

Cloning and Expression of Human Immunodeficiency Virus-1 Epitopes in *Escherichia coli*

Yoo, Hyang-Sook, Won-Hee Jang, Huei-Dong Park, Sang Won Hyun, Sang-Uk Nham and Young-Ik Lee*

Lab. of Molecular Genetics,

Genetic Engineering Research Institute,

Korea Institute of Science and Technology, Daejeon, Daeduk, Korea

대장균에서 사람의 면역결핍 바이러스-1 epitopes 클로닝과 발현에 대한 연구

유향숙 · 장원희 · 박희동 · 현상원 · 남상욱 · 이영익*

한국과학기술연구원, 유전공학연구소, 분자유전학실

ABSTRACT: Human immunodeficiency virus type 1 (HIV-1) causes a deadly infectious disease, Acquired Immunodeficiency Syndrome (AIDS). As a first step to develop a reliable and fast diagnostic procedure for HIV-1 infection, we cloned various immunodominant epitopes of HIV-1 in bacterial expression vectors containing *tac* or *trp* promoter. While the protein level of direct expression of gp160 was low, *trp* E fused gp120, gp41 and p17-p24 were produced at high levels (15-30% of total bacterial proteins) in *E. coli*. Since gp120 and gp41 contain relatively conserved regions which can react with antibodies in the plasma from most of HIV-1 infected individuals, these expression clones were used for large preparations of HIV-1 antigens.

KEY WORDS □ HIV-1 epitopes/*tac* and *trp* expression vector/immunodominant epitopes/hydrophobic region

Acquired immunodeficiency syndrome (AIDS) is a serious infectious disease which is prevalent in all over the world. This disease is mostly caused by Human Immunodeficiency Virus type 1 (HIV-1) which infects and kills a type of immune cells, T4 cells (Papovic *et al.*, 1984; Gallo *et al.*, 1984; Groopman *et al.*, 1984). The destruction of T4 cells leads to severe immune deficiency which in turn causes many opportunistic infections in AIDS patients (Gluckman *et al.*, 1986; Fauci, 1988). The genome of HIV provirus consists of 9749 bp and contains several genes including two long terminal repeats (LTR), core protein gene (GAG), RNA-dependent DNA polymerase gene (POL), envelope glycoprotein

encoding gene (ENV), trans-activating factor gene (TAT) and negative effector gene (NEF) (Muesing *et al.*, 1985; Ratner *et al.*, 1985).

The coat of this virus is consisted of env gene products which provoke an immune response in HIV-1 infected individuals in most cases. Env gene encodes a precursor peptide (gp160) which is later processed to an external envelope glycoprotein (gp120) and a transmembrane protein (gp41) (Veronese *et al.*, 1985; Chakrabarti *et al.*, 1986). While there are highly variable regions in the viral env gene products, C-terminal portion of gp120 and N-terminal region of gp41 are known to be quite antigenic and less variable (Starcich *et al.*, 1986). These proteins, therefore, have been used as antigens to detect HIV-1 antibody in HIV-1 infected individuals. Gag gene

*Correspondence should be addressed.

also produces p17, p24 and p15 protein through proteolytic cleavage of 512 amino acids precursor (Robey *et al.*, 1985; Ratner *et al.*, 1985). Although these proteins mostly exist in the inner core of the virus, they are able to elicit an immune response in HIV-1 infected persons. It has been shown that the immune response against HIV-1 begins before the depletion of T4 cells and the plasma levels of antibodies against each gene products changes according to the stages of HIV-1 infection (Seligmann *et al.*, 1984).

Methods to detect the infected individuals provide a vital key to curtail wide spread of the virus. Several diagnostic tests were already developed in a way to detect plasma antibody against HIV-1 (Chang *et al.*, 1985; Dowbenko *et al.*, 1985). However these methods including enzyme linked immunosorbent assay were time consuming and labor intensive. As a first step to develop the rapid and simple diagnostic procedure, we attempted to clone immunodominant regions of HIV-1 genes and expressed in *E. coli*. We report here successful expression of gp41 and p24 and gp120 in *E. coli* using various expression vectors under different promoters.

