schumacher, K. M., McMahon, J. B., Makowski, L., and Mori, T. (2006) *Proteomics* **6**, 4210–4222
- Binz, H. K., Amstutz, P., and Pluckthun, A. (2005) *Nat. Biotechnol.* **23**, 1257–1268
- New England Biolabs, Inc. (2002) *Instruction Manual for the Ph.D.-C7C Phage Display Peptide Library Kit*, p. 21, New England Biolabs, Inc., Ipswich, MA

The Trp Cage Motif as a Scaffold for the Display of a Randomized Peptide Library on Bacteriophage T7

Richard E. Herman, Douglas Badders, Mark Fuller, Ekaterina G. Makienko, Michael E. Houston, Jr., Steven C. Quay and Paul H. Johnson

J. Biol. Chem. 2007, 282:9813-9824.

doi: 10.1074/jbc.M610722200 originally published online January 30, 2007

Access the most updated version of this article at doi: [10.1074/jbc.M610722200](https://doi.org/10.1074/jbc.M610722200)

Alerts:

- [When this article is cited](#)
- [When a correction for this article is posted](#)

[Click here](#) to choose from all of JBC's e-mail alerts

This article cites 26 references, 6 of which can be accessed free at <http://www.jbc.org/content/282/13/9813.full.html#ref-list-1>