Available online at www.sciencedirect.com
SCIENCE DIRECT
Protein Expression and Purification 46 (2006) 233–239
Protein Expression & Purification
www.elsevier.com/locate/yprep

Overexpression and purification of scytovirin, a potent, novel anti-HIV protein from the cultured cyanobacterium *Scytonema varium*

Changyun Xiong [a], Barry R. O'Keefe [a,*], Istvan Botos [b,1], Alexander Wlodawer [b], James B. McMahon [a]

[a] *Molecular Targets Development Program, National Cancer Institute-Frederick, National Institutes of Health, Frederick, MD, USA*
[b] *Protein Structure Section, Macromolecular Crystallography Laboratory, National Cancer Institute, Frederick, MD, USA*

Received 24 June 2005, and in revised form 12 September 2005
Available online 24 October 2005

Abstract

Scytovirin (SVN) is a novel anti-human immunodeficiency virus (HIV) protein isolated from aqueous extracts of the cultured cyanobacterium *Scytonema varium*. The protein consists of a single 95-amino acid chain with significant internal sequence duplication and 10 cysteines forming five intrachain disulfide bonds. A synthetic gene that encodes scytovirin was constructed, and expressed in *Escherichia coli*, with thioredoxin (TRX) fused to its N-terminus (TRX-SVN). Most of the expressed protein was in soluble form, which was purified by a polyhistidine tag affinity purification step. SVN was then cleaved from TRX with enterokinase and separated from the TRX partner by C18 reversed-phase HPLC. This production method has proven superior to earlier synthetic attempts and recombinant procedures using a standard expression system. The current system resulted in yields of 5–10 mg/L of purified SVN for structural studies and for preclinical development of SVN as a topical microbicide for HIV prophylaxis.
Published by Elsevier Inc.

Keywords: Scytovirin, SVN; Microbicide; HIV; anti-HIV; Cyanobacterium; Recombinant expression

Scytovirin (SVN),[2] a potent anti-human immunodeficiency virus (HIV) protein, was originally isolated from aqueous extracts of the cyanobacterium *Scytonema varium* [1]. SVN, with a molecular mass of 9713 Da, contains five intrachain disulfide bonds and binds to HIV-1 proteins gp120, gp160, and gp41 but not to the cellular receptor CD4 [1]. Low nanomolar concentrations of SVN inactivate laboratory strains and primary isolates of HIV-1. The inhibition has been shown to involve selective interactions between SVN and high mannose oligosaccharide-bearing glycoproteins [1]. Specifically, SVN interacts with oligosaccharides containing α1–2, α1–6 trisaccharide units [2].

In general, proteins purified from natural product extracts have yields of only microgram quantities per gram of starting material [3]. Recombinant DNA technology provides a ready means of re-supply of bioactive proteins which can be used for follow-up studies such as structure determination by NMR or X-ray crystallography and to further preclinical development. One potential difficulty in the production of recombinant SVN, is that all of the 10 Cys residues in the protein are involved in creating disulfide bonds and must find their specific partners to generate a correctly folded protein. Initial attempts using standard techniques for recombinant production of SVN were found to be inadequate, producing significant amounts of misfolded protein, so additional efforts were required for successful recombinant expression.

It has been shown that the 11 kDa thioredoxin (TRX) fusion partner not only enhances the solubility of many expressed fusion proteins, but appears to catalyze the formation of disulfide bonds in the cytoplasm [4,5]. In this report, we used the pET32 system for successfully produc-

* Corresponding author. Fax: +1 301 846 6872.
E-mail address: okeefe@ncifcrf.gov (B.R. O'Keefe).
[1] Current address. Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD, USA.
[2] *Abbreviations used:* SVN, scytovirin; TRX, thioredoxin; ECL, enhanced chemiluminescence; PMSF, phenylmethylsulfonylfluoride.

1046-5928/$ - see front matter. Published by Elsevier Inc.
doi:10.1016/j.pep.2005.09.019

ing recombinant SVN from *Escherichia coli*. The recombinant SVN fusion protein was purified in milligram quantities by metal chelation affinity chromatography, rEK digestion and reversed-phase HPLC. The final product retained full biological activity in both a gp160 binding assay and in a cell-based cytoprotection assay for HIV.