METHODS

Bacterial expression vectors

Expression vector pATH1, pATH2 and pATH3 were obtained from Dr. Thomas Cunningham at The University of Tennessee, U.S.A. pTTQ19 vector was from Amersham and pCT10 vector was a kind gift from Dr. Myeong-Hee Yu at Korea Institute of Science Technology, Genetic Engineering Research Institute.

Construction of HIV-1 expression clones.

HIV-1 gene containing vector pSP64-BH10 was obtained from R. Gallo at National Institute of Health, U.S.A. Various regions of the viral gene were digested with appropriate restriction enzymes and ligated into expression vectors indicated above. Plasmid DNA preparation and bacterial transformation were performed as described in Molecular Cloning (Sambrook *et al.*, 1989). SDS-Poly Acrylamide Gel Electrophoresis (SDS-PAGE) for identifying expression levels was performed as reported before (Laemmli, 1970).

Bacterial strains and media

E. coli strain HB101, RR1, and JM109 were used as recipients of plasmids containing HIV-1 genes. For the induction of trp promoter, bacterial

cultures were grown overnight in M9 media supplemented with 20 $\mu\text{g/ml}$ of L-tryptophan. M9 media contains 42 mM Na_2HPO_4 , 22 mM KH_2PO_4 , 8.6 mM NaCl 18.7 mM NH_4Cl , 0.05% Casamino acid, 0.1 mM CaCl, 1.0 mM MgSO_4 , 0.2% glucose, 10 $\mu\text{g/ml}$ thiamine and 20 $\mu\text{g/ml}$ ampicillin. The overnight culture was diluted 1:10 with fresh M9 medium without tryptophan and grown for 2 hours in intensive aeration. 3-beta-indoacrylic acid was added at the final concentration of 10 $\mu\text{g/ml}$ and the cultures were incubated 4 more hours before harvesting the cells.

Induction of tac promoter containing plasmid such as pCT10 and pTTQ19 vector was carried out by incubation at 37°C for 4 hours in the presence of 1.0 mM isopropyl beta-D-thiogalactoside (IPTG) as previously described (DeBore *et al.*, 1983).

RESULTS AND DISCUSSIONS

In order to obtain various epitopes of HIV-1 in large quantity, we constructed expression plasmids for transformation of *E. coli*. First we cloned various fragments of gp160 gene in direct expression vectors such as pTTQ19 or pCT10. Figure 1 shows strategies of construction for pYYM-env1. pYYM-env2. 2.7 kb Kpn I fragment corresponding to gp160 was inserted into pTTQ19 (pYYM-env1) and 2.1 kb Bam HI fragment was put into pCT10 vector (pYYM-env2). The transcription termination in these expression vectors was provided by the termination codons in rrnBt112 and lac I^q adjacent to the inserted HIV-1 genes. Bacterial cultures containing proper expression vectors were induced in the presence of 1.0 mM IPTG. Cell pellets were subjected to SDS-PAGE to ascertain the inducibility of each genes. As shown in Figure 2, pYYM-env1 and pYYM-env2 produced a little elevated level of 110 kD and 100 kD protein, respectively, however, the levels of expression were not sufficient enough to use as antigens.

As the second effort to get higher levels of expression, we constructed trp E fusion plasmids in which various sizes of env and gag gene were linked to 3'-end of trp E gene. Fig. 3 shows the scheme of pYYM-env4 construction which represents the generation of other plasmids containing different regions of env gene, pYYM-env5, pYYM-env6 and pYYM-env7 (Fig. 6B). The sizes

