Analysis of Sequence Requirements for Biological Activity of Cyanovirin-N, a Potent HIV(Human Immunodeficiency Virus)-Inactivating Protein^{1,2}

Toshiyuki Mori,* Robert H. Shoemaker,* Robert J. Gulakowski,* Benjamin L. Krepps,* James B. McMahon,* Kirk R. Gustafson,* Lewis K. Pannell,† and Michael R. Boyd*.

*Laboratory of Drug Discovery Research and Development, Development Therapeutics Program, Division of Cancer Treatment, Diagnosis and Centers, National Cancer Institute-FCRDC, Frederick, Maryland 21702-1201; and †Laboratory of Analytical Chemistry, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, Maryland 20892-0805

Received July 28, 1997

Site-directed mutagenesis of DNA constructs coding for the novel, HIV-inactivating proteins cyanovirin-N (CV-N) and FLAG-cyanovirin-N (F-CV-N) was performed using mutagenic oligonucleotide primers in the polymerase chain reaction or by a restriction site elimination maneuver. The mutant constructs were expressed in Escherichia coli and the recombinant protein products were tested for binding to the HIV surface envelope glycoprotein gp120 and for antiviral activity against infectious HIV. Results showed an overall very high correlation (r²> 0.9) between the relative gp120 binding affinities and the anti-HIV activities of CV-N, F-CV-N, and the various mutants. An outlier, however, was a mutant which lacked one of the internal disulfide linkages normally present in CV-N and which showed modest gp120 binding but no antiviral activity against HIV. These findings are consistent with the view that gp120 binding is a necessary but not sufficient requirement for the HIV-inactivating activity of CV-N and related proteins; the sequence speci-

Abbreviations used: HIV, human immunodeficiency virus; gp120, HIV envelope glycoprotein gp120; CD4, CD4 cellular receptor; CV-N, cyanovirin-N; F-CV-N, CV-N bearing the FLAG octapeptide leader sequence (AspTyrLysAspAspAspAspLys); *E. coli, Escherichia coli*; PCR, polymerase chain reaction; IPTG, isopropyl-1-thio- β -D-galactopyranoside; SDS-PAGE, sodium dodecyl sulfate – polyacrylamide gel electrophoresis; ELISA, enzyme-linked immunosorbent assay; PBS, phosphate-buffered saline; TPBS, PBS containing 0.05% Tween-20; BSA, bovine serum albumin.

ficities for gp120 binding and anti-HIV activity are not identical. © 1997 Academic Press

The discovery of cyanovirin-N (CV-N) from extracts of the cultured cyanobacterium Nostoc ellipsosporum has recently been described (1,2). Low nanomolar concentrations of CV-N irreversibly inactivated diverse Tlymphocyte-tropic, macrophage-tropic and dual-tropic strains and primary isolates of HIV type 1, as well as strains of HIV type 2 and simian immunodeficiency virus. The inhibition appeared to be mediated, at least in part, through high-affinity interactions of CV-N with the viral surface envelope glycoprotein gp120 (generically termed "gp120") that differed from gp120 interactions with either the cellular receptor CD4 or with antibodies directed against viral gp120 neutralizing determinants (1). The exceptional range of anti-HIV activity of CV-N against both lab-adapted and primary clinical isolates of HIV contrasted markedly with that of soluble CD4 which, in accord with prior observations of others (3-5), lacked activity against primary isolates of HIV-1. CV-N also prevented fusion and transmission of HIV-1 between infected and uninfected T-cells (1). The U.S. National Cancer Institute is currently conducting preclinical development of CV-N as a potential anti-HIV virucide.

CV-N is composed of a linear sequence of 101 amino acid residues (Fig. 1) having no post-translational modifications (1,2) and containing 4 cysteine residues that form 2 intramolecular disulfide bonds (2). The active protein does not appear as a dimer or higher oligomer (2). The biological activity of CV-N is highly resistant to physicochemical denaturation (1). There is no significant sequence homology between CV-N and any other proteins or transcription products of

¹ A DNA coding sequence for cyanovirin-N is deposited in the Gen-Bank database (Accession No. L48551).

² This is part 39 in the series HIV-Inhibitory Natural Products; for part 38 see Gustafson, K. R., Sowder, R. C., II, Henderson, L. E., Cardellina, J. H., II, McMahon, J. B., Rajamani, U., Pannell, L. K., and Boyd, M. R. (1997) Submitted for publication.