Methods and materials

Materials

Restriction endonucleases and T4 DNA ligase were obtained from New England Biolabs (Beverly, MA). *Pfu* DNA polymerase was obtained from Stratagene (La Jolla, CA). Rabbit anti-SVN polyclonal antibody was produced in our laboratory as described previously [1]. Reagents for sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) were obtained from Invitrogen (Carlsbad, CA). All other chemicals were of analytical reagent grade from Sigma (St. Louis, MO). The plasmid vector pET32C(+) and *E. coli* strain BL21 *TRX*B(DE3)pLysS were from Novagen (Madison, WI).

Construction of the synthetic gene for scytovirin

The deduced amino acid sequence was back-translated to a DNA sequence using an *E. coli* codon preference table, and supplemented with a termination codon. The sequence of the synthetic gene and the sequence of the encoded protein are shown in Fig. 1A.

As SVN shows strong internal sequence duplication, two separate fragments, fragment 1 (coding for scytovirin protein sequence 1–47), and fragment 2 (coding for scytovirin protein sequence 48–95) were prepared. The coding sequences for fragments 1 and fragment 2 were initially synthesized as 4 overlapping, complementary oligonucleotides which were assembled to form a double-stranded coding sequence. The two fragments were then linked together to form the synthetic gene by a gene splicing method [6].

The *Nco*I-*Xho*I-digested PCR fragment of SVN was inserted into *Nco*I-*Xho*I-digested pET32C(+) vector. The linker sequence between the enterokinase digestion site and SVN DNA was deleted from the construct by site-directed mutagenesis method (GeneTailor Site-Directed Mutagenesis System, Invitrogen) and the resulting plasmid vector pET(SVN) was transformed into *E. coli* strain DH5α. The sequence of the gene was confirmed by DNA sequencing.

Expression of SVN fusion protein in E. coli

The plasmid pET(SVN) under the control of a T7 promoter was transformed into *E. coli* BL21 *TRX*B(DE3)pLysS. The cells containing the construct were grown at 37 °C in LB medium (Luria–Bertani broth) containing 100 μg/ml ampicillin, 34 μg/ml chloramphenicol, and 15 μg/ml kanamycin to an OD_{600} of 0.5–0.8. The cells were induced with isopropyl-1-thio-β-D-galactopyranoside (IPTG) to a final concentration of 1 mM, then incubated for an additional 4 h, and collected by centrifugation. The cells were either prepared immediately for SDS–PAGE assay or frozen at −80 °C.

Cell fractionation and SDS–PAGE

Frozen cells were thawed in 50 mM phosphate buffer containing 0.4 mg/ml lysozyme and 1 mM phen-

