Vol. 164, No. 1, 1989 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
October 16, 1989 Pages 30-38

SUBSTITUTION MUTATIONS OF THE HIGHLY CONSERVED ARGININE 87 OF HIV-1 PROTEASE RESULT IN LOSS OF PROTEOLYTIC ACTIVITY

John M. Louis[+], C. A. Dale Smith[+], Ewald M. Wondrak[*], Peter T. Mora[+], and Stephen Oroszlan[*]

[+]Division of Cancer Biology and Diagnosis, National Cancer Institute, N.I.H., Bethesda, MD 20892

[*]Laboratory of Molecular Virology and Carcinogenesis, BRI-Basic Research Program, NCI-Frederick Cancer Research Facility, Frederick, MD 21701

Received August 10, 1989

The 297bp gene coding for the HIV-1 protease was chemically synthesized and expressed in *E. coli* (1). Single amino acid substitutions (Arg 87 -> Lys; Arg 87 -> Glu) were introduced in the C-terminally located conserved region GlyArgAsn of the protease gene in the wild-type clone. The products of the mutant and the wild-type clones were expressed at approximately similar levels at 30 minutes of induction but the mutant protease proteins accumulated as a function of time of induction unlike the wild-type protease which declined after 60 minutes. The mutants were completely devoid of proteolytic activity as determined in assays employing as substrates a synthetic nonapeptide and a *gag* related recombinant polyprotein.

The human immunodeficiency virus type-1 (HIV-1) (2-4) encodes a protease (PR) which is responsible for the specific cleavage of the polyproteins, $Pr55^{gag}$ and $Pr160^{gag\text{-}pol}$, into the mature *gag* derived structural proteins p17 (MA), p24 (CA), p7 (NC) and p6 and the pol derived protease (PR), reverse transcriptase (RT) and the endonuclease, the integration protein (IN) (5-10). Protease deficient mutants of HIV were shown to produce immature noninfectious virus particles (11).

Recent studies on retroviral proteases confirmed earlier predictions that they are dimers (12), forming a symmetrical structure and thereby becoming similar to the semi-symmetrical cellular aspartic proteases (13, 14). The dimeric structure of HIV-1 protease has a characteristic aspartic proteinase-like active site formed from two AspThrGly sequences (amino acids 25-27). In addition to the AspThrGly sequence, retroviral proteases have a second highly conserved region GlyArgAsp/Asn (amino acids 86-88 in the HIV-1 located in the C terminal domain) that is unique to the viral enzymes and is not present in cellular aspartic proteinases (9, 12).

For this reason, we have conducted a mutational analysis of the Arg 87 and the Asn 88 in the HIV-1 protease. Our approach was to use a synthetic gene construct and to

0006-291X/89 $1.50

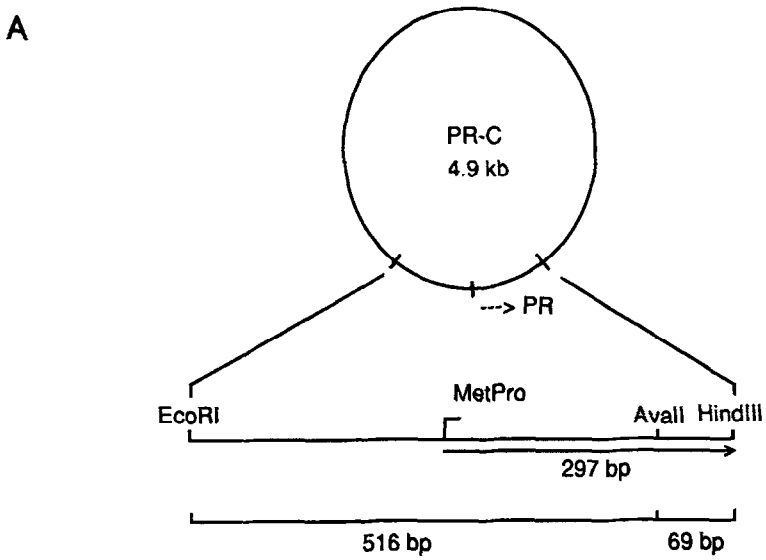
introduce substitution mutations replacing Arg -> Lys; Arg -> Glu; Asn -> Glu. Although the mutant proteins were expressed to a significantly higher level than that obtained with the wild-type protease in *E. coli*, all the substitutions resulted in the loss of proteolytic activity.