³ To whom correspondence should be addressed at NCI-FCRDC, Building 1052, Room 121, Frederick, MD 21702-1201. Fax: (301) 846-



FIG. 1. Amino acid sequence of CV-N. To facilitate visualization of internal sequence homology, the sequence has been separated into two portions. The upper portion of the sequence represents amino acids 1-50 (Domain 1) and the lower portion of the sequence represents amino acids 51-101 (Domain 2). Dark and light boxes indicate identical and similar amino acids, respectively. Disulfide bonds are indicated by linking cysteine residues. To facilitate alignment, a gap (*) has been inserted between positions 16 and 17.

DNA sequences that are accessible in current protein or nucleic acid databases (1,2). Interestingly, a sequence comparison between the first 50 (Domain 1; D1) and last 51 (Domain 2; D2) amino acids of CV-N show high sequence similarity (32% identity and 26% conservative changes, thus 58% overall homology) (Fig. 1), suggesting that the cyanobacterial gene coding for CV-N might have arisen through tandem gene duplication (2).

A DNA coding sequence corresponding to the chemically-deduced protein sequence has been synthesized, and a recombinant protein indistinguishable from natural CV-N was successfully expressed in *E. coli* (1). Therefore, in the present investigation, a series of DNA constructs coding for mutant forms of CV-N were produced and expressed in *E. coli* to explore functional domains and sequence requirements for gp120 binding and anti-HIV activity of CV-N.

MATERIALS AND METHODS

Materials. Restriction endonucleases, T4 polynucleotide kinase and T4 DNA ligase were obtained from Pharmacia Biotech (Uppsala, Sweden). Reagents for SDS-PAGE were obtained from Bio-Rad (Hercules, CA). All other chemicals are analytical reagent grade.

Amplification. Oligonucleotides -3C (5'-TCT CTG CTC GAG TTA CAG GGT ACC GTC GAT GTT AGC GA-3'), -8C (5'-TCT CTG CTC GAG TTA GTT AGC GAT GTG GTC GTC CAG GT-3'), -1N (5'-AGA TAT CAT ATG AAA AAG ACA GCT ATC GCG ATT GCA GTG GCA CTG GCT GGT TTC GCT ACC GTT GCG CAA GCT GGT AAA TTC TCC CAG ACC TGC TA-3'), -2N (5'-AGA TAT CAT ATG AAA AAG ACA GCT ATC GCG ATT GCA GTG GCA CTG GCT GGT TTC GCT ACC GTT GCG CAA GCT AAA TTC TCC CAG ACC TGC TAC AA-3'), -3N (5'-AGG AGA TAT CAT ATG AAA AAG ACA GCT ATC GCG ATT GCA GTG GCA CTG GCT GGT TTC GCT ACC GTT GCG CAA GCT TTC TCC CAG ACC TGC TAC AAC T-3'), N26 (5'-CATCATAACGGTTCTGGCAAATATTC-3'), C24 (5'-CTG TAT CAG GCT GAA AAT CTT CTC-3'), D1-1 (5'-GAT GAC AAG CTT GGT AAA TTC-3'), D1-2 (5'-CTG CCA TTT CAG GGA ACC-3'), D1-3 (5'-CTT GGT AAA TTC TCC CAG AC-3'), D1-4 (5'-AGT CCT CGA GTT ACT GCC ATT TCA GGG AAC C-3'), D2-1 (5'-CAT GAA GCT TTC CAA CTT CAT CGA AAC CTG-3'), D2-2 (5'-TTC GTA TTT CAG GGT ACC GT-3'), D2-3 (5'-CCG TCC AAC TTC ATC GAA A-3'), D2-4 (5'-CTC TCA CTC GAG TTA TTC GT-3'), #8 (5'-GAT AGC GGA GTT GTA ACT AGT CTG GGA GAA TTT AC-3'), #22 (5'-GTT GGT ACG TTC ACT AGT GGA GGT CAG AAC G-3'), #58 (5'-GCC AGC TGG GTG TTT CTA GAG GTT TCG ATG AAG-3'), and #73 (5'-GAG CAC GGG TTT TTG ATT CAG CAG CCA G-3') were

prepared using a DNA synthesizer (Applied Biosystems, Foster City, CA). PCR was carried out using 10 ng of plasmid pPBS7 or pPBS7(–) as template (1) and reagents as per the manufacturer's instructions (Stratagene, La Jolla, CA) and 20 pmol of primer. Each PCR totaled 15 cycles consisting of denaturation at 95°C for 1 min, annealing at 65°C for 1 min and polymerization at 75°C for 2 min in each cycle. The amplified fragments were purified on agarose (FMC, Rockland, ME).