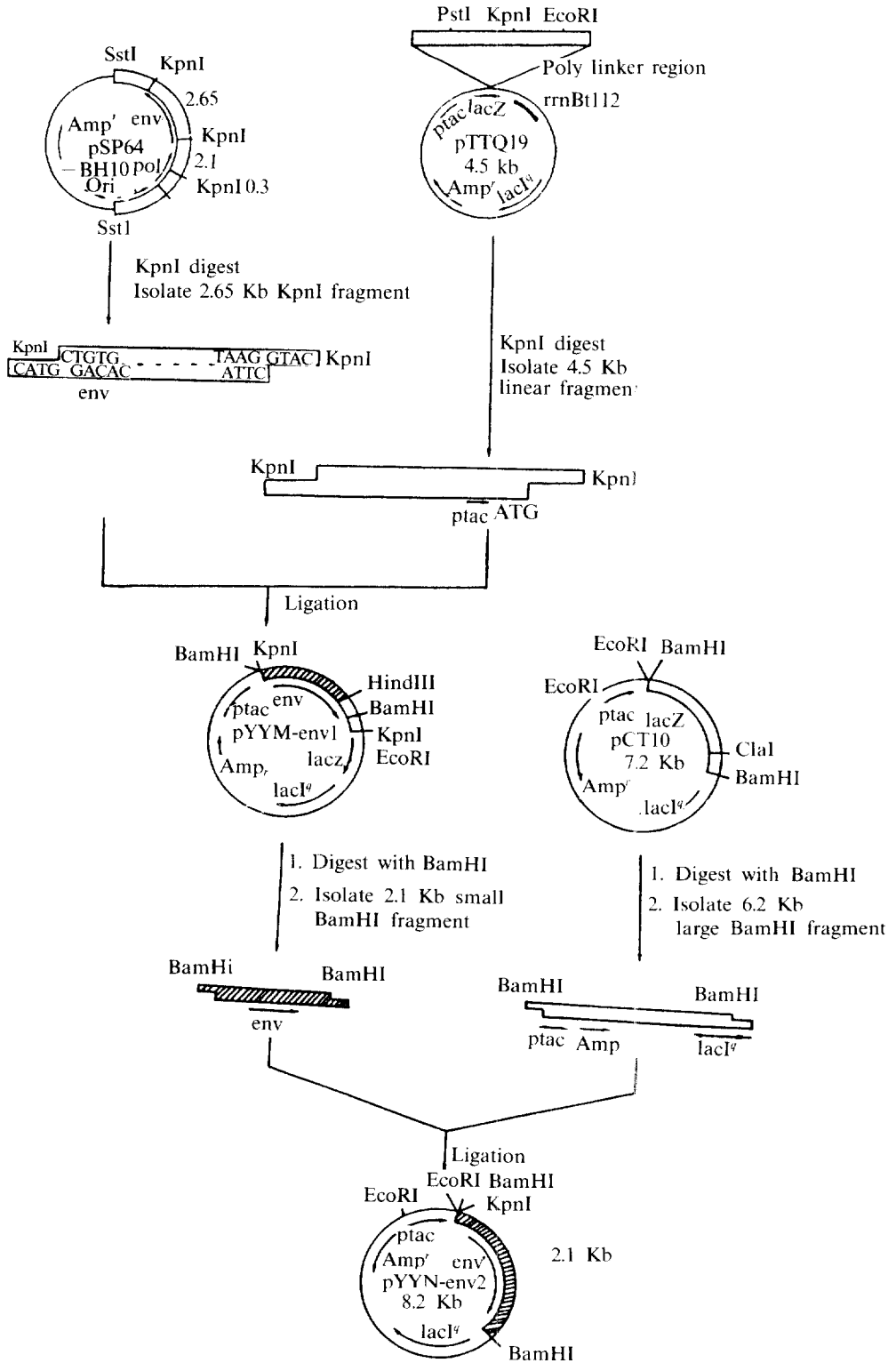


Fig. 1. The strategies of the plasmids pYYM-env1 and pYYM-env2 construction.

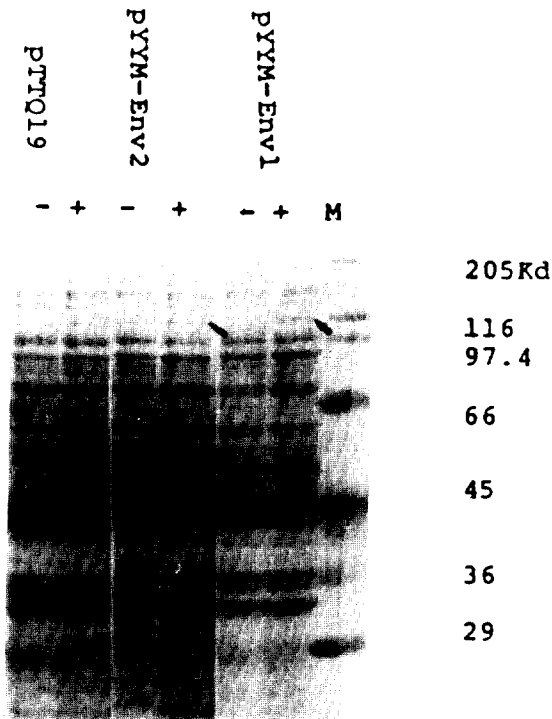


Fig. 2. SDS-PAGE analysis of expression of pYYM-env1 and pYYM-env2 plasmids in *E. coli* RR1 strain. Cells were incubated in the presence (+) or absence (-) of IPTG for 4 hours. Arrows point out expressed HIV-1 proteins.