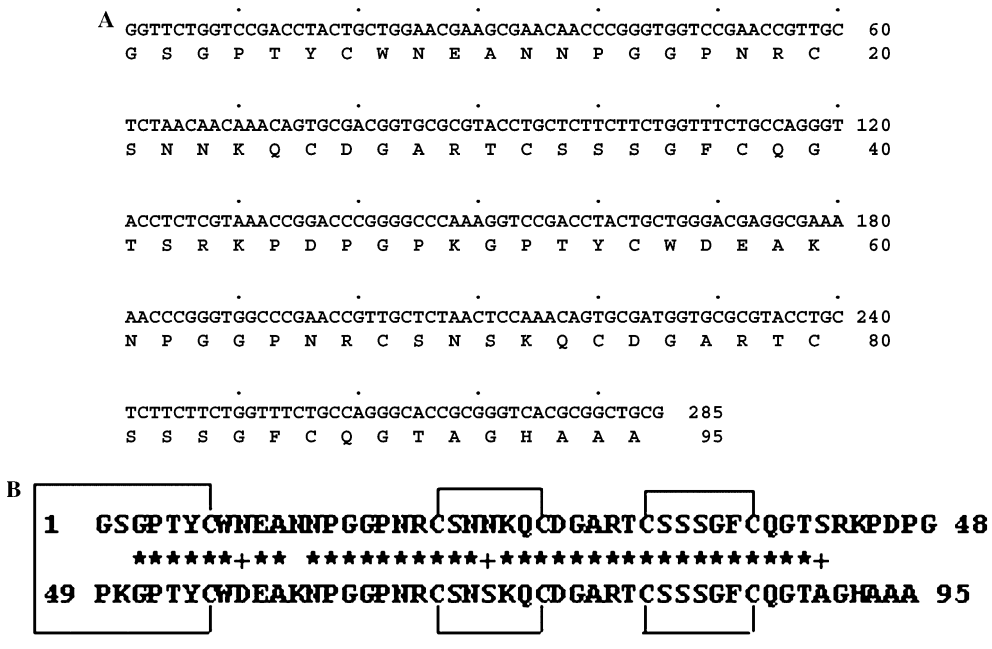

Fig. 1. Nucleotide and amino acid sequence of scytovirin showing domain sequence alignment. (A) Nucleotide sequence for the synthetic gene scytovirin and the encoded protein sequence. (B) Sequence alignment between the two domains of SVN. Amino acids are represented by single-letter codes. Identical sites are represented by ∗. Conserved sites are represented by +. Disulfide bonds are marked with solid lines above the sequence.

ylmethylsulfonylfluoride (PMSF). Efficient lysis occurred over several minutes. DNase I (20 μg/ml) and $MgCl_2$ (10 mM) were then added to digest DNA. Crude lysates were centrifuged to separate soluble and insoluble protein fractions, and analyzed on 10% Bis–Tris SDS–polyacrylamide gel under reducing conditions.

Protein purification

The expressed proteins were initially purified by histidine tag metal affinity chromatography (BD) according to manufacturer's protocol. Briefly, 25 ml cleared lysate from 1 liter of LB cultured bacterial pellet was incubated with 2 ml affinity resin that had been equilibrated with 10 column volumes of buffer A (50 mM phosphate buffer, pH 7.5, 300 mM NaCl, 5 mM imidazole, and 1 mM PMSF). The protein-bound resin was then washed with 10 column volumes of buffer A. Scytovirin fusion protein was then eluted with 10 column volumes of 150 mM imidazole in buffer A.

The resulting sample was filtered through an Amicon Centriprep 10 kDa molecular mass cutoff filter to a volume of 2 ml to desalt and concentrate the material. It was then digested for 16 h at room temperature with recombinant enterokinase (rEK, Novagen), 1 unit of enterokinase/10 μg of recombinant protein. rEK-cleaved scytovirin was purified using reversed-phase HPLC. The protein was bound to a Dynamax C18 300A column eluted with a gradient of 0–60% acetonitrile in 0.05% aqueous TFA over 60 min at a flow rate of 3 mL/min with UV monitoring at 210 nm and 280 nm.

Mass spectroscopic analysis

The purified recombinant scytovirin was analyzed by LC–MS using a Hewlett Packard high-performance liquid chromatography/electrospray ionization quadrupole mass spectrometer (model 1100D) running a C8 Zorbax column (2.1 × 110 mm) eluted with a linear gradient from 0 to 100% acetonitrile over 70 min in H_2O with 5% (v/v) acetic acid in the mobile phase at a flow rate of 0.2 mL/min and with UV monitoring at 280 nm. After mass spectral deconvolution according to manufacturer's protocols, the molecular weight of scytovirin was calculated to be 9712.2 Da, matching the mass determined for the native protein [1].

Immunoblotting

Protein samples were separated by Bis–Tris SDS–PAGE using a 10% (w/v) polyacrylamide gel and electroblotted on to a PVDF membrane. After blocking in 10 mM Tris–HCl, pH 8.0, containing 150 mM NaCl, 0.05% Tween 20, and 5% (w/v) non-fat milk for 1 h at room temperature, the membrane was incubated with rabbit anti-SVN polyclonal antibody at 1:5000 dilution overnight at 4 °C. The immune complexes on the membrane were then reacted with horseradish peroxidase-conjugated goat anti-rabbit IgG at 1:2000 dilution for 1 h at room temperature. Immunodetection was achieved by enhanced chemiluminescence (ECL). Densitometric scanning of ECL-blots was performed on a Molecular Dynamics 300S computing densitometer (Sunnyvale, CA) using ImageQuant V3.0 software.

