Published in final edited form as:

Protein Expr Purif. 2009 August ; 66(2): 172–180. doi:10.1016/j.pep.2009.03.002.

***E. coli* expression of a soluble, active single-chain antibody variable fragment containing a nuclear localization signal**

Hairong Xiong^{a,d}, Shuyi Li^{a,b}, Zhanqiu Yang^d, Richard R. Burgess^c, and William S. Dynan^{a,*}

^aInstitute of Molecular Medicine and Genetics, Medical College of Georgia, Augusta, GA 30912

^bApeliotus Technologies, 1456 N. Morningside Dr. NE, Atlanta, GA, 30306

^cMcArdle Laboratory for Cancer Research, University of Wisconsin, Madison, WI, 53706

^dState Key Laboratory of Virology, Institute of Medical Virology, Wuhan University School of Medicine, Wuhan 430071, China

Abstract

Single-chain antibody variable fragment (scFv) proteins consist of an antibody heavy chain variable sequence joined via a flexible linker to a light chain variable sequence. Prior work has shown that ScFv 18-2 binds the DNA-dependent protein kinase catalytic subunit (DNA-PKcs) and sensitizes cancer cells to radiation following nuclear microinjection. A potential clinical delivery strategy is based on modification of the scFv so that it can be taken up into cells and imported to the nucleus. This will require development of an expression system for a nuclear localization signal (NLS)-tagged scFv derivative. We found, however, that addition of the highly basic NLS severely compromised expression in the host-vector system used for the parental scFv. After testing a variety of host strains, fusion partners, and NLS sequences and placements, successful expression was obtained with a construct containing a stabilizing N-terminal maltose binding protein tag and a single, optimized, C-terminal NLS moiety. Amylose affinity-purified ScFv 18-2 NLS protein was stable to storage at 4 °C in the presence of glycerol or trehalose, bound selectively to an epitope peptide, and was cleavable at an engineered Factor Xa protease site. Following lipid-mediated uptake into cultured cells, NLS-tagged ScFv 18-2, unlike the parental ScFv 18-2, localized predominantly in the cell nucleus.

Keywords

scFv; nuclear localization signal; fusion tag; DNA-dependent protein kinase catalytic subunit

Introduction

A single-chain antibody variable fragment (scFv) is composed of the variable regions of the heavy and light chains of an immunoglobulin joined by a flexible linker [1,2]. Because of their

© 2009 Elsevier Inc. All rights reserved.

Corresponding author. Address: Institute of Molecular Medicine and Genetics, Medical College of Georgia, Room CB-2803, 1120 15th Street, Augusta GA 30912, wdynan@mcg.edu (W. Dynan), Phone: 706.721.8756, FAX: 706.434.6440.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Declaration of competing interests

H.X. and W.S.D. are consultants to Apeliotus Technologies, which owns intellectual property relating to the parental ScFv 18-2.

small size and antigen-binding properties, scFvs have a wide range of therapeutic and diagnostic applications. The present studies used ScFv 18-2, which binds the DNA-dependent protein kinase catalytic subunit (DNA-PKcs) and blocks DNA end joining in an *in vitro* assay [3]. Inhibition of DNA repair, particularly DNA double-strand break repair, affords a possible means to increase the sensitivity of cancer cells to radiotherapy [4–6]. The *in vivo* activity of ScFv 18-2 has been validated by nuclear microinjection studies, which show that it is capable of co-localizing with DNA-PKcs, sensitizing cells to an otherwise sublethal dose of ionizing radiation and extending the persistence of γ -H2AX foci, a marker of unrepaired DNA double-strand breaks [3].

In this study we explored ways to express and purify an improved version of ScFv 18-2 that is better able to reach its target inside the cell nucleus. Several reports describe expression of scFvs in mammalian cells via gene transfer (for example, [7–10]). This approach has been termed “intracellular immunization” [11]. Although proteins of less than 40 kDa can sometimes enter the nucleus by passive diffusion through nuclear pores, at least some scFvs require a nuclear localization signal (NLS) to promote entry [12]. The NLS engages receptors that direct proteins to the nucleus via the nuclear pore complex [13]. Other localization sequences have been shown to direct intracellularly expressed scFvs to the endoplasmic reticulum, mitochondria, or secretory apparatus [11].

Expression via gene transfer is limited to those scFvs that fold properly inside mammalian cells. Most scFvs require formation of intrachain disulfide bonds for folding, which is hindered by the intracellular reducing environment [12]. Perhaps for this reason, our preliminary attempts to express ScFv 18-2 inside mammalian cells by gene transfer were not fruitful (S.L. and W.S.D., unpublished observations). We have therefore turned to a different approach based on expression of an NLS-tagged ScFv 18-2 in conventional *Escherichia coli* host-vector systems and transfer of the expressed protein into mammalian cells. Expression in *E. coli* is the quickest and most consistent of various systems that have been tested, although scFv proteins have also been expressed in the yeasts, *Saccharomyces cerevisiae* and *Pichia pastoris*, with varying degrees of success [14,15].