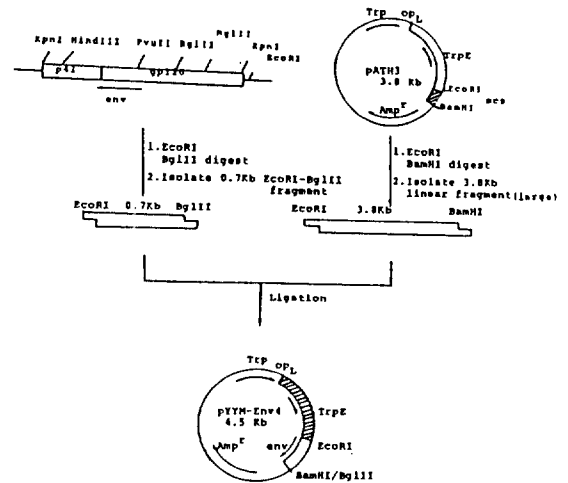


Fig. 3. The scheme of the plasmid pYYM-env4 construction. HIV-1 env gene was cut out from pSP64-BH10 by *EcoRI* and *BglIII* and ligated into pATH3 previously digested with *EcoRI* and *BamHI*.

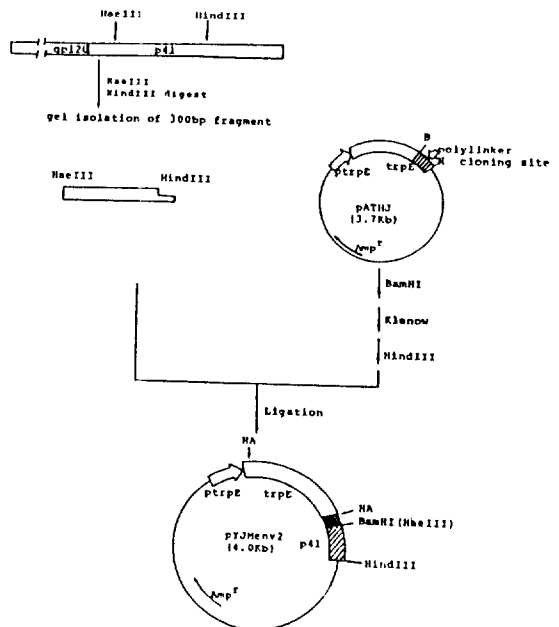


Fig. 4. A construction scheme of the plasmid pYJM-env2. pATHJ is a vector derived from pATH2 with a synthetic Asn-Gly linker.

and the enzyme sites of these plasmids are shown in Fig. 6B. In Fig. 4, gp41 encoding gene was inserted into a pATHJ vector which was originated from pATH2 and contained a hydroxylamine cleavage site in between trp promoter and a multicloning site. Different regions of gag gene (p17 or p17+p24) were isolated and ligated into the trp E vector pATH3 as shown in Fig. 5. 430 bp *Cla I* and *Rsa I* fragment of p17 gene was ligated with a synthetic oligomer containing Asn-Gly linker and 15 amino acids of p17 gene. Resulting plasmid pYYM-gag1 contains whole gag p17 gene at the 3'-end of trp E gene and the hydroxylamine cleavage site. Asn-Gly sequence in between 3'-end of Trp E and at the beginning of gag p17 initiation codon. The plasmid pYYM-gag2 containing p17 and p24 was constructed from the plasmid pYYM-gag1 and pSP64-BH10 (Fig. 5). Expression of fused genes was controlled by trp promoter in the pATH plasmid and cells containing these plasmids were incubated in

tryptophan free medium supplemented by tryptophan analogue (3-beta-indolacrylic acid). All the fusion proteins except the clones pYYM-env7 and pYYM-env6 were induced at high levels (Fig.