Plasmids, bacterial strains, and cell lines. The plasmids encoding various mutant forms of CV-N are shown in Fig.2. The method of construction of plasmids encoding native CV-N, and FLAG-CV-N (F-CV-N) which contains the FLAG (8) leader sequence (AspTyrLysAsp-AspAspAspLys) was described previously (1). The plasmids pPBS7-(-1N), pPBS7(-2N) and pPBS7(-3N), encoding deletion mutants of CV-N which were truncated by 1, 2 or 3 amino acids from the N-terminus respectively, were constructed by insertion of Nde I-Xho I digested PCR fragments (amplified using -1N/C24, -2N/C24 and -3N/C24, respectively) into an Nde I-Xho I fragment of plasmid pPBS7(-). The plasmids pPBS7(-3C) and pPBS7(-8C), encoding deletion mutants of F-CV-N which were truncated by 3 or 8 amino acids from the Cterminus respectively, were constructed by insertion of Hind III-Xho I digested PCR fragments (amplified using -3C/N26 and -8C/N26 respectively) into a *Hind III-Xho I* fragment of plasmid pPBS7. The plasmid pD1D1 was constructed by insertion of a Hind III digested PCR fragment (amplified using D1-1 and D1-2) and an XhoI digested PCR fragment (amplified using D1-3 and D1-4) into a Hind III-Xho I fragment of plasmid pPBS7. In the cases of pD2D2 and pD2D1, the first amino acid after the FLAG peptide was changed from Pro to Leu to allow connection using the Hind III site. The plasmid pD2D2 was constructed by insertion of a Hind III digested PCR fragment (amplified using D2-1 and D2-2) and an Xho I digested PCR fragment (amplified using D2-3 and D2-4) into a Hind III-Xho I fragment of plasmid pPBS7. The plasmid pD2D1 was constructed by insertion of a Hind III digested PCR fragment (amplified using D2-1 and D2-2) and an Xho I digested PCR fragment (amplified using D1-3 and D1-4) into a Hind III-Xho I fragment of plasmid pPBS7. The plasmids pC>S#8-22 and pC>S#58-73 encoding mutant FLAG-CV-N's in which either cysteines 8 and 22 or cysteines 58 and 73 were replaced with serine residues were constructed using the phosphorylated primers #8, #22 and #58, #73, respectively, according to U. S. E. Mutagenesis Kit from Pharmacia Biotech (Uppsala, Sweden). All plasmid constructs were confirmed by DNA sequencing. The constructs were used to transform $E.\ coli\, DH5\alpha F'IQ$ (Gibco BRL, Gaithersburg, MD) grown in LB broth with ampicillin (50 μ g/ml) and 0.4% glucose at 37°C. When the cultures reached an appropriate density (Ab₆₀₀ nm, 0.2-0.3), isopropyl-1-thio- β -D-galactopyranoside (IPTG) was added to a final concentration of 0.5 mM for induction of gene expression. All recombinant proteins were secreted into the periplasmic space of *E. coli* by virtue of the outer membrane protein A signal peptide (9). Cells were harvested 60 min after induction and, where appropriate, periplasmic fractions were prepared according to the pFLAG system manufacturer's recommendations (International Biotechnologies, Inc., New Haven, CT). Sequence components of the mutant CV-N proteins are shown schematically in Figure 2.

Protein purification and analysis. Recombinant CV-N and mutant proteins, CV-N(-1N), CV-N(-2N) and CV-N(-3N) were purified by HPLC as described previously for natural CV-N (1, 2). Purification of the F-CV-N and the mutants F-C>S#8-22, F-C>S#58-73, F-CV-N(-3C), F-CV-N(-8C), F-D1D1, F-D2D2 and F-D2D1 was facilitated by presence of the FLAG peptides. The FLAG-proteins were first concentrated from periplasmic extracts by affinity chromatography, then further purified by HPLC on a 1.0×25 cm MICROSORB C_{18} column (Rainin Instrument Co, Inc., Emeryville, CA), using a Bio-CAD SPRINT perfusion Chromatography System (PerSeptive Biosystems, Cambridge, MA) eluted with a gradient of increasing concentration of CH $_3$ CN in H_2 O (0.05% TFA, v/v in the mobile phase) at a flow rate of 2 ml/min. Column operation was monitored by absorbance at 210 nm. The peak corresponding to the target protein

was collected and stored at -70°C. All proteins were at least 95% pure by HPLC analysis.