We have investigated a number of strategies for high-level *E. coli* expression of ScFv 18-2 containing an NLS sequence. A classical, or monopartite, NLS is composed of three to five basic amino acid residues, whereas a bipartite NLS is composed of two basic regions separated by a spacer of variable length [13,16,17]. We modified ScFv 18-2 by adding a monopartite nuclear localization signal derived from SV40 T antigen [18] but were unable to obtain satisfactory activity or protein yield in the host-vector system that was used for the parental scFv. We explored a number of different strategies, including refolding from inclusion bodies, insertion of NLS elements in different positions and copy numbers, and addition of different fusion partners. The best results, among the strategies tested, were obtained using an N-terminal maltose binding protein (MBP) tag and a single C-terminal NLS, joined to the body of the scFv by a short flexible linker. We demonstrate that this derivative, MBP-ScFv 18-2 NLS LC1, but not the parental ScFv 18-2, is capable of undergoing nuclear import when introduced into mammalian cells.

Materials and Methods

Plasmid vectors

The starting plasmid was pCANTAB 5E ScFv 18-2 [3], which targets expression to the periplasmic space using a phage g3 signal sequence. A C-terminal E tag allows immunodetection and immunoaffinity purification. To create N-terminal NLS derivatives, annealed NLS SfiI-F and NLS SfiI-R oligonucleotides (Supplementary Table 1) were inserted in-frame at a unique SfiI site upstream of the scFv coding sequence. The ligation reaction

mixture was transformed into the *E. coli* TOP10 strain. Colony PCR was performed using one primer internal to the insert and one primer flanking the insert. Positive colonies were identified and sequenced. Clones containing one, two, or three copies of the insert were designated NLS pCANTAB 5E ScFv 18-2 NLS N1, N2, and N3, respectively. To create analogous C-terminal NLS derivatives, annealed oligonucleotides NLS1 NotI-F and NLS1 NotI-R (Supplementary Table 1) were inserted at a unique NotI site downstream of the scFv coding sequence. Clones were identified by PCR and sequencing as before, and clones containing one or two copies of the NLS were designated pCANTAB 5E ScFv 18-2 NLS C1 and C2, respectively. Alternatively, annealed NLS3 NotI-F and NLS3 NotI-R oligonucleotides (encoding a triple tandem NLS, see Supplementary Table 1) were inserted at the NotI site. A clone bearing one copy of the triple NLS sequence was designated pCANTAB 5E ScFv 18-2 NLS C3. To create a derivative with an NLS flanking the scFv coding sequence on both ends, the annealed NLS1 NotI-F and NLS1 NotI-R oligonucleotides were inserted at the NotI site of pCANTAB 5E NLS N1. The desired clone was designated pCANTAB 5E ScFv 18-2 NLS N1C1.

To create clones with a slightly longer linker between the scFv and the NLS sequence (corresponding to the 4-residue sequence that separates the scFv coding region and E tag in the parent clone), annealed NLS L1 NotI-F and NLS L1 NotI-R oligonucleotides were inserted at the NotI site of pCANTAB 5E ScFv 18-2. Clones containing one and two copies of the NLS were identified and designated as pCANTAB 5E ScFv 18-2 NLS LC1 and LC2, respectively. A variant was also constructed with a slightly longer spacer between the scFv coding region and the NLS. Annealed NLS GlyC1 NotI-F and NLS GlyC1 NotI-R oligonucleotides (Supplementary Table 1) were inserted at the NotI site of pCANTAB 5E ScFv 18-2, and a clone containing a single copy of the insert was designated pCANTAB 5E ScFv 18-2 NLS GlyC1.

Creation and testing of glutathione-S-transferase (GST)-ScFv 18-2 derivatives is described in Supplementary Data. MBP-ScFv 18-2 derivatives were constructed using pMAL-p4X (New England Biolabs, Ipswich, MA). The MBP in this commercial vector has been engineered to increase its affinity for amylose, facilitating affinity purification. The vector also contains an engineered protease site between the MBP and scFv coding sequences, facilitating cleavage of the MBP tag. To express ScFv 18-2 in this system, pCANTAB 5E ScFv 18-2 NLS LC1 and LC2 were used as PCR template with N-EcoRI and C-HindIII primers (Supplementary Table 1). The resulting clones were designated pMAL-p4X ScFv 18-2 NLS LC1 and LC2, respectively.

Protein expression and purification

The pCANTAB 5E series clones were transferred to *E. coli* strains HB2151 or TG1 for protein expression. Each strain was cultured in 2-YT medium with 2% glucose (2-YT AG) at 30 °C to an OD₆₀₀ of ~ 0.8 to 1.0. (Note that the glucose in this medium maintains catabolite repression of scFv expression). Cultures were centrifuged and resuspended in the same volume of 2-YT without glucose but containing 1 mM isopropylthiogalactoside (IPTG). This medium (designated 2-YT AI) induces scFv expression. Cultures were incubated overnight at 30 °C. Periplasmic proteins were extracted as described [3]. Soluble cytoplasmic proteins were extracted from the residual pellet using 5 ml/g wet weight pellet of Bugbuster Protein Extraction Reagent (Novagen, Gibbstown, NJ). Insoluble proteins were extracted by heating in SDS-PAGE sample buffer. Periplasmic, soluble, and insoluble fractions were analyzed by SDS-PAGE with immunoblotting with rabbit polyclonal anti-E tag (1:500 or 1:1000, Abcam, Cambridge, MA).