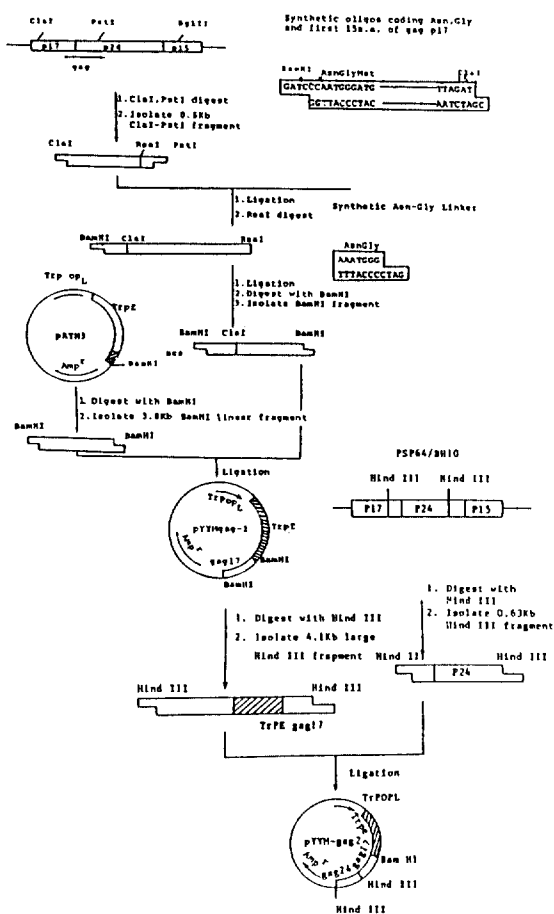


Fig. 5. Construction strategies of the plasmids pYYM-gag1 and pYYM-gag2.

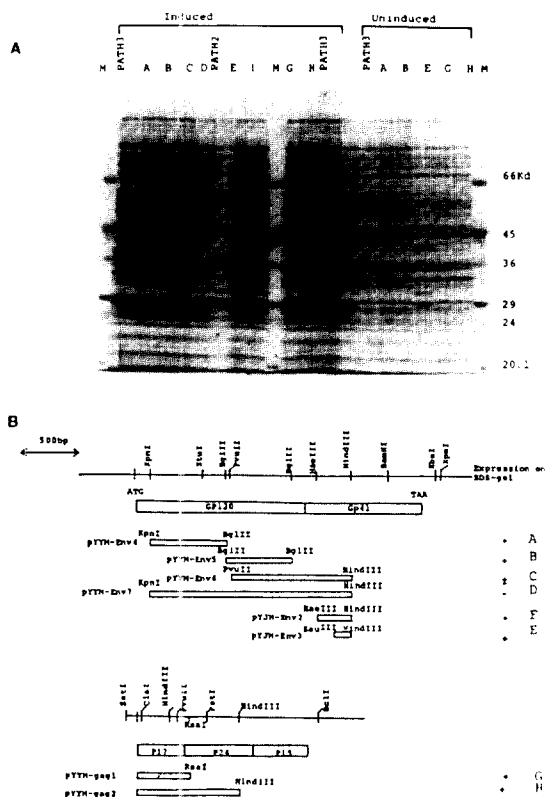


Fig. 6. Expression of several trp E fusion plasmids in JM109.

A panel: SDS-PAGE analysis of trp E fusion plasmids. A. pYYM-env4; B. pYYM-env5; C. pYYM-env6; D. pYYM-env7; E. pYJM-env3; F. pYJM-env3; G. pYYM-gag1; H. pYYM-gag2. B panel: subcloning of HIV-1 env and gag gene fragments into trp E fusion vector..

6B). The levels of expressed proteins were 15-30% of total bacterial proteins judged by densitometer scanning. The sizes of induced proteins appeared to be right judged from the size of insert DNA in the plasmids. Also these expressed fusion proteins were highly immunoreactive to AIDS patients serum, which indicates that these bacterial produced proteins contains HIV-1 antigen (data not shown). This demonstrates that the trp E fused genes are appropriate clones for further experiments to obtain large amount of HIV-1 proteins.