MATERIALS AND METHODS

Construction of mutant protease genes: The chemical synthesis of the exact coding sequence of the HIV-1 protease, its cloning and expression in *E.coli* has been described before (1). Plasmid DNA from the wild type clone PR-C (4.9kb), which had the correct coding sequence of 297bp, was digested with EcoRI and Hind3. The fragments derived (Figure 1A), which are of sizes 4.3kb (fragment 1) and 0.585kb were purified by preparative agarose gel electrophoresis (1, 15, 16). The 585bp fragment contains the entire protease coding sequence and in addition 288bp located upstream of the ATG, derived from the plasmid. The 585bp fragment was further digested with AvaII to obtain fragments of 516bp (fragment 2) and 69bp (fragment 3). The 69bp fragment contains the sequence coding for the second conserved domain Gly-Arg-Asn. Fragments to replace this 69bp (fragment 3) were synthesized (Applied Biosystems, Model 380B) with the appropriate base changes corresponding to the desired mutations (see Mut3-11, Mut4-1 and Mut5-10 in Figure 1B). Base overhangs for an AvaII site at the 5′ end and a TER/Hind3 towards the 3′ end were provided for selective ligation. Appropriate complementary fragments were annealed and ligated with equimolar concentrations of fragment 1 and fragment 2 under conditions described before (1). A control ligation was performed with fragments 1, 2, and 3 (PR-C+). The constructs were then used to transform *E. coli* JM105 (1, 17), and twenty colonies from each transformation were screened by immunoblotting, using protease specific antisera (1). Plasmid DNA was isolated from random positive colonies expressing the correct size product and used for sequencing the entire coding sequence of 297bp by the dideoxy method (15, 16). Figure 1C shows a portion of the sequence carrying the substitutions and confirmed that the respective constructs have the intended coding sequences that are shown in Figure 1B. Two representative clones corresponding to each one of the three substitution mutations shown above were used for studies on expression and activity assays.

Cell growth and detection of the expressed product: The wild type clones (PR-C and PR-C+) and the mutant clones were grown at 37°C in LB (Luria-Bertani) medium to a density of 0.4 to 0.42 OD (600nm) and then induced with IPTG (isopropylbeta-D-thiogalactopyranoside) for the desired time. The cells were collected and a solublized extract was prepared as before (1). The extract was centrifuged at 45K rpm (Beckmann 50Ti) for 1h and the supernatant was used for further analysis. Cell extracts containing equal amounts of protein (50 or 75μg) were electrophoresed on 5-12.5% linear gradient sodium dodecyl sulfate (SDS) polyacrylamide gels. The proteins were electrophoretically transferred (18) onto nitrocellulose paper (0.22μ, Schleicher & Schuell) and probed using antisera raised against the chemically synthesized HIV-1 protease (1).

Assay of proteolytic activity: Both the synthetic peptide and the *gag* related polyprotein were used as substrates. The nonapeptide, Val-Ser-Gln-Asn-Tyr-Pro-Ile-Val-Gln-amide (1, 9) was incubated with appropriate aliquots of the cell extracts (see legend to Figure 3A). The levels of substrate and cleavage products were determined using reversed-phase high pressure liquid chromatography (RP-HPLC) as described before (1). Peak areas were integrated to determine the amounts of substrate and cleavage products. When purified *E. coli* expressed GAG4 fusion protein [obtained from J. Rusche, Repligen Corp. (PvuII-BglII fragment of HIV-1: 317 amino acids)] was used as the substrate, 5 μl of bacterial lysates (protein concentration adjusted to 5.5mg/ml) were added to 5ul of 2X reaction buffer (1x buffer contained 50mM potassium phosphate, pH 5.5, 100mM sodium chloride, 5mM