The pMAL series clones were transferred into *E. coli* strain K12TB1. Each strain was cultured at 30 °C in 2-YT medium containing 1% glucose and 100 µg/ml ampicillin until it reached an OD₆₀₀ of about 0.5. IPTG was added to 0.33 mM, and incubation was continued at 18 °C for

18–20 h. Cells were harvested by centrifugation at 4000 *g* for 20 min and the pellet (typically 10 g wet weight /l culture) was resuspended in 30 mM Tris-HCl (pH 8.0 at 20 °C), 20% sucrose, (400 ml/l culture). EDTA was added to 1 mM and the preparation was incubated for 5–10 minutes at room temperature with shaking. The supernatant was centrifuged at 8000 *g* for 20 minutes at 4 °C, the pellet was resuspended in 400 ml/l culture of ice-cold osmotic shock buffer (5 mM MgSO₄), and the suspension was shaken for 10 min in an ice bath. The material was centrifuged at 8000 *g* for 20 min at 4 °C. Tris-HCl (pH 7.4 at 20 °C) was added to 8 mM, and the crude extract was loaded onto amylose resin (6 ml/l culture) at a flow rate of about 0.1 column volumes/minute. The column was washed with 8–12 volumes of 20 mM Tris-HCl (pH 7.4 at 20 °C), 200 mM NaCl, 1 mM EDTA and was eluted with the same buffer containing 10 mM maltose. Fractions were collected, dialyzed against PBS (pH 7.8), pooled, and concentrated to 1–2 mg/ml using a 10 kDa molecular weight cutoff Vivaspin 6 centrifugal concentrator (Sartorius Stedim Biotech, Goettingen, Germany). The yield was ~2 mg/l culture. The concentrated MBP-ScFv 18-2 NLS protein was stored at –20 °C in PBS containing 20% glycerol.

Peptide ELISA

Peptide A (DNA-PKcs residues 2001–2025, biotin-KKKYIEIRKEAREAANGSDSGPSYM) or Peptide C (DNA-PKcs residues 2031–2055, biotin-LADSTLSEEMSQFDFSTGVQSYSYS) were incubated in individual wells of a Reacti-Bind NeutrAvidin Coated Strip Plate (Thermo Fisher Scientific, Rockford, IL) at room temperature for 1 h (100 µl of 5 µg/ml solution/well). Plates were washed with PBS, blocked with 3% BSA in PBS, and incubated 1 h at room temperature with 100 µl primary antibody (mouse anti-E-tag monoclonal antibody (mAb), GE Healthcare Bio-Sciences Corp., Piscataway, NJ) 1:1000 in PBST (PBS with 0.5% Tween-20) with 1% BSA and 1% gamma-globulin. They were washed with PBST and incubated 1 h at room temperature with 100 µl secondary antibody (alkaline phosphatase-conjugated goat-anti-mouse IgG (1:5000 in PBST with 1% BSA and 1% gamma-globulin, Sigma-Aldrich, St. Louis MO). After washing with PBST, color development was performed using the BluePhos Microwell Phosphatase substrate kit (Kirkegaard & Perry Laboratories, Gaithersburg, MD). Absorbance was determined at 595 nm.

Cellular uptake

ScFvs were delivered to HeLa cells using a BioPORTER QuikEase Protein Delivery kit (Sigma-Aldrich, St. Louis, MO). HeLa cells were grown in DMEM with 10% fetal bovine serum and subcultured 2–3 times weekly. Cells were seeded in 8-well glass slides (Lab-Tek Chamber Slide, Nunc, Rochester, NY) and grown to 70–80% confluency. ScFvs were diluted to 250 µg/ml in a final volume of 40 µl PBS and added to dry BioPORTER Reagent. After allowing 3–5 min for rehydration, 460 µl of serum-free DMEM was added. Cells were washed once with serum-free DMEM, 80 µl serum-free DMEM was added, and 80 µl of scFv/BioPORTER mixture was transferred onto the cells. The cells were incubated at 37 °C for 4 h, fixed, and stained for immunofluorescence as described [3]. Antibodies used were: rabbit anti-E tag (1:200, Abcam, Cambridge, MA), mouse mAb anti-DNA-PKcs (Ab-2 (Clone 25-4), 1:100, Thermo Scientific, Fremont, CA), Alexa Fluor 594 goat-anti-rabbit IgG, and Alexa Fluor 488 goat-anti-mouse IgG (Molecular Probes, Inc., Eugene, OR). The slides were mounted using VECTASHIELD mounting medium with DAPI (Vector Laboratories, Burlingame, CA) and confocal projection images were collected using a Zeiss LSM 510 Meta Confocal Microscope.

Results

Design of NLS-containing ScFv 18-2 derivatives

Recombinant scFvs used in this study are derivatives of ScFv 18-2, which recognizes an epitope in a central regulatory region of DNA-PKcs (amino acids 2001–2025) [3]. The parent ScFv 18-2 was expressed in pCANTAB 5E, which contains a bacteriophage g3 signal peptide that targets expression to the *E. coli* periplasm. The coding sequence contains heavy and light chain variable sequences from mAb 18-2 [19], joined by a (G₄S)₃ linker. A C-terminal E-tag permits detection and immunopurification using commercial anti-E tag antibody.

To create NLS-containing derivatives, we inserted oligonucleotides encoding 1–3 copies of the SV40 NLS into the parent pCANTAB 5E vector, either at an SfiI site at the N terminus of the scFv coding sequence or at a NotI site between the coding sequence and the E tag (Fig. 1A). Depending on the position of insertion and whether additional linker sequences were present, these constructs were termed the ScFv 18-2 NLS N, C, LC, or Gly C series (see Supplementary Data for details).