There are numerous reports to show high level expression of hybrid gene such as insulin A, B chain, human growth hormone and beta-globin in *E. coli* (Goeddel *et al.*, 1979; Nagai *et al.*, 1985; Szoka *et al.*, 1986). Heterologous fusion gene expression in *E. coli* is generally higher than direct

expression provably due to the stability of the gene product in bacterial cytoplasm and proper folding of the proteins. In addition to this advantage, fusion proteins can be produced to facilitate purification procedure (Marston, 1987). Bacterial fusion proteins, however, need to be cleaved. To overcome this problem, we inserted the synthetic oligonucleotide encoding Asn-Gly linker which provides a hydroxylamine cut site in between trp E and HIV-1 gene fragments. Thus we could purify HIV-1 antigen aside from Trp E protein. Preliminary study in our lab. showed that the fusion protein derived from the plasmid pYJM-env2 was successfully cleaved by the chemical treatment of hydroxylamine.

Numerous reports indicated that some genes

were hard to be expressed in *E. coli* (Marston, 1986). The lack of inducibility of the genes was explained by several factors such as a harmful sequence to bacteria and/or the inappropriate size of the gene to be expressed. Certain genes encoding very hydrophobic amino acid sequence may be toxic to the host cells after being expressed at high level. Also production of large size of heterologous gene products can be a big burden to the host bacteria, which will work as a negative force to the cells. These might be the reasons why the plasmid pYYM-env6 and pYYM-env7 were not expressed well.

N-terminus portion of gp41 is an external region of transmembrane proteins which binds to the gp120 with hydrophobic interaction (Kowalski *et al.*, 1987). Synthetic oligopeptide containing 12 amino acid (bp 598-609) which falls

in this region showed strong antibody reactivity due to its cyclic structure formed by a disulfide bond, suggesting immunodominant epitopes useful to diagnose HIV-1 infection (Gnann *et al.*, 1987). In addition, N-terminus portion of gp41 has few glycosylation sites, which is appropriate for *E. coli* expression. pYYM-env2 in this report is the plasmid that contains above mentioned N-terminal region of gp41, and will be used for further experiments.

In this paper we report that gp120, gp41 and p17-24 were produced at high levels in the forms of trp E fusion in *E. coli*. The levels of expression were 15-30 percent of total bacterial proteins, which allow us to use the bacterial clones as constant sources for mass production of those proteins.

적 요

Human immunodeficiency virus-1(HIV-1)은 후천성 면역결핍증(AIDS)을 일으키는 바이러스로서 이 바이러스의 감염여부를 진단하는 것은 이 전염병의 확산을 막는 중요한 방법의 하나이다. 보다 빠르고 확실성이 높은 AIDS 진단시약개발을 위하여 본 실험실에서는 HIV-1의 여러 epitopes를 대장균의 trp 또는 tac promoter가 있는 발현벡터에 넣어 클로닝하였고 이 클로닝된 플라스미드들을 발현 시켜보았다. gp120은 tac promoter에 직접 연결되어 발현되었는데 그 양은 많지 않았다. 하지만 gp120, gp41 그리고 p17-p24 등은 trp E 단백질과 융합된 형태로서 발현되는 양이 많았는데 전체 대장균 단백질량의 15%-30% 까지 이르렀다. gp120, gp41은 비교적 유전적 변이가 적은 표면항원을 포함하고 있고 또한 박테리아에서 발현 level이 상당히 높으므로 본 연구에서 제조한 이러한 플라스미드들은 AIDS 환자의 혈액에 존재하는 항체를 식별할 수 있는 항원을 다량 생산할 수 있는 공급원으로 유용히 사용되리라 본다.