B **69bp FRAGMENTS:**

PR-C+: 5'AvaII---TAATTGGAAGAAATCTGTTG---TER/HindIII3'
GlyArgAsn

Mut3-11: ---TAATTGGAAaAAATCTGTTG---
Lys

Mut4-1: ---TAATTGGAggAAATCTGTTG---
Glu

Mut5-10: ---TAATTGGAAGAgAaCTGTTG---
Glu

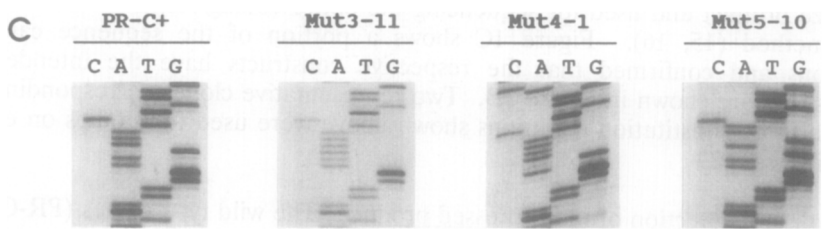


FIGURE 1: Strategy for the construction of mutated protease genes for expression. PR-C denotes the original wild-type clone (1). 1A shows the plasmid construct in clone PR-C which contains the 297bp coding sequence for the protease gene in the expression vector PKK233-2 (Pharmacia). 1B shows the wild-type sequence in PR-C+ and also the intended sequences in clones Mut3-11, Mut4-1 and Mut5-10 that were synthesized with replaced bases (shown in lower case) corresponding to the mutations used in this study. The constructs bearing the wild-type and mutant protease genes were sequenced by the dideoxy method (16). In 1C, PR-C+, Mut3-11, Mut4-1 and Mut5-10 show the region of interest to correspond exactly to the wild-type nucleotide sequence and to the intended substitution mutations shown between dotted lines in 1B.

dithiothreitol, 5mM EDTA and 10μg/ml aprotinin) containing the purified GAG4. These reaction mixtures were incubated at 37°C for various time periods (legend to Figure 3B) and terminated by addition of 0.2 volume 5X Laemmli buffer. The samples were subjected to 10-20% gradient SDS-polyacrylamide gel electrophoresis, transferred to nitrocellulose (18) and immunoblotted with specific rabbit antisera to p24 of HIV-1 (kindly provided by L. Arthur and J. Bess, Biological Products Lab., Program Resources Inc., NCI-FCRF).

RESULTS AND DISCUSSION

The following mutants were constructed: Mut3-11 (Arg87 -> Lys), Mut4-1 (Arg87 -> Glu), and Mut5-10 (Asn88 -> Glu). The wild type and the mutant clones were grown, induced with 1mM IPTG for 30 min, lysed and subjected to immunoblot analysis as described in Materials and Methods. As seen in Figure 2A, all the clones expressed a single unfused product of 11.5 kDa in size. This compared well with the product size of the previously reported synthetic protease gene product (1), the chemically synthesized HIV-1 protease (9, 19), and the protease isolated from the virions (8). The products expressed in the mutants were identical in size to the wild-type product and approximately equal amounts were observed after 30 min of induction. Figure 2B shows an immunoblot of the gene products of the wild-type and mutant clones as a function of time, using 1mM IPTG. In a previous report using the wild-type clone PR-C, we observed that the expressed protease accumulated to a maximum at 30 min of induction, after which the level rapidly declined (1). A similar pattern of accumulation of the protease product was observed with the reconstructed wild-type clone PR-C+, with a maximum at 45 to 60 min of induction. In contrast, all the mutant clones showed an accumulation of product as a function of time