Constructs were also made in the pMAL-p4X vector, which directs periplasmic expression of MBP-tagged derivatives (Fig 1A(c)). A complete map of the vector encoding one of the MBP constructs, which had the best overall solubility and activity of the proteins that were evaluated, is shown in Fig. 1B.

Comparison of ScFv 18-2 and NLS C3

Initial studies used ScFv 18-2 NLS C3, which contains three tandem NLS sequences to promote robust nuclear import (Fig. 1A). Expression was attempted in parallel with parental ScFv 18-2 in the same *E. coli* HB 2151 strain as previously described [3]. Periplasmic fractions were analyzed with or without IPTG induction. Periplasmic extracts of the induced parental strain contained ScFv 18-2, with substantial cleavage at a site near the C terminus of the V_H domain, consistent with prior results [3] (Fig. 2A). By contrast, periplasmic extracts prepared under the same conditions from the ScFv 18-2 NLS C3 strain contained little or no immunoreactive scFv. Similar results were obtained in the *E. coli* RosettaBlue strain, which expresses rare tRNAs to aid translation of mammalian cDNAs. To determine if the ScFv 18-2 NLS C3 might be expressed but aberrantly targeted, we analyzed total cell extract and culture medium from induced cultures (Fig. 2B). ScFv 18-2 NLS C3 was present in the total extract, but was not detectable in the medium. Expression in the HB2151 and RosettaBlue strains was comparable, although there was less proteolysis in the RosettaBlue strain. To compare epitope-binding activity of ScFv 18-2 with and without the NLS, we performed a peptide ELISA as described in Materials and Methods. The ScFv 18-2 NLS C3 was far less active than the parent ScFv 18-2, even though 100-fold more scFv was present in the assay (Fig. 2C). Results suggest that modification of the parent scFv by addition of the 3-copy NLS leads both to aberrant targeting and severely diminished epitope-binding activity.

Recovery and refolding of ScFv 18-2 NLS C3 from inclusion bodies

We tried an alternative approach based on intracellular expression in inclusion bodies, followed by *in vitro* refolding. Details of these experiments are given in Supplementary Data. Although protein was recovered in good yield using an optimized refolding buffer, the product was not well behaved: it showed high nonspecific binding activity, interacted anomalously with Ni-NTA agarose, and lost epitope-binding activity when concentrated. The results suggested that ScFv 18-2 NLS C3 might be unstable under native conditions and that other NLS insertion positions or copy numbers should be explored.

Comparison of ScFv 18-2 derivatives with different NLS placements and copy numbers

A series of NLS-containing derivatives were expressed in pCANTAB5E and analyzed for expression in inclusion bodies and periplasmic extracts. Expression of 5 of the 7 constructs was detected in the inclusion bodies (Fig. 3A, lanes N1, N2, N3, N1C1, C3) and expression of all but one of these (N1C1) was also detected in periplasmic extract. Peptide ELISA showed that the periplasmic extracts had no epitope-binding activity, however, except for NLS C3, which had slight activity consistent with Fig. 2 (Fig. 3B). Results suggested that the new constructs were misfolded, unstable in *E. coli*, or both.

In some cases, misfolding of a tagged protein can be caused by insufficient length of the linker between a structured domain and a tag. We therefore varied the length of the linker between the VL domain and the NLS in the C1 and C2 derivatives. Expression was slightly improved relative to the corresponding C1 and C2 constructs (Fig. 3C). There was also detectable epitope-specific binding activity (Fig. 3D). Yields, however, were insufficient for further characterization.

Evaluation of ScFv 18-2 derivatives fused with GST or MBP

We tested two fusion partners, GST and MBP, which have been reported to promote recovery of recombinant proteins, including scFvs, in soluble form [20–22]. We found that his₆-GST-ScFv 18-2 NLS C3 was more soluble than the corresponding non-GST-tagged protein. However, the soluble GST fusion protein flowed through a glutathione-agarose affinity column and was refractory to cleavage of the GST moiety at an engineered HRV 3C protease site. Another GST fusion construct without the his₆ tag showed no epitope-specific binding activity (Supplementary Data). Because the GST fusion proteins did not show promise, they were not characterized further.

To create MBP fusion proteins, sequences encoding ScFv 18-2 NLS LC1 and LC2 derivatives (which appeared to show promise in Fig 3) were transferred to pMAL-p4X and expressed in the *E. coli* host K12TB1. Strains showed IPTG-inducible expression of a 75 kDa species that was immunoreactive with both anti-MBP and anti-E tag antibodies (Fig. 4). Although expression was seen in inclusion bodies, ample soluble product was also present. Peptide ELISA showed that the NLS LC1 derivative, but not the NLS LC2 derivative, had epitope-specific binding activity (Fig. 4). Of the many NLS-containing constructs evaluated, this appeared to be the most promising and was therefore selected for further characterization. For brevity, we shall henceforth refer to the MBP-ScFv 18-2 NLS LC1 derivative as “ScFv 18-2 NLS.”