Disulfide bonds determination

The presence and connectivity of disulfide bonds were determined as previously described [1]. Briefly, to a 100 μg sample of recombinant, nonreduced SVN was added to 60 μL of 100 mM ammonium bicarbonate (pH 8.0), 6 μL of acetonitrile, and 0.6 μL of a 40 μM solution of trypsin in H_2O. The mixture was incubated at 37 °C for 16 h and then analyzed by LC–MS as above. The resulting peptides were evaluated using peptide recognition software to detect all possible disulfide-linked peptide fragments [7].

Anti-HIV assays

An XTT-tetrazolium-based assay was used to determine the anti-HIV activity of SVN on acute HIV-1 infection in CEM-SS cells as previously described [8]. Eight serial 0.5 log dilutions of SVN in complete medium were performed and the protein solutions were then added to exponentially growing CEM-SS human lymphocyte cells. Cell cultures were infected with freshly thawed solutions of HIV-1_{RF} and allowed to incubate for 7 days. Metabolic reduction of the tetrazolium salt, XTT, to a colored formazan product was used to determine cellular viability at the end of the 7-day incubation period. Optical densities of all test and appropriate control cultures were measured at 450 nm.

ELISA protocols

To determine the binding of SVN for gp160, 100 ng each of gp160 and bovine serum albumin (BSA) were evaluated by ELISA as previously described [9]. Briefly, the proteins were bound to a 96-well plate, which was then rinsed three times with PBS containing 0.05% Tween 20 (TPBS) and blocked with BSA. Between subsequent steps, the plate was again rinsed with TPBS (3×). The wells were incubated with serial dilutions of SVN, followed by incubation with a 1:1000 dilution of an anti-SVN rabbit polyclonal antibody preparation. The amount of bound SVN was determined by adding a 1:2000 dilution of donkey anti-rabbit antibodies conjugated to horseradish peroxidase. After addition of the horseradish peroxidase substrate buffer and color formation, the reaction was stopped by 2 M H_2SO_4 and absorbance was measured at 450 nm for each well.

Results

Expression and purification of recombinant SVN

SVN shows strong internal sequence duplication. When amino acids 1–48 and 49–95 are aligned, 36 residues (75%)

are identical and three (6%) represent conservative amino acid changes (Fig. 1B). The bonds formed by the C20–C26 and C32–C38 disulfides correspond closely to those defined by the C68–C74 and C80–C86 disulfide links [1]. Based on this fact, two fragments covering SVN's whole sequence were synthesized separately. The two fragments were subse-