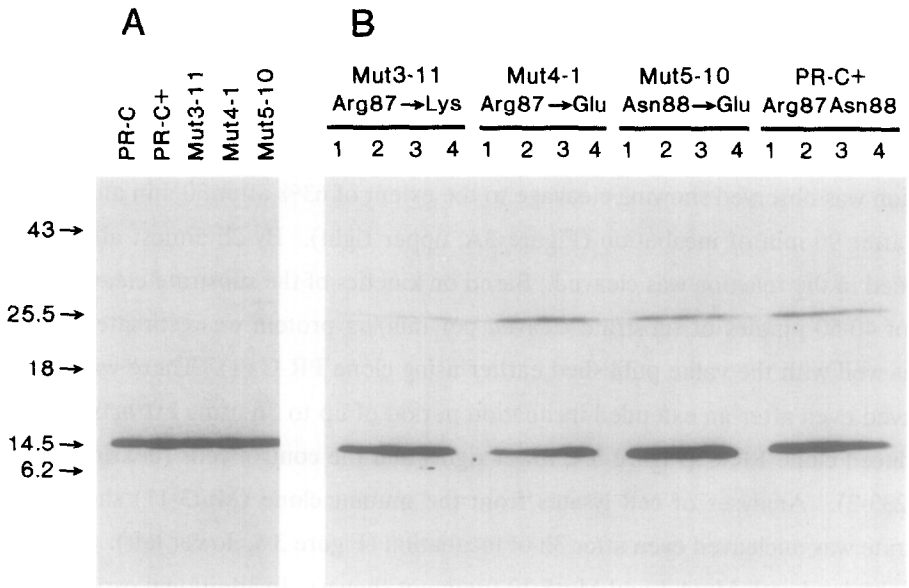


FIGURE 2: Expression of the cloned wild-type and mutated protease genes in *E. coli*. Bacteria bearing the corresponding plasmid construct were grown, and the total cell lysates were electrophoresed and immunoblotted as described in Materials and Methods. Figure 2A shows the comparison of the amounts of product expressed in the wild-type (PR-C and PR-C+) and the mutant clones (Mut3-11, Mut4-1 and Mut5-10) after 30 min of induction using 1mM IPTG. Figure 2B shows induction of the gene product as a function of time: time in min 0, 30, 60, 120 is denoted as 1 to 4 respectively. The migration of molecular weight markers (in kDa) are shown on the left.

up to 2h. The amount of product was two to three fold higher at 2h than after 30 min of induction. Duplicate mutant clones analyzed, carrying the same base changes, showed the same pattern of induction. In addition, *E. coli* transformed with the mutant constructs maintained the same level of growth as that of the control cells (bearing just the plasmid, PKK233-2) in the presence of added IPTG. The growth of the wild-type clone was inhibited in the presence of inducer. The short half-life of the active enzyme may be due to self-degradation, or degradation caused by *E. coli* proteases. However, we could not detect smaller size products using the antiserum raised against the synthetic protease.

The specific proteolytic activity of the expressed products derived from wild-type clones PR-C and PR-C+, and the mutant clones Mut3-11, Mut4-1, and Mut5-10 were analyzed by using a synthetic peptide as well as a *gag* polyprotein substrate. Equal aliquots of cell lysates which had similar amounts of expressed protein (estimated by immunoblotting and densitometric scanning) were used in this assay.

The synthetic nonapeptide, Val-Ser-Gln-Asn-Tyr-Pro-Ile-Val-Gln-amide spanning the HIV-1 p17-p24 cleavage site (Tyr-Pro) in the natural *gag* polyprotein was used in quantitative assays as described previously (1, 9). The samples were incubated at 37°C, terminated at the desired time and analyzed by RP-HPLC as described in legend to Figure 3A. Extracts from wild-type clone PR-C+, after an incubation period of 15 min, showed cleavage of the substrate to the extent of 52% (Figure 3A, upper left). As determined by amino acid analysis, the cleavage products, the tetrapeptide Pro-Ile-Val-Gln-amide (P_1) and the pentapeptide Val-Ser-Gln-Asn-Tyr (P_2) corresponded to the expected tyrosylproline cleavage. Furthermore, the recovered products P_1 and P_2 co-eluted with the independently synthesized appropriate tetra- and pentapeptides respectively. A progression in the enzyme reaction was observed showing cleavage to the extent of 63% after 30 min and greater than 85% after 90 min of incubation (Figure 3A, upper right). By 2h almost all the substrate provided in the reaction was cleaved. Based on kinetics of the substrate cleavage, an initial rate of 40-50 pmoles of substrate cleaved per min/μg protein was estimated. This value agrees well with the value published earlier using clone PR-C (1). There was no cleavage observed even after an extended incubation period of up to 3h using extracts derived from uninduced clone PR-C (Figure 3A, lower right) and the control cells (bearing the plasmid PKK233-2). Analyses of cell lysates from the mutant clone (Mut3-11) showed that the substrate was uncleaved even after 3h of incubation (Figure 3A, lower left). Similar results were obtained with Mut4-1 and Mut5-10 (data not shown). In additional experiments using extracts from uninduced clones PR-C and the mutants, there were no degradation products observed, and the substrate remained at the same level even after 6h of incubation under these assay conditions.

The recombinant protein GAG4 containing the p17-p24 and the p24-p15 cleavage sites (figure 3B) was also used to measure activity of the wild-type and mutant protease.