Purification, stability, and cleavage of ScFv 18-2 NLS

Periplasmic extract containing soluble MBP fusion product was subjected to amylose affinity chromatography. Approximately half of the 75 kDa fusion protein bound to the column. (Fig. 5A). The eluted protein was immunoreactive with anti-MBP and anti-E tag antibodies, indicating that it was intact at the N and C termini. Peptide ELISA showed that periplasmic fraction (onput), and eluted protein had epitope-specific binding activity, although some activity was also lost in the column flow-through (Fig. 5B). Amylose affinity-purified proteins were subjected to further anti E-tag affinity chromatography. This resulted in purification to near homogeneity, with removal of an anti-MBP reactive protein fragment (Fig. 5C). Although the eluted protein was active (Fig. 5D) the yield was low. Single-step purification of periplasmic extract on the anti-E tag column also resulted in poor yield (data not shown), so this method of purification was not routinely used. Serial dilution of the amylose affinity-purified ScFv 18-2 NLS showed that epitope-specific binding activity was nearly as high as parent ScFv 18-2 on a molar basis (Fig. 5E). The mean protein yields obtained in five sequential purifications are given in Table 1.

To determine optimal storage conditions, purified ScFv 18-2 NLS was incubated for one week at 4 °C in PBS, which is the standard storage buffer for the parent ScFv 18-2 [3]. This was compared with storage at various temperatures using PBS supplemented with the biocompatible polyol stabilizers, glycerol or trehalose. Although the scFv lost all activity in PBS, it retained activity in 20–50% glycerol or 15% trehalose at all temperatures tested (Fig. 5F).

To further investigate whether ScFv 18-2 NLS was expressed in a native state, it was incubated with Factor Xa, which cleaves at an engineered site between the MBP and V_H domains. Loss of the 75 kDa fusion protein was observed, with appearance of a smaller protein species that was reactive with anti-E tag antibody (Fig. 5G). The position corresponds to the expected migration of cleaved scFv with a short residual N-terminal linker. Although half to two thirds of the MBP-scFv substrate was cleaved under the conditions used, little or no diminution of binding activity was evident in a peptide ELISA comparing input scFv to the Factor Xa cleavage mixtures (Fig. 5H). The cleaved ScFv 18-2 NLS also maintained its specificity for the target peptide, compared to the control peptide. The results suggest that the cleavage product remains active. Because the MBP tag does not appear to interfere with activity, however, we have not routinely attempted to remove it. Indeed, it has been reported that MBP has molecular chaperone properties that may stabilize scFv folding and increase expression [22].

Cellular uptake

To test if ScFv 18-2 NLS is capable, in principle, of entry into the cell nucleus, we performed a cellular uptake experiment using HeLa cells and BioPORTER Reagent, a lipid-based formulation that allows intracellular delivery of proteins and peptides into a range of cell types [23]. Cells were fixed and stained with DAPI to indicate the boundaries of the nucleus, with anti-E tag to indicate the localization of the scFv, and with anti-DNA-PKcs. Results show that the ScFv 18-2 NLS is taken up by the majority of cells, and that much, if not all, staining is within the boundaries of the nucleus (Fig. 6A). In many cases (inset) nuclei showed one or more intense foci of co-localized scFv and DNA-PKcs. Some cells also showed punctate staining of the cytoplasm, probably reflecting vesicular trapping. A parallel experiment (Fig. 6B) using ScFv 18-2 (without NLS) showed that HeLa cells also took up this protein in the presence of BioPORTER, although staining was, on average, less intense and most of the protein was in cytoplasmic vesicles. There was also some speckled staining within the nuclear boundaries, but no evident co-localization with DNA-PKcs foci. Because of low signal intensity, a confocal projection image was used, and we cannot exclude the possibility that these speckles are above or below the nucleus. In contrast to our prior microinjection study, we did not observe cells in which the DNA-PKcs had been drawn from the nucleus to the cytoplasm [3]. This may reflect the shorter time scale of the experiment (4 h versus 6 h) or an effect of the BioPORTER reagent.

Discussion

We describe here a systematic investigation of strategies for high-level *E. coli* expression of an scFv containing an NLS. Parental ScFv 18-2 is expressed at a modest level in the *E. coli* periplasm (1–2 mg per liter of culture) with variable proteolytic cleavage during expression [3]. Addition of NLS sequences reduced periplasmic expression to nearly undetectable levels. Extensive re-engineering was needed to obtain a derivative that was stable, capable of epitope-specific binding, and available in sufficient quantity for biological studies. In addition to MBP, glycerol or trehalose was required to stabilize the ScFv 18-2 NLS for freeze-thawing and storage, whereas these were not required for the parental scFv. Most immunoglobulins are naturally slightly acidic, and the pI of parental ScFv 18-2 is about 6.0. The lysine, arginine-

rich NLS shifts the pI to about 9.0, and we speculate that this may be an underlying factor that contributes to instability of the protein and difficulty in expression.

Consistent with this explanation, reducing the number of NLS moieties from three to one and fusion to MBP correlated with better stability, specificity, and expression. Addition of a linker sequence between the structured scFv domains and the NLS was also beneficial. By contrast, several strategies that reportedly increase yield or activity of other *E. coli*-expressed scFvs, including overexpression of rare tRNAs and mutation of thioredoxin reductase and glutathione reductase genes, had no apparent benefit. Refolding strategies were also not fruitful, although they have been widely used for purification of other recombinant proteins from *E. coli*, including scFvs [24]. Optimization of refolding conditions resulted in some initial recovery of active ScFv 18-2 NLS C3, but the protein showed high nonspecific binding in a peptide ELISA, the his₆-tag was inaccessible for immobilized metal affinity chromatography, and the protein was not readily concentrated. Likewise, addition of a his₆-GST tag or GST tag to the N terminus increased the yield of soluble cytoplasmic scFv, but the soluble protein was not active. Taken together, results demonstrate an inability to predict which approaches will be beneficial based on the prior literature or soluble expression levels alone and emphasize the need for empirical characterization of both expression levels and binding selectivity of many different variants.