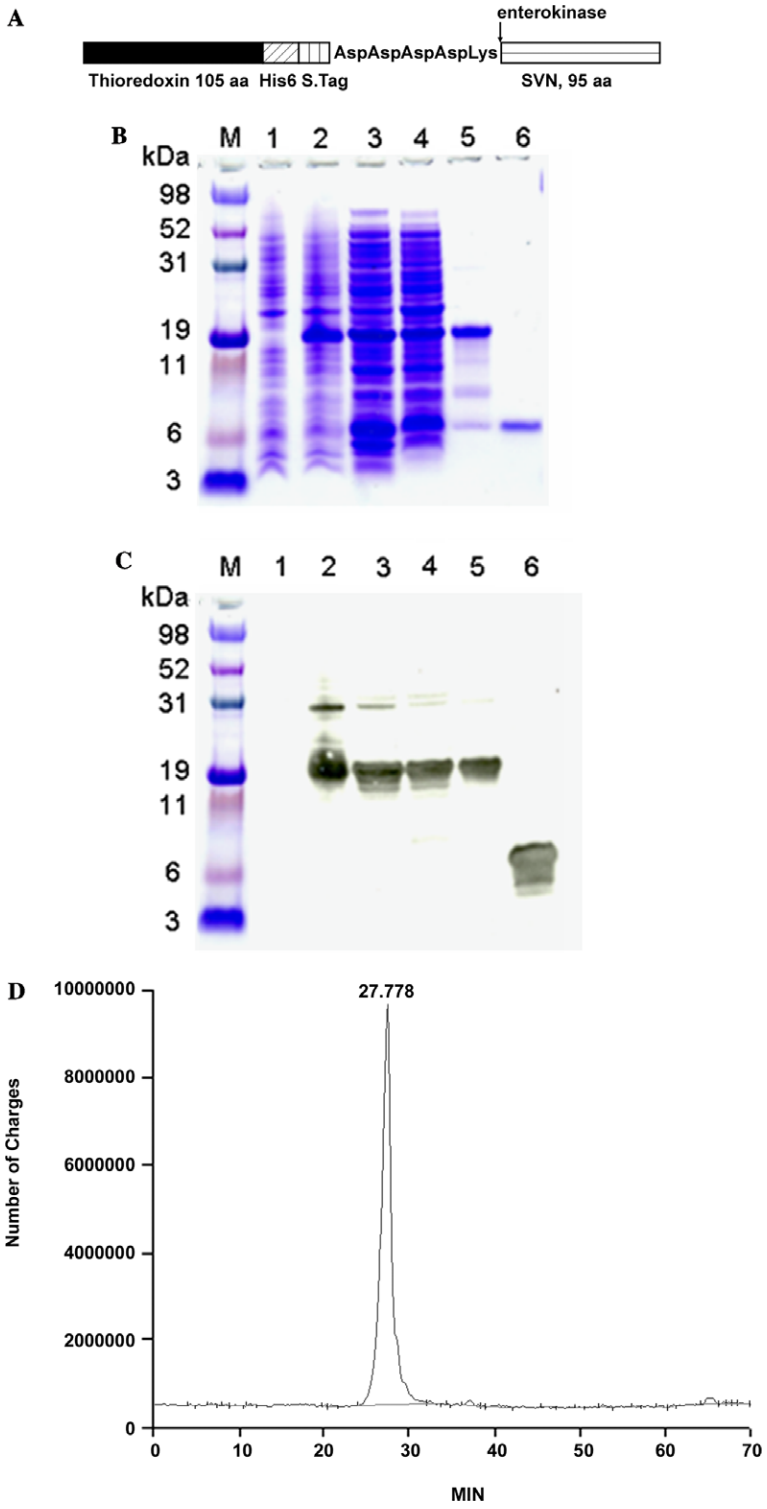

Fig. 2. Schematic of scytovirin fusion protein; SDS–PAGE, immunoblot and LC–MS analysis of the purification of scytovirin from *E. coli TRXB*BL21(DE3)pLysS. (A) Schematic used to produce scytovirin fusion protein in pET32c(+) vector. His-Tag and S-Tag between thioredoxin and scytovirin provides a method for detection and purification. The enterokinase digestion site is used for generation of scytovirin. (B) Coomassie blue-stained gel. (C) Immunoblot analysis with anti-scytovirin polyclonal antibody. Lane 1, whole cells without IPTG induction; lane 2, whole cells with IPTG induction; lane 3, soluble protein; lane 4, insoluble protein; lane 5, scytovirin fusion protein after His tag affinity purification; lane 6, scytovirin protein after reversed-phase HPLC purification followed with rEK digestion of scytovirin fusion protein. (D) LC–MS analysis of HPLC-purified SVN ($m/z = 9712.2$).