REFERENCES

1. Chakrabarti, S., M. Robert-Guroff, F. Wong-Staal, R.C. Gallo, and B. Moss, 1986. *Nature*, **320**, 535-537.
2. Chang, T.W., I. Kato, S. McKinney, P. Chanda, A.D. Barone, F. Wong-Staal, R.C. Gallo, and N.T. Chang, 1985. *Biotechnology*, **3**, 985-990.
3. DeBoer, H.A., L.J. Comstock and M. Vasser 1983. *Proc. Natl. Acad. Sci. USA*, **80**, 21-25.
4. Dowbenko, D.J., J.R. Bell, C.V. Benton, J.E. Groopman, H. Nguyen, D. Vitterlein, D.J. Capon, and Lasky 1985. *Proc. Natl. Acad. Sci. U.S.A.*, **82**, 7748-7752.
5. Fauci, A.S. 1988. *Science*, **239**, 617-622.
6. Gallo, R.C., S.Z. Salahuddin, M. Popovic, G.M. Shearer, M. Kaplan, B.F. Haynes, T.J. Parker, R. Redfield, J. Oleske, B. Safai, F. White, P. Foster and P.D. Markham, 1984. *Science*, **224**, 500-503.
7. Gnann Jr. J.W., J.A. Nelson and M.B. Oldstone, 1987. *J. Virology*, **61**, 2639-2641.
8. Goeddel, D.V., D.G. Kleid, F. Bolivar, H.L. Heynecker, D.G. Yansur, R. Crea, T. Hirose, A. Kraszewski, K. Itakura, and A.D. Riggs, 1979. *Proc. Natl. Acad. Sci. U.S.A.*, **86**, 106-110.
9. Groopman, J.E., S.Z. Salahuddin, M.G. Sarngadharan, P.D. Markham, M. Gonda, A. Sliski, and R.C. Gallo 1984. *Science*, **226**, 447-449.
10. Kowalski, M., J. Potz, I. Basitipour, T. Dorfman, W.C. Goh, E. Terwilliger, A. Daytonk, C. Rosen, W. Haseltin and J. Sodroski, 1987. *Science*, **387**, 1351-1355.
11. Laemmli, U.K. 1970. *Nature*, **227**, 680-685.
12. Marston, F.A.O. 1986. *Biochemical J.*, **240**, 1-12.
13. Marston, F.A.O. 1987. in DNA Cloning Vol. III, A Practical Approach. E.D. Glover, D.M. IRL press, Oxford, England.
14. Muesing, M.A., D.H. Smith, C.D. Cabradilla, C.V. Benton, L.A. Lasky, and D.J. Capon 1985. *Nature*, **313**, 450-458.
15. Nagai, K. M.F. Perutz and C. Poyaart, 1985. *Proc. Natl. Acad. Sci. U.S.A.*, **82**, 7252-7255.
16. Ratner, L., W. Haseltine, R. Patarca, K.J. Livak, B. Starcich, S.F. Josephs, E.R. Doran, A. Rafalski, E.A. Whitehorn, K. Baumeister, L. Ivanoff, S.R. Petteway, M.L. Pearson, J.A. Lautenberger, T.S. Papas, J. Ghrayeb, N.T. Chang, R.C. Gallo, and F. Wong-Staal, 1985. *Nature*, **313**, 227-284.
17. Robey, W.G., B. Safai, S. Oroszlan, L.O. Arther, M.A. Gonda, R.C. Gallo, and P.J. Fischinger 1985.

- Science*, **228**, 593-595.
18. Popovic, M. M.G. Sarngadharan, E. Read, and R. C. Gallo, 1984. *Science*, **224**, 497-500.
 19. Sambrook, J., E.F., Fritsch and T. Maniatis, 1989. in *Molecular Cloning, A Laboratory Manual*, 2nd Eds., Cold Spring Harbor Laboratory, New York.
 20. Seligmann, M., L. Chess, J. Fahey, A. Fauci, P.J. Lachmann, J. L'Age-Stehr, J. Ngu, A.J. Pinching, F.S. Rosen, T.J. Spira, and J. Wybran, 1984. *New Eng. J. Med.*, **311**, 1286-1291.
 21. Starcich, B.R., B.H. Hahn, G.M. Shaw, P.D. Mcneely, S. Modrow, H. Wolf, E.S. Park, W.P. Parks, S.F. Josephs, R.C. Gallo, and F. Wong-Staal 1986. *Cell*, **45**, 637-648.
 22. Szoka, P.R., A.B. Scareiber, H. Chan and J. Murthy, 1986. *DNA*, **5**, 11-20.
 23. Veronese, F. A.L. DeVico, T.D. Copeland, S. Oroszlan, R.C. Gallo, and M.G. Sarngadharan, 1985. *Science*, **229**, 1402-1405.

(Received January 5, 1991)

(Accepted February 25, 1991)