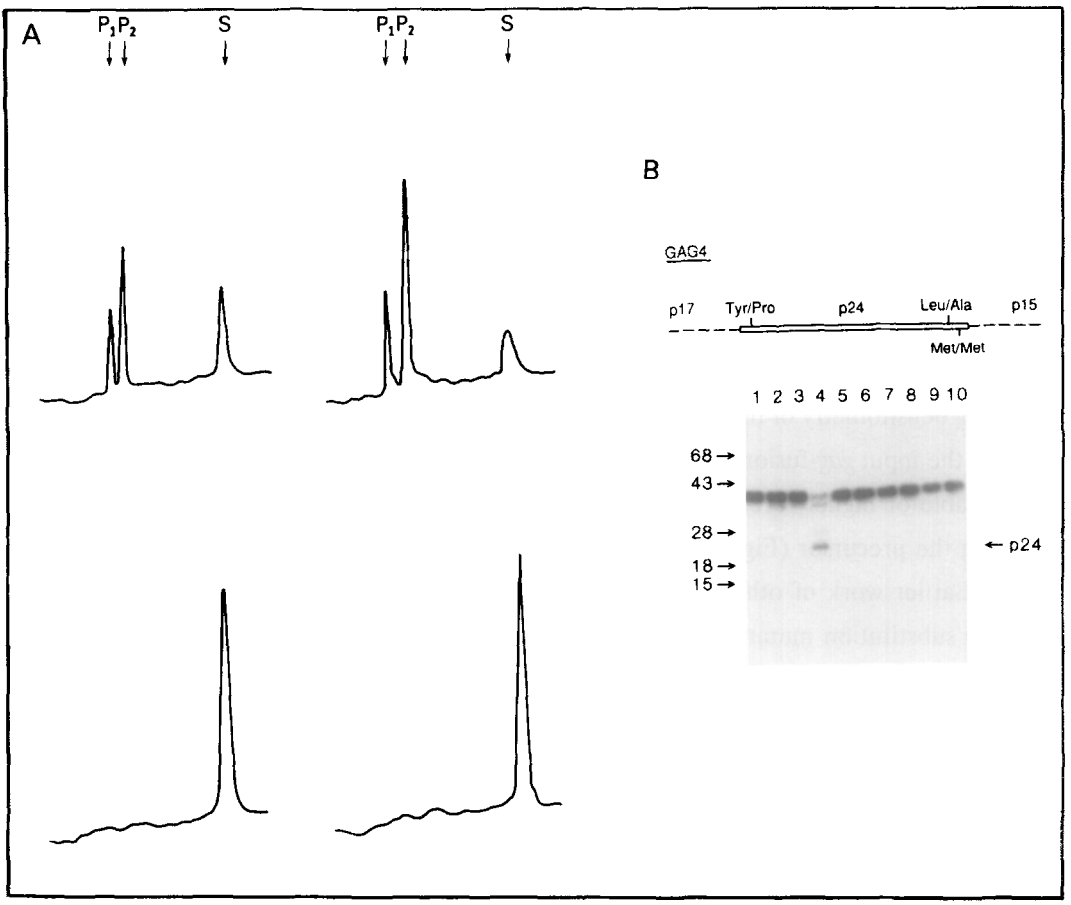


FIGURE 3: Specific proteolytic activity of the expressed wild-type and the mutated proteases determined using the synthetic nonapeptide and the HIV-1 *gag* polyprotein as substrates. [A] Protease assays in a total volume of 10μl were carried out with 2μl of bacterial lysate containing 11.0μg of total protein, added to 3μl of substrate (4.35 nmoles) in double distilled water and to 5μl of 2X reaction buffer, providing a final concentration of 0.25 M potassium phosphate, pH 7.0, 5% glycerol (V/V), 5mM dithiothreitol, 1mM EDTA and 1M sodium chloride, and incubated at 37°C for the required time. The substrate and the cleavage products were separated by RP-HPLC as described (1). Upper left and right of Figure 3A show the RP-HPLC profiles of reactions incubated for a period of 15 and 30 min respectively using extracts from the induced wild type clone PR-C+. Lower left and right profiles were samples incubated for 3h with extracts of induced clones Mut3-11 and uninduced PR-C respectively. S denotes the substrate, and P1 and P2 the cleavage products. No other major peaks were eluted using 0-20% acetonitrile gradient. [B] Various bacterial extracts were mixed with GAG4 precursor in 2X reaction buffer (see Materials and Methods) and incubated at 37°C. An example of an immunoblot of such an assay is shown. Lanes 1,3,5,7 and 9 show the control reactions (0 time) using extracts from uninduced PR-C and 30 min induced PR-C, mut3-11, mut4-1 and mut5-10, respectively. Lanes 2,4,6,8 and 10 show the same samples after an incubation period of 2h. The marker proteins (in kDa) and the position of p24 are indicated on the left and right respectively. At the top, the GAG4-*E. coli* fusion protein containing the protease cleavage sites at p17, p24, and p15 is shown, as discussed in the text. The boxed area represents the *gag* sequence.