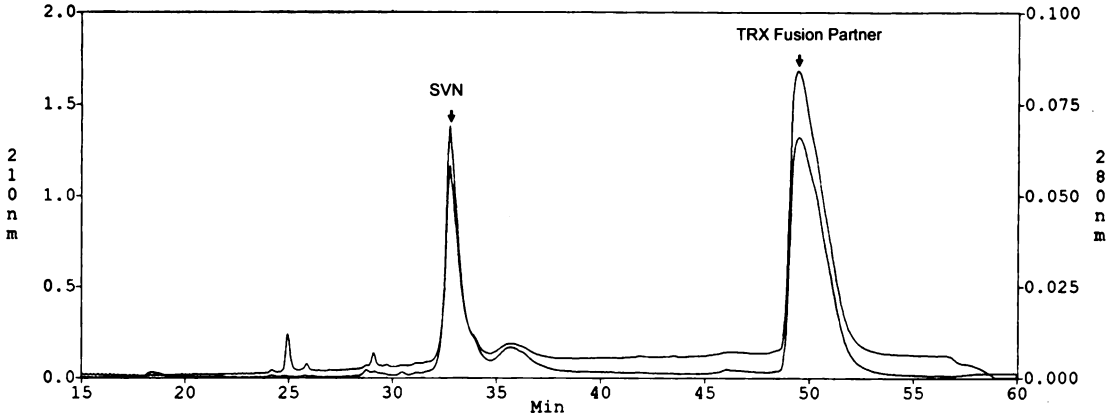

Fig. 3. Reverse-phase HPLC purification of rEK-cleaved scytovirin. One milliliter of 1 mg/ml rEK digested fusion protein suspended in 20 mM Tris–HCl, pH 7.5, was injected onto a C18 column as described in Methods and materials.

quently linked together to form the whole SVN synthetic gene by gene splicing PCR.

The pET32c(+) vector is designed for cloning and high-level expression of protein sequences fused with the 109 aa TRX Tag thioredoxin protein. Cloning sites are available for producing fusion proteins also containing cleavable His-Tag and S-Tag sequences for detection and purification. The enterokinase digestion site between SVN and its fusion protein partner provides an easy way to separate SVN from the fusion partner (Fig. 2A).

The expression level of the fusion protein from a T7 promoter was as high as 50% of the total cellular protein, and about 50% of that was soluble (Fig. 2B). Immunoblotting assays (Fig. 2C) showed that the expressed proteins were recognized by anti-SVN polyclonal antibodies.

The SVN fusion protein was approximately 90% pure following one step of metal affinity chromatography. LC–MS analysis revealed that the molecular weight of this fusion protein is 26637.03 Da, which matches the calculated value of 26637 Da. Following enterokinase digestion overnight at room temperature, SVN was purified from the TRX fusion partner by reversed-phase HPLC to yield the homogeneous protein. Cleaved SVN matching the native sequence, eluted at 34 min, while the TRX fusion fragment eluted at 50 min (Fig. 3). The resulting purified SVN was shown to be homogeneous by both SDS–PAGE (Fig. 2B) and LC–MS (Fig. 2D). LC–MS showed that the molecular weight of the recombinant SVN was 9712.2 Da, which is consistent with the native scytovirin (MW = 9713 Da). Approximately, 5 mg of HPLC-purified scytovirin per liter of cell culture was obtained by this procedure.

Disulfide bond determination

The disulfide bond pattern of native SVN was previously reported [1]. LC–MS and peptide recognition software [7] were used to identify the theoretical disulfide-bonded fragments of SVN. The values of six fragments of trypsin-digested recombinant SVN (m/z = 1317.80, 1553.14, 2719.56, 3850.84, 3158.05, and 2508.41 Da) were input into the software to establish the disulfide bond pattern (Fig. 4). The data retrieved from the software matched the previously determined pattern and demonstrated that the recombinant SVN had the same disulfide pattern as the native SVN [1].

Fig. 4. Disulfide-bonding pattern of recombinant scytovirin. The values of six fragments of trypsin-digested recombinant SVN (m/z = 1317.80, 1553.14, 2719.56, 3850.84, 3158.05, and 2508.41 Da) were input into the software (http://sx102a.niddk.nih.gov/peptide/) to establish the disulfide-bonding pattern.

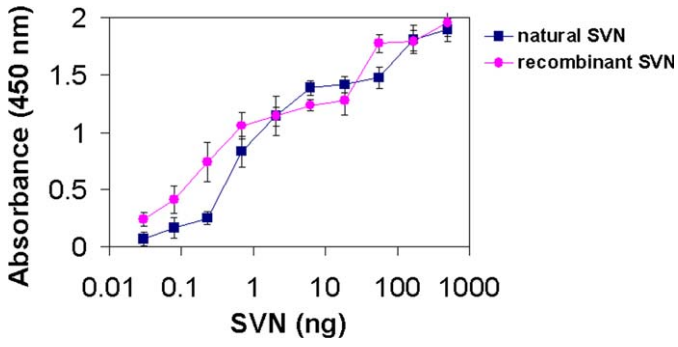

Fig. 5. ELISA study of the concentration-dependent binding of recombinant or natural scytovirin to gp160. Glycosylated gp160 was bound to an ELISA plate and then treated with scytovirin. The amount of bound scytovirin was determined by the absorbance at 450 nm as described in Methods and materials.