Mass spectrometry. Electrospray ionization mass spectrometry of the HPLC purified proteins from all constructs showed molecular ions consistent with the calculated values.

Protein determination. Protein content was determined by the Bradford method (6) using bovine serum albumin as the standard (Pierce, Rockford, IL).

SDS-PAGE analysis and immunoblotting. The purified samples were analyzed on 16.5% Tris-Tricine gels (Bio-Rad, Hercules, CA) by sodium dodecyl sulfate/polyacrylamide gel electrophoresis (SDS-PAGE) under reducing conditions. For staining with Coomassie-Blue, 400-2700 ng of the purified samples was loaded. For immunoblotting, 80-540 ng of the purified samples was loaded and transferred to nitrocellulose membranes (Schleicher & Schuell, Keene, NH) following SDS-PAGE, according to conventional procedures. In separate experiments, blotting was performed with a monoclonal antibody against the FLAG peptide (International Biotechnologies, Inc., New Haven, CT) or with polyclonal antibodies to CV-N (1). The respective detection antibodies were conjugated with horseradish peroxidase (Sigma, St. Louis, MO). The SDS-PAGE profile of all of the final products indicated a high degree of purity and the expected molecular masses (around 11~12 kDa); all of the purified proteins were immunoreactive with the anti-CV-N antibodies; the FLAG-con-

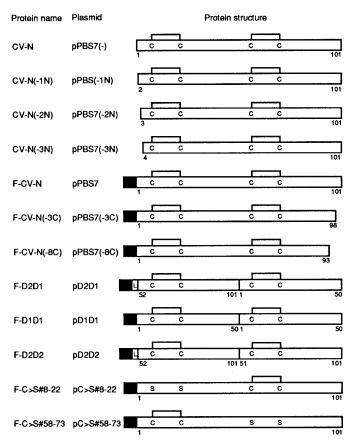


FIG. 2. Schematic diagram of CV-N mutants encoded by the respective plasmid constructs. Numbers represent amino acid positions that define spans within the intact CV-N-encoding sequence. Cys or Ser indicated in each mutant are located at positions 8, 22, 58 and/or 73 of the CV-N sequence. Closed and open boxes indicate FLAG octopeptides and CV-N, respectively. L, Leucine.

TABLE 1
Anti-HIV Activities of CV-N Mutants and F-CV-N Mutants

Protein	EC ₅₀ (nM) ^a
CV-N mutants	
CV-N control	0.9 ± 0.4
CV-N(-1N)	2.0 ± 0.7
CV-N(-2N)	8.3 ± 4.0
CV-N(-3N)	140.7 ± 36.7
F-CV-N mutants	
F-CV-N control	3.6 ± 1.4
F-CV-N(-3C)	149.5 ± 17.2
F-CV-N(-8C)	Inactive b
F-D2D1	217.7 ± 9.5
F-D1D1	Inactive c
F-D2D2	Inactive b
F-C>S#8-22	Inactive b
F-C>S#58-73	Inactive b

^a Values are means \pm S.D. of at least triplicate determinations.

taining mutants also reacted with anti-FLAG antibody (data not shown). Consistent with the previous results (1), both the control CV-N and the F-CV-N potently inhibited HIV-1 (EC $_{50}$'s 0.9 and 3.6 mM, respectively) (Table 1).

Anti-HIV assays. For comparisons of the anti-HIV activities of the recombinant proteins, the XTT assay was employed, as described by Gulakowski, et al (7).

Studies of interactions of the CV-N mutants with gp120. An enzyme-linked immunosorbent assay (ELISA) was used to determine the relative binding of the CV-N mutants (16 μ g/well) to gp120 (100 μ g/well), as described previously for CV-N (1).