Initial standardization experiments showed that the cleavage of the GAG4 precursor by the wild-type enzyme, was optimal at pH 5.5 in the presence of 100mM sodium chloride. Figure 3B shows an example of a typical immunoblot analysis using antiserum against the

p24 of HIV-1, after incubation of the GAG4 with the various bacterial cell lysates for 2h. As seen in lanes 1 and 2, cell lysates derived from uninduced wild-type clone PR-C were incapable of cleaving the GAG4 precursor, which migrates to a position in the gel corresponding to 41 kDa. Incubation of cell lysates derived from induced clone PR-C with the GAG4 resulted in the appearance of two specific cleavage products, a major 24 kDa band and minor 26 kDa band (Figure 3B: lane 4). It is assumed that the 24 kDa band is derived from GAG4 by cleavage at Tyr/Pro to generate the N terminus and Leu/Ala to generate the C terminus and corresponds to the mature p24 seen in the virion (8). The p26 may be a product resulting from cleavage at Tyr/Pro and Met/Met (see Figure 3B). By scanning densitometry of the autoradiograms and integration, we estimate that greater than 85% of the input *gag*-fusion protein was cleaved in 2h. The mutant cell extracts which had comparable or higher levels of expressed protease protein (Figure 2A) were incapable of cleaving the precursor (Figure 3B: lanes 6, 8, 10).

Earlier work of others on the mutational analysis of the HIV-1 protease were by random substitution mutations between residues 22 to 29 (20) and by specific changes of the conserved active site Asp 25 to Asn (21), to Thr (22) or to Ala (23). These mutations resulted in the elimination of the protease activity in all cases. It is also known that a deletion of 17 amino acids at the carboxyl terminus renders the enzyme inactive suggesting that this region containing the GlyArgAsn sequence is essential for protease function (24). Other mutations such as substitutions in residues 34-46 did not abolish activity (20). We introduced mutations specifically into the conserved GlyArgAsn domain (amino acids 86-88). As already pointed out, this region is conserved in all retroviral proteases but is absent in cellular aspartic proteinases. We chose to make a relatively minor change by changing the Arg 87 to Lys (maintaining a positive charge on the side chain) and a major change by replacing Arg87 with Glu (shortened and negatively charged side chain). Both these mutants did not show any specific proteolytic activity.

Our observations on such substitution mutational inactivation of the enzyme points out the importance of Arg87. This is in agreement with recent reports showing the possible interaction of Asp 29 with the highly conserved Arg 87 in forming the specific 3-D structure of the HIV-1 protease (25, 26). The Arg 87 which is conserved in the protease of all retroviruses, unlike in pepsins, interacts by ion pairing with the similarly highly conserved Asp 29. It should be pointed out that residue Asn 88 is unique to HIVs and in the related simian viruses as it is Asp in other retroviruses. It is possible that the mutation Asn 88 -> Glu alters the stereochemical structure of the dimeric protease by disturbing the essential Asp 29 to Arg 87 interaction. Further characterization of these mutant proteases described in this report may help in understanding the functional role of the GlyArgAsn structural element of HIV-1 protease and in the design of specific inhibitors capable of recognizing the viral protease but not the related pepsin-like proteases of the host.

ACKNOWLEDGMENTS

We thank J. Rushe for the GAG4 recombinant protein and J. Bess and L. Arthur for antiserum to HIV-1 p24. We also wish to thank S. Bricker and V. W. McFarland for excellent technical assistance, T. Copeland for the synthetic nonapeptide substrate, C. Hixson for amino acid analysis, D. Davies, E. Padlan, I. Weber and A. Wlodawer for useful discussions and A. Rabson for support. This work was sponsored in part by the National Cancer Institute, DHHS under contract number NO1-CO-74101 with BRI. EMW was supported by a Deutsche Forschungsgemeinschaft fellowship: WO 375/1-1.