ELISA study of scytovirin–gp160 interaction

We have reported that scytovirin binds to the HIV-1 viral coat proteins gp120, gp160, and gp41 [1]. To compare the binding activity of native and recombinant SVN, we tested both recombinant SVN and native SVN for their ability to bind to HIV-1 gp160 (Fig. 5). The results indicated that both proteins bound to gp160 with similar concentration dependence.

Concentration dependence of antiviral activity of scytovirin

Native SVN displayed potent anticytopathic activity against laboratory strains and primary isolates of HIV-1 with EC_{50} values ranging from 0.3 to 22 nM [1]. In side-by-side in vitro *XTT*-based anti-HIV assays using CEM-SS host cells and HIV-1_{RF}, the recombinant scytovirin potently inhibited HIV-induced cytopathicity with an EC_{50} of 4.5 nM, while native SVN displayed a range of EC_{50} values from 0.3–7 nM.

Discussion

High-level expression of an SVN fusion protein was achieved by synthesizing a synthetic gene for the protein and expressing it under the control of T7 promoter in *E. coli*. The pET-32 vector used here contained a 109 aa TRX.Tag thioredoxin protein which increased the solubility of the target protein. The host cell BL21 *TRXB* (DE3)pLysS possesses a thioredoxin reductase (*TRXB*) mutation which has been shown to allow the formation of disulfide bonds in the *E. coli* cytoplasm [10]. The combination of this vector and host cell resulted in the correct disulfide bond formation of the soluble recombinant SVN. As SVN contains five disulfide bonds in a protein weighing only 9.7 kDa, it was necessary to use this expression system for proper disulfide bond formation, since previous attempts using other systems were not successful at producing properly folded, bioactive SVN. The purified recombinant protein was found to have anti-HIV activity, gp160-binding activity, and the correct disulfide-bonding pattern, indicating that it had folded into an active conformation. Furthermore, mass spectrometry demonstrated that the protein had the expected sequence and that there were no obvious posttranslational modifications.

The presence of the histidine tag and S tag between the fusion protein partners provided an easy way to capture and purify the target protein. The fusion protein was found to be 90% pure following histidine tag affinity chromatography. The inclusion of a rEK digestion site provided a simple method for cleaving off the recombinant SVN and resulted in homogenous, biologically active SVN.

Western blotting results showed that a small fraction of the fusion protein existed in dimeric form, whereas the molecular weight determined by mass spectrometry indicated a monomer (data not shown). This result confirms those previously described for native SVN, in which evidence of dimerization appears in both SDS–PAGE and Western blots, and further indicates the potential presence of homodimerization in SVN.

SVN was naturally produced in the cyanobacterium *S. varium* [1]. Previous efforts in our laboratory and by others have failed to produce significant additional quantities of SVN from this organism. Furthermore, synthetic efforts to produce substantial quantities of SVN have also failed due to the production of misfolded protein with variable disulfide bond formation [Dr. Wuyuan Lu, personal communication]. The ability to produce adequate quantities of SVN will allow for three-dimensional studies of its structure by both X-ray crystallography and solution-phase NMR. In addition, further in vivo antiviral studies with SVN will now be possible due to the availability of this production system. As SVN has been shown to be potently active against HIV, microbicidal studies, similar to those recently published with the cyanobacterial protein cyanovirin-N [11,12], will now be possible. The efficient recombinant production of SVN will also allow for studies on its spectrum of antiviral activity and biochemical studies of its specific molecular target. The production system reported here has demonstrated advantages in terms of increased expression, enhanced solubility, protection from proteolysis, improved folding, and protein purification via affinity chromatography [13].