Cellular uptake experiments provide evidence that the final ScFv 18-2 NLS derivative can, in principle, be taken up into the cell nucleus. Uptake was observed only in the presence of the BioPORTER reagent. The BioPORTER reagent is useful only to promote nonspecific uptake in cell culture studies and does not offer a comprehensive solution for *in vivo* delivery. An alternative method is to use peptide transduction domains such as occur in HIV Tat and other proteins [25]. Like the BioPORTER reagent, however, this mechanism is also nonspecific. Protein transduction domains are also highly basic, like the NLS, and may be expected to similarly aggravate problems of protein instability and expression. For these reasons we are currently exploring alternative approaches for cellular uptake based on receptor-mediated endocytosis [26].

Although there have been a number of reports of ectopic expression of scFvs by gene transfer, this approach is limited by the well-recognized problem of obtaining proper folding of immunoglobulin domains in the intracellular reducing environment. Strategies have been devised to select scFvs from a eukaryotic expression library based on intracellular epitope recognition [8,27], but these are not applicable to scFvs derived from well-characterized preexisting mAbs. Furthermore, like all gene therapy approaches, ectopic expression of intracellular antibodies is limited by a lack of safe and efficient vectors for clinical applications. A more general approach may be to produce scFv proteins using conventional host-vector systems and to confer on them drug-like functionalities that would allow them to home to the target tissue, cross the plasma membrane, and enter specific subcellular compartments. Determination of requirements for *E. coli* expression of an NLS-tagged scFv contributes toward this ultimate objective.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Dr. Yoshihiko Takeda for materials and advice during the initial phase of this study, John Edwards (Apeliotus Technologies, Atlanta, GA) for advice and encouragement, Dr. Quangsheng Du and his laboratory for advice and for the gift of pGEX-4T1 vector, and Sir Richard Roberts (New England Biolabs, Ipswich, MA) for suggesting the use of the MBP expression system. This work was supported by US Public Health Service Award CA 98239, the NIH

Nanomedicine Roadmap Initiative (EY018244), a Georgia Research Alliance collaboration planning grant, and Apeliotus Technologies. WSD is an Eminent Scholar of the Georgia Research Alliance. This article is in partial fulfillment of the requirements for a Ph.D. degree at Wuhan University (H. Xiong).

References

1. Bird RE, Hardman KD, Jacobson JW, Johnson S, Kaufman BM, Lee SM, Lee T, Pope SH, Riordan GS, Whitlow M. Single-chain antigen-binding proteins. *Science* 1988;242:423–426. [PubMed: 3140379]
2. Skerra A, Pluckthun A. Assembly of a functional immunoglobulin Fv fragment in *Escherichia coli*. *Science* 1988;240:1038–1041. [PubMed: 3285470]
3. Li S, Takeda Y, Wragg S, Barrett J, Phillips A, Dynan WS. Modification of the ionizing radiation response in living cells by an scFv against the DNA-dependent protein kinase. *Nucleic Acids Res* 2003;31:5848–5857. [PubMed: 14530433]
4. Madhusudan S, Middleton MR. The emerging role of DNA repair proteins as predictive, prognostic and therapeutic targets in cancer. *Cancer Treat. Rev* 2005;31:603–617. [PubMed: 16298073]
5. Martin NM. DNA repair inhibition and cancer therapy. *J. Photochem. Photobiol. B* 2001;63:162–170. [PubMed: 11684463]
6. Zaremba T, Curtin NJ. PARP inhibitor development for systemic cancer targeting. *Anticancer Agents Med. Chem* 2007;7:515–523. [PubMed: 17896912]
7. Maciejewski JP, Weichold FF, Young NS, Cara A, Zella D, Reitz MS Jr, Gallo RC. Intracellular expression of antibody fragments directed against HIV reverse transcriptase prevents HIV infection in vitro. *Nat. Med* 1995;1:667–673. [PubMed: 7585149]
8. Vascotto F, Visintin M, Cattaneo A, Burrone OR. Design and selection of an intrabody library produced de-novo for the non-structural protein NSP5 of rotavirus. *J. Immunol. Methods* 2005;301:31–40. [PubMed: 15907924]
9. Mary MN, Venot C, Caron de Fromental C, Debussche L, Conseiller E, Cochet O, Gruel N, Teillaud JL, Schweighoffer F, Tocque B, Bracco L. A tumor specific single chain antibody dependent gene expression system. *Oncogene* 1999;18:559–564. [PubMed: 9927213]
10. Filesi I, Cardinale A, van der Sar S, Cowell IG, Singh PB, Biocca S. Loss of heterochromatin protein 1 (HP1) chromodomain function in mammalian cells by intracellular antibodies causes cell death. *J. Cell Sci* 2002;115:1803–1813. [PubMed: 11956312]
11. Persic L, Righi M, Roberts A, Hoogenboom HR, Cattaneo A, Bradbury A. Targeting vectors for intracellular immunisation. *Gene* 1997;187:1–8. [PubMed: 9073060]
12. Sibling AP, Nordhammer A, Masson M, Martineau P, Trave G, Weiss E. Nucleocytoplasmic shuttling of antigen in mammalian cells conferred by a soluble versus insoluble single-chain antibody fragment equipped with import/export signals. *Exp. Cell Res* 2003;286:276–287. [PubMed: 12749856]
13. Pouton CW, Wagstaff KM, Roth DM, Moseley GW, Jans DA. Targeted delivery to the nucleus. *Adv. Drug Deliv. Rev* 2007;59:698–717. [PubMed: 17681634]
14. Miller KD, Weaver-Feldhaus J, Gray SA, Siegel RW, Feldhaus MJ. Production, purification, and characterization of human scFv antibodies expressed in *Saccharomyces cerevisiae*, *Pichia pastoris*, and *Escherichia coli*. *Protein Expr. Purif* 2005;42:255–267. [PubMed: 15946857]
15. Feldhaus MJ, Siegel RW. Yeast display of antibody fragments: a discovery and characterization platform. *J. Immunol. Methods* 2004;290:69–80. [PubMed: 15261572]
16. Christie RJ, Grainger DW. Design strategies to improve soluble macromolecular delivery constructs. *Adv. Drug Deliv. Rev* 2003;55:421–437. [PubMed: 12628325]
17. Hodel MR, Corbett AH, Hodel AE. Dissection of a nuclear localization signal. *J. Biol. Chem* 2001;276:1317–1325. [PubMed: 11038364]
18. Kalderon D, Richardson WD, Markham AF, Smith AE. Sequence requirements for nuclear location of simian virus 40 large-T antigen. *Nature* 1984;311:33–38. [PubMed: 6088992]
19. Carter T, Vancurova I, Sun I, Lou W, DeLeon S. A DNA-activated protein kinase from HeLa cell nuclei. *Mol. Cell Biol* 1990;10:6460–6471. [PubMed: 2247066]
20. Frangioni JV, Neel BG. Solubilization and purification of enzymatically active glutathione S-transferase (pGEX) fusion proteins. *Anal. Biochem* 1993;210:179–187. [PubMed: 8489015]