RESULTS

Effect of deletion of N-terminal and C-terminal amino acids on the anti-HIV activity of CV-N. To study whether the N- and C-terminus of CV-N were necessary for anti-HIV activity, a series of deletion mutants was constructed in which 1, 2 or 3 amino acids were deleted from the N-terminus of CV-N ((CV-N(-1N), CV-N(-2N) or CV-N(-1N)], or 3 or 8 amino acids were deleted from the C-terminus of F-CV-N [F-CV-N(-3C) or F-CV-N(-8C)] (Fig. 2). Anti-HIV activity comparable to the full-length CV-N was retained in the mutant with deletion of 1 N-terminal amino acid [CV-N(-1N)] (Table 1). Deletion of 2 N-terminal amino acids [CV-N(-2N)] resulted in about a 9-fold decrease in anti-HIV potency, while deletion of 3 N-terminal amino acids [CV-N(-3N)] resulted in about a 156-fold decrease in potency (EC₅₀'s 8.3 and 140.7, respectively; Table 1) relative to CV-N. Deletion from F-CV-N of 3 C-terminal amino acids [F-CV-N(-3C)] resulted in approximately a 42-fold decrease in potency (EC₅₀ 149.5 nM) relative to F-CV-N, while deletion of 8 C-terminal amino acids [F-CV-N-(-8C)] completely eliminated activity (Table 1).

 $[^]b$ No anti-HIV activity was detectable at the highest tested concentration of 200 nM.

^c No anti-HIV activity was detectable at the highest tested concentration of 1000 nM.

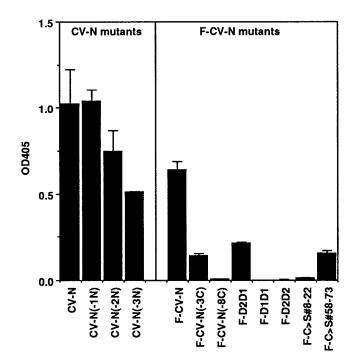


FIG. 3. ELISA assay of relative binding of CV-N, F-CV-N and mutants to gp120. All data are corrected for background antibody absorption in the absence of captured protein (typically $< 0.1 \ OD_{405}$). Results are the average OD405 (\pm S.D.) from triplicate wells.

Effect of circular permutation of two highly homologous regions on the anti-HIV activity of CV-N. To investigate the effects of circular permutation of the two highly homologous regions (first 50 amino acids (Domain 1) and last 51 (Domain 2)) (Fig. 1) or for tandem connection of identical domains, the F-CV-N mutants F-D2D1, F-D1D1 and F-D2D2 were constructed (Fig, 2). The F-D1D1 and F-D2D2 mutants were essentially devoid of anti-HIV activity (Table 1), however, some anti-HIV activity, albeit at greatly diminished potency, (EC $_{50}$ =217.7 nM) was retained in the D2D1 mutant (Table 1).

Effect of deletion of disulfide bonds on the anti-HIV activity of CV-N. To determine the relative importance of each of the two disulfide bonds within CV-N for the bioactivity, we constructed FLAG-mutants in which either the first two or last two cysteines were changed to serines to eliminate opportunity for the respective disulfide bonds (F-C>S#8-22 and F-C>S#58-73) (Fig. 2). Both of these mutants were devoid of anti-HIV activity (Table 1).

Effects on gp120 binding. The gp120-binding of the N-terminal deletion mutant CV-N(-1N) was essentially the same as that of the intact CV-N (see Fig. 3). In contrast, the gp120-binding of the N-terminal deletion mutants CV-N(-2N) and CV-N(-3N) was approximately 73% and 50%, respectively, relative to CV-N. The F-CV-N control showed gp120-binding activity of approxi-

mately 62% of the non-FLAG CV-N. Relative to F-CV-N, the gp120-binding affinity of the C-terminal deletion mutant F-CV-N(-3C) was about 22%, whereas that of the other C-terminal deletion mutant F-CV-N(-8C) was essentially nil. The circularly permuted mutant F-D2D1 showed modest gp120-binding affinity; however, neither of the tandem mutants F-D1D1 or F-D2D2 showed any appreciable gp120-binding. The disulfide linkage deletion mutant F-C>S#8-22 showed little or no detectable gp120-binding. In contrast, the other disulfide linkage deletion mutant F-C>S#58-73 showed modest gp120-binding, roughly comparable to that of the F-CV-N(-3C) mutant.

DISCUSSION

In the present study, we have compared the gp120-binding and anti-HIV activities of CV-N, F-CV-N and a series of corresponding mutants which were truncated at the N- or C-terminus, missing a disulfide bond, circularly permutated in two highly homologous domains, or tandemly duplicated in the latter domains. Addition of the FLAG-octapeptide leader sequence to CV-N resulted in a modest decrease both in gp120 binding and in anti-HIV activity. Nonetheless, the F-CV-N retained sufficient potency to serve as an appropriate control for additional CV-N mutants in which the FLAG peptide was incorporated to facilitate purification.