REFERENCES

1. Louis M. J., Wondrak, E., Copeland, T., Smith, C. A. D., Mora, P. T. and Oroszlan, S. (1989) Biochem. Biophys. Res. Comm. **159**, 87-94
2. Barre-Sinoussi, F., Chermann, J. C., Rey, F., Nugeyre, M. T., Chamaret, S., Gruest, J., Dauguet, C., Axler-Blin, C., Vezinet-Brun, F., Rouzioux, C., Rozenbaum, W. and Montaganier, L. (1983) Science **220**, 868-870
3. Gallo, R. C., Salahuddin, S. Z., Popovic, M., Shearer, G. M., Kaplan, M., Haynes, B. F., Palker, T. J., Redfield, R., Oleske, J., Safai, B., White, G., Foster, P. and Markham, P. (1984) Science **224**, 500-502
4. Fauci, A. S. (1988) Science **239**, 617-622
5. Veronese, F.D., Copeland, T. D., DeVico, A. L., Rahman, R., Oroszlan, S., Gallo, R. C. and Sarngadharan, M. G. (1986) Science 231, 1289-1291
6. Lightfoote, M. M., Coligan, J. E., Folks, T. M., Fauci, A. S., Martin, M. A. and Venkatesan, S. (1986) J. Virol. **60**, 771-775
7. Veronese, F.D., Rahman, R., Copeland, T. D., Oroszlan, S., Gallo, R. C., and Sarngadharan, G. (1987) AIDS Res. Human Retroviruses **3**, 253-264
8. Henderson, L. E., Copeland, T. D., Sowder, R. C., Schultz, A. M. and Oroszlan, S. (1988) in Human Retrovirusus, Cancer and AIDS: Approaches to Prevention and Therapy. (ed. D. Bolognesi) pp. 135-147, Alan R. Liss, NY, NY
9. Copeland, T. D. and Oroszlan, S. (1988) Gene Anal. Tech. **5**, 109-115
10. Leis, J., Baltimore, D., Bishop, J. M., Coffin, J., Fleissner, E., Goff, S. P., Oroszlan, S., Robinson, H., Skalka, A. M., Temin, H. M., and Vogt, V. (1988) J. Virol. **62**, 1808-1809
11. Kohl, N. E., Emini, E. A., Schleif, W. A., Davis, L. J., Heimbach, J. C., Dixon, A. F., Scolnick, E. M., and Sigal, I. S. (1988) Proc. Natl. Acad. Sci. USA **85**, 4686-4690
12. Pearl, L. H. and Taylor, W. R. (1987) Nature **329**, 351-354
13. Miller, M., Jaskolski, M., Mohana Rao, J. K., Leis, J. and Wlodawer, A. (1989) Nature **337**, 576-579
14. Navia, M. A., Fitzgerald, P. M. D., McKeever, B. M., Chih-Tai Leu, Heimbach, J. C., Herber, W. K., Sigal, I. S., Darke, P. L. and Springer, J. P. (1989) Nature **337**, 615-620
15. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) Molecular Cloning : A Laboratory Manual. pp. 366-380, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
16. Berger, S.L., and Kimmel, A.R. (1987) in Guide to Molecular Cloning Techniques. Meth. Enzymol. **152**, pp. 94-562, Academic press, NY
17. Hanahan, D. (1983) J. Mol. Biol. **166**, 557-580
18. Burnette, W. N. (1981) Anal. Biochem. **112**, 195-203
19. Schneider, J. and Kent, S. B. H. (1988) Cell **54**, 363-368
20. Loeb, D. D., Hutchison III, C. A., Edgell, M. A., Farmerie, W. G. and Swanstrom, R. (1989) J. Virol. **63**, 111-121

21. Kohl, N. E., Emini, E. A., Schleif, W. A., Davis, L. J., Heimbach, J. C., Dixon, A. F., Scolnick, E. M., and Sigal, I. S. (1988) Proc. Natl. Acad. Sci. USA **85**, 4686-4690
22. Seelmeier, S., Schmidt, H., Turk, V., and von der Helm, K. (1988) Proc. Natl. Acad. Sci. USA **85**, 6612-6616
23. Mous, J., Heimer, E. P., and Le Grice, S. F. J. (1988) J. Virol. **62**, 1433 -1436
24. Hansen, J., Billich, S., Schulze, T., Sukrow, S., and Moelling, K. (1988) EMBO J. **7**, 1785-1791
25. Weber, I. T., Miller, M., Jaskolski, M., Leis, J., Skalka, A. M. and Wlodawer, A. (1989) Science **243**, 928-931
26. Blundell, T., and Pearl, L. H. (1989) Nature **337**, 596-597