21. Chi XS, Landt Y, Crimmins DL, Dieckgraefe BK, Ladenson JH. Isolation and characterization of rabbit single chain antibodies to human Reg Ia protein. *J. Immunol. Methods* 2002;266:197–207. [PubMed: 12133637]
22. Shaki-Loewenstein S, Zfania R, Hyland S, Wels WS, Benhar I. A universal strategy for stable intracellular antibodies. *J. Immunol. Methods* 2005;303:19–39. [PubMed: 16045924]
23. Zelphati O, Wang Y, Kitada S, Reed JC, Felgner PL, Corbeil J. Intracellular delivery of proteins with a new lipid-mediated delivery system. *J. Biol. Chem* 2001;276:35103–35110. [PubMed: 11447231]
24. Lamberski JA, Thompson NE, Burgess RR. Expression and purification of a single-chain variable fragment antibody derived from a polyol-responsive monoclonal antibody. *Protein Expr. Purif* 2006;47:82–92. [PubMed: 16216525]
25. Wadia JS, Dowdy SF. Protein transduction technology. *Curr. Opin. Biotechnol* 2002;13:52–56. [PubMed: 11849958]
26. Zhao X, Li H, Lee RJ. Targeted drug delivery via folate receptors. *Expert Opin. Drug Deliv* 2008;5:309–319. [PubMed: 18318652]
27. Visintin M, Settanni G, Maritan A, Graziosi S, Marks JD, Cattaneo A. The intracellular antibody capture technology (IAC): towards a consensus sequence for intracellular antibodies. *J. Mol. Biol* 2002;317:73–83. [PubMed: 11916379]

A. Expressed protein sequences

a. Parent ScFv 18-2

b. ScFv 18-2 NLS N and C series; LC series; GlyC

c. MBP ScFv 18-2 NLS LC series

B. pMAL-p4X ScFv 18-2 NLS LC1 vector

Fig. 1.

Single-chain antibody variants characterized in this study. A. Domain structure of expressed proteins. (a) Parent ScFv 18-2 expressed using pCANTAB 5 E vector system [3]. The g3 signal peptide directs product to the periplasmic compartment and is cleaved during synthesis. Heavy chain (V_H), gly-ser linker, and light chain (V_L) sequences are shown. E tag allows recognition by specific antibody for detection and immuno-affinity purification. (b) ScFv 18-2 NLS N and C series; LC series; GlyC series. (c) MBP ScFv 18-2 NLS LC series. SV40 nuclear localization signal has sequence DPKKKRKV. For DNA sequence details, see Supplementary Table 1. B. Structure of pMAL-p4X ScFv 18-2 NLS LC1 vector. Vector portion is adapted from manufacturer's technical data sheet.

Fig. 2.

Comparison of pCANTAB 5E ScFv 18-2 and ScFv 18-2 NLS C3. A. Periplasmic extract was prepared as described in Materials and Methods using indicated vectors and *E. coli* host strains. Molecular weights are indicated in kDa. IPTG was added to 1 mM where indicated. Proteins were resolved by SDS-PAGE and visualized by immunoblotting using anti-E tag antibody as described in Materials and Methods. Each lane contains an equivalent amount of extract (2.6% of periplasmic extract from a 50 ml culture). Arrow denotes full-length ScFv 18-2. Asterisk denotes a cleavage product produced *in situ* during expression [3]. B. As in Panel A, except that culture medium and total cell extract were analyzed. Each lane contained an equivalent amount of medium (0.26% of 10 ml culture) or total cell lysate (1.2% of 10 ml culture). C. Peptide ELISA. Biotinylated peptides representing ScFv 18-2 epitope (DNA-PKcs amino acids 2001–2025) or adjacent control sequence (amino acids 2031 to 2055) were bound to streptavidin-coated microwells and ScFv 18-2 epitope-binding activity in periplasmic extract

was determined by ELISA as described in Materials and Methods. Wells contained 20 μg of total periplasmic extract from ScFv 18-2 NLS C3-expressing cells or 0.2 μg from ScFv 18-2 expressing cells as indicated. Wells with no scFv served as negative controls. Assays were performed in duplicate and range is shown.