Acknowledgments

This research was supported by the Intramural Research Program of the National Cancer Institute, National Institutes of Health (USA). We are grateful to Carrie Saucedo and Jennifer Wilson for assistance in anti-HIV assays and technical assistance in making figures.

References

[1] H.R. Bokesch, B.R. O'Keefe, T.C. McKee, L.K. Pannell, G.M. Patterson, R.S. Gardella, R.C. Sowder 2nd, J. Turpin, K. Watson, R.W. Buckheit Jr., M.R. Boyd, A potent novel anti-HIV protein from the cultured cyanobacterium *Scytonema varium*, Biochemistry 42 (9) (2003) 2578–2584.

[2] E.W. Adams, D.M. Ratner, H.R. Bokesch, J.B. McMahon, B.R. O'Keefe, P.H. Seeberger, Oligosaccharide and glycoprotein microarrays as tools in HIV glycobiology: glycan-dependent gp120/protein interactions, Chem. Biol. 11 (6) (2004) 875–881.

[3] B.R. O'Keefe, Biologically active proteins from natural product extracts, J. Nat. Prod. 64 (10) (2001) 1373–1381.

[4] E.J. Stewart, F. Aslund, J. Beckwith, Disulfide bond formation in the *Escherichia coli* cytoplasm: an in vivo role reversal for the thioredoxins, EMBO J. 17 (19) (1998) 5543–5550.

[5] V.S. Stoll, A.V. Manohar, W. Gillon, E.L. MacFarlane, R.C. Hynes, E.F. Pai, A thioredoxin fusion protein of VanH, a D-lactate dehydrogenase from *Enterococcus faecium*: cloning, expression, purification, kinetic analysis, and crystallization, Protein Sci. 7 (5) (1998) 1147–1155.

[6] R.M. Horton, Z.L. Cai, S.N. Ho, L.R. Pease, Gene splicing by overlap extension: tailor-made genes using the polymerase chain reaction, Biotechniques 8 (5) (1990) 528–535.

[7] http://sx102a.niddk.nih.gov/peptide.

[8] R.J. Gulakowski, J.B. McMahon, P.G. Staley, R.A. Moran, M.R. Boyd, A semiautomated multiparameter assay for anti-HIV drug screening, J. Virol. Methods 33 (1991) 87–100.

[9] B.R. O'Keefe, S.R. Shenoy, D. Xie, W. Zhang, J.M. Muschik, M.J. Currens, I. Chaiken, M.R. Boyd, Analysis of the interaction between the HIV-inactivating protein cyanovirin-N and soluble forms of the envelope glycoproteins gp120 and gp41, Mol. Pharmacol. 58 (2000) 982–992.
[10] A.I. Derman, W.A. Prinz, D. Belin, J. Beckwith, Mutations that allow disulfide bond formation in the cytoplasm of *Escherichia coli*, Science 262 (5140) (1993) 1744–1747.
[11] M.R. Boyd, K.R. Gustafson, J.B. McMahon, R.H. Shoemaker, B.R. O'Keefe, T. Mori, R.J. Gulakowski, L. Wu, M.I. Rivera, C.M. Laurencot, M.J. Currens, J.H. Cardellina 2nd, R.W. Buckheit Jr., P.L. Nara, L.K. Pannell, R.C. Sowder 2nd, L.E. Henderson, Discovery of cyanovirin-N, a novel human immunodeficiency virus-inactivating protein that binds viral surface envelope glycoprotein gp120: potential application to microbicide development, Antimicrob. Agents Chemother. 41 (7) (1997) 1521–1530.
[12] C.C. Tsai, P. Emau, Y. Jiang, M.B. Agy, R.J. Shattock, A. Schmidt, W.R. Morton, K.R. Gustafson, M.R. Boyd, Cyanovirin-N inhibits AIDS virus infections in vaginal transmission models, AIDS Res. Hum. Retroviruses 20 (1) (2004) 11–18.
[13] D.R. Smyth, M.K. Mrozkiewicz, W.J. McGrath, P. Listwan, B. Kobe, Crystal structures of fusion proteins with large-affinity tags, Protein Sci. 12 (7) (2003) 1313–1322.