Deletion of the first amino acid (Leu) of the N-terminus of CV-N had relatively little effect on gp120 binding or anti-HIV activity. However, deletion of 2 N-terminal amino acids (Leu-Gly) resulted in a more substantial decrease of gp120 binding and a modest decrease in anti-HIV potency. Deletion of 3 N-terminal amino acids (Leu-Gly-Lys) resulted in a very pronounced decrease in gp120 binding and anti-HIV activity.

Similar to the N-terminus, the C-terminus of the active protein was also very sensitive to deletions. For example, the F-CV-N (-3C) mutant showed only weak gp120 binding and anti-HIV activity, and the F-CV-N (-8C) mutant neither bound gp120 nor gave detectable anti-HIV activity.

The circularly permuted protein, D2D1, showed some, albeit relatively weak, anti-HIV and gp120-binding activities. In contrast, neither of the tandem mutants, D1D1 or D2D2 showed detectable gp120 binding or anti-HIV activity.

Mutants F-C>S#8-22 and F-C>S#58-73, in which cysteine residues were replaced with serines, thereby lacking either one of the two disulfide linkages normally present in CV-N or F-CV-N, were inactive against HIV. This result is consistent with other observations that, when the cysteine residues of CV-N were reduced and alkylated with vinyl pyridine, the antiviral activity was abolished (2). Interestingly, whereas the inactive F-C>S#8-22 mutant showed little or no

binding to gp120, the other mutant F-C>S#58-73, although also devoid of antiviral activity against intact HIV, did show modest binding to gp120.

The present results, therefore, are consistent with earlier conclusions (1) that gp120 is a molecular target of CV-N. There was an overall very high correlation (r² > 0.9) of the relative gp120 binding activities of CV-N, F-CV-N and the various mutants thereof, with their corresponding anti-HIV potencies. None of the mutants that showed anti-HIV activity failed to bind gp120. However, the F-C>S#58-73 mutant (which lacked one of the internal disulfide linkages) was inactive against HIV yet retained modest gp120 binding affinity. Thus, while the binding of gp120 may be essential, it is not necessarily a sufficient requirement for anti-HIV activity of CV-N and related proteins; the sequence specificity for gp120 binding and anti-HIV activity do not appear to be identical.

ACKNOWLEDGMENTS

We thank Barry R. O'Keefe for consultation on the ELISA studies and Ricardo Bernal for technical support.

REFERENCES

- Boyd, M. R., Gustafson, K. R., McMahon, J. B., Shoemaker, R. H., O'Keefe, B. R., Mori, T., Gulakowski, R. J., Wu, L., Rivera, M. I., Laurencot, C. M., Currens, M. J., Cardellina, J. H., II, Buckheit, R. W., Jr., Nara, P. L., Pannell, L. K., Sowder, R. C., II, and Henderson, L. E. (1997) Antimicrob. Agents Chemother. 41, 1521–1530.
- Gustafson, K. R., Sowder, R. C., Henderson, L. E., Cardellina, J. H., II, McMahon, J. B., Rajamani, U., Pannell, L. K., and Boyd, M. R. (1997) Submitted for publication.
- Daar, E., Li, W. L., Moudgli, T., and Ho, D. (1990) Proc. Natl. Acad. Sci. USA 87, 6574-6578.
- Moore, J. P., Mckeating, J. A., Huang, Y., Ashkenazi, A., and Ho, D. O. (1992) J. Virol. 66, 235–243.
- Orloff, S. L., Bandea, C. L., Kennedy, M. S., Allaway, G. P., Maddon, P. J., and Mcdougal, J. S. (1995) *AIDS Res. Hum. Retrovir.* 11, 335–342.
- 6. Bradford, M. M. (1976) Anal. Biochem. 72, 248-254.
- Gulakowski, R. J., McMahon, J. B., Staley, P. G., Moran, R. A., and Boyd, M. R. (1991) J. Virol. Methods 33, 87-100.
- Hopp, T. P., Prickett, K. S., Price, V., Libby, R. T., March, C. J., Cerretii, P., Urdal, D. L., and Conlon, P. J. (1988) *Biotechnology* 6, 1205–1210.
- Movra, N. R., Nakamura, K., and Inouye, M. (1980) J. Biol. Chem. 255, 27–29.