Fig. 3.

Characterization of pCANTAB 5E ScFv 18-2 variants containing different NLS sequences and placements. **A.** Expression of ScFv 18-2 NLS variants in inclusion bodies or periplasmic extract. The indicated proteins were resolved and visualized by immunoblotting as in Fig. 2. Each lane contained an approximately equivalent amount of protein solubilized from inclusion bodies or periplasmic extract (0.5% of 50 ml culture). **B.** Peptide ELISA performed as in Fig. 2 using 20 μ g of total periplasmic extract from ScFv 18-2 NLS-expressing cells or 0.2 μ g from ScFv 18-2-expressing cells as indicated. **C.** Expression of ScFv 18-2 variants containing additional linker sequences between the VL and NLS sequences. Experiment was performed as in Panel A. **D.** Peptide ELISA using periplasmic extracts of variants containing additional linker sequences.

Fig. 4.

A. Expression of MBP-ScFv 18-2 NLS LC1 and LC2. Cultures were grown in the presence or absence of IPTG, and total cell extracts (T), inclusion bodies (IB) or periplasmic extract (P) were analyzed by SDS-PAGE and anti-E tag or anti-MBP immunoblotting. B. Peptide ELISA performed as in Fig. 2 using 20 μ g of total periplasmic extract from ScFv 18-2 NLS-expressing cells or 0.2 μ g of purified ScFv 18-2 as indicated.

Fig. 5. Purification, stability, and cleavage of ScFv 18-2 NLS. **A.** Amylose affinity purification. Proteins in lanes labeled ON (onput) and FT (flow-through) represent 0.002% of a 2 L culture. Eluted fractions represent 0.02 % of a 2 L culture. Proteins were resolved by SDS-PAGE and detected by anti-E-tag or anti-MBP immunoblotting. Positions of MBP-ScFv 18-2 NLS and contaminating MBP are indicated. **B.** Peptide ELISA performed as in Fig. 2, using 0.4 μ g of total periplasmic extract or purified ScFv 18-2 NLS, or 0.2 μ g of purified ScFv 18-2 as indicated. **C.** Anti-E-tag affinity chromatography of MBP-ScFv 18-2 NLS LC1. Onput and flow through contained about 0.8 μ g total protein (0.3% of concentrated eluate from amylose affinity column). Proteins were analyzed as in Panel A. **D.** Peptide ELISA performed as in

Panel B. E. Titration of MBP-ScFv 18-2 NLS LC1 and ScFv 18-2. F. Optimization of storage conditions. Peptide ELISA was performed as in Panel B using 1.5 μg of amylose affinity-purified MBP-ScFv 18-2 NLS LC1 or 0.2 μg ScFv 18-2 as indicated. G. Cleavage of MBP-ScFv 18-2 NLS LC1. Reactions contained 20 μg of protein substrate and 0 or 4 μg Factor Xa in a volume of 20–24 μl of PBS with 20% glycerol. Incubation was for 18 h at room temperature (RT) or at 4 $^{\circ}\text{C}$ as indicated. Products were analyzed by SDS-PAGE and anti-E tag immunoblotting. Arrows denote uncleaved substrate and cleaved product as indicated. Image analysis using ImageJ (<http://rsb.info.nih.gov/ij>) indicated that 68% of the substrate was cleaved in the RT reaction and 54% at 4 $^{\circ}\text{C}$. H. Peptide ELISA using 3.75 μg of MBP-ScFv18-2 NLS LC1 (Input), Factor Xa cleavage mixture, or 0.2 μg of ScFv 18-2.

A. ScFv 18-2 NLS

B. ScFv 18-2 (without NLS)

Fig. 6.

Cellular uptake and nuclear entry of scFvs. BioPORTER protein delivery reagent was used to facilitate uptake into HeLa cells as described in Materials and Methods. A. ScFv 18-2 NLS. B. ScFv 18-2 (without NLS). Scale is same in all fields (note scale bar in B). White squares denote region that is enlarged in insets. Red, blue, and green channel images were collected separately and merged as indicated. DAPI staining indicates cell nuclei.

Table 1Purification of MBP-ScFv 18-2 NLS LC1^a

Step	Volume (ml)	Total protein (mg)
Periplasmic extract	800	65 ± 13
Amylose-agarose eluate	44.8 ± 3.4	4.5 ± 2.0
Dialyzed and concentrated	1.7 ± 0.4	4.0 ± 1.4

^aValues reflect mean and standard deviation from 5 consecutive purifications. Each purification began with 2 L of culture and was performed as described in Materials and Methods. The colloidal blue-stained gel in Fig. 5A is representative of final purity. Densitometry of the image indicated that the 75 kDa band containing MBP-ScFv 18-2 NLS LC1 accounted for about 33% of the stainable material in the sample.