

Human T-cell Leukemia Virus Type I (HTLV-I)의 Gag-Pro Transframe 단백질 정제를 위한 재조합 DNA의 제작

남 석 현

아주대학교 자연과학대학 생명과학과

HTLV-I의 *gag-pro* 유전자 중첩영역내에서 -1 ribosomal frameshifting이 일어나는 자리를 결정하기 위하여 *gag-pro* 중첩영역의 일부를 SP6 promoter를 가진 벡터내에 클로닝하였다. 그 결과 닭의 prelysozyme에서 유래한 5개의 아미노산을 코딩하는 합성유전자와 141 bp로된 *gag-pro* 중첩영역의 뒤에 *Staphylococcus aureus*의 protein A 유전자단편이 연결된 hybrid 유전자를 보유한 플라스미드를 제작하였다. 이 DNA 클론을 주형으로 SP6 RNA polymerase의 작용에 의해 한 종류의 mRNA를 다량으로 합성하였다. *In vitro*에서 합성된 mRNA로 무세포계에서 단백질을 합성한 결과 21 kDal의 단백질이 생성되었고 IgG-Sepharose를 사용한 affinity chromatography로 합성된 단백질을 순수하게 정제할 수 있었다. 본 연구에서 설명한 *in vitro* 실험계는 Gag-Pro transframe 단백질의 신속한 정제 및 일차구조의 결정에 유익하게 사용될 것으로 보이며 이와 같은 실험의 결과 mRNA에서 ribosomal frameshifting이 일어나는 정확한 site를 결정할 수 있을뿐 아니라 *pro* 유전자의 발현에 필요한 frameshift를 유도하는 tRNA의 동정도 가능하게 될 것이다.

KEY WORDS HTLV-I, Ribosomal Frameshifting, Transframe protein

Retrovirus 입자가 재감염성을 보유하기 위해서는 바이러스의 캡시드 구성단백질을 코딩하는 *gag* 유전자에서 발현된 전구단백질이 Matrix단백질(MA)과 Capsid단백질(CA) 그리고 Nucleocapsid단백질(NA)로 정확히 process되어야 하며 마찬가지로 *pol* 유전자에서 발현된 전구단백질 역시 reverse transcriptase (RT)와 integrase(IN)로 processing에 의하여 분해되어야 한다(16). 이와 같은 Gag, Pol 구조단백질의 processing은 바이러스 genome내의 protease 유전자(*pro*)가 코딩한 효소작용의 결과임이 밝혀졌다(10). 지금까지 보고된 각종 retrovirus의 *pro* 유전자의 발현양식을 살펴보면 우선 termination codon의 suppression에 의해 *gag* 유전자의 initiation codon에서 개시된 polypeptide 합성이 *pro* 유전자까지 연장되는 방식(Molony-murine leukemia virus)과 유전자의 중첩영역에서 ribosome이 -1 방향으로 미끄러진 결과 triplet을 따라 번역해 오던 ribosome이 0 frame의 termination codon을 피하여 -1 frame인 *pro* 유전자로 진입하여 계속 polypeptide를 합성하는 방식(mouse mammary tumor virus). 그리고 *gag* open reading frame(ORF)의 일부가 protease이기 때문에 정상적인 *gag* 유전자의 translation기구에 의존하는 방식(Rous sarcoma virus)의 3가지로 대별할 수 있다(2, 3, 19). HTLV-I은 바이러스의 genome구조에서 볼때 -1 ribosomal frameshifting을 통하여 *pro* 유전자가 발현한다고 생각되었다. 왜냐하면 *pro* 유전자의 5'말단은 *gag*의 3'말단과 중첩되어 있고 *pro* 유

전자의 5'말단에는 initiator로 사용될만한 methionine codon을 찾아볼 수 없었기 때문이다(1, 7, 12). 이 추측은 vaccinia virus expression vector를 사용하여 조식배양세포에서 HTLV-I의 *gag*와 *pro* 유전자영역을 발현시키거나 *in vitro*에서 합성한 한 종류의 *gag-pro* mRNA를 무세포 단백질 합성계에서 translation시킨 실험결과에서 확인할 수 있었다(8, 9). 그러나 *gag-pro* mRNA에서 triplet을 인식하며 polypeptide 합성을 진행하던 ribosome이 *gag* frame에서 *pro* frame으로 switch하는 정확한 지점은 아직 실험적으로 밝혀져 있지 않다. 본 연구에서는 *gag-pro* 중첩영역에서 ribosomal frameshifting에 의하여 codon phase의 변동이 생긴 부분을 transframe 단백질의 일차구조해석을 통하여 결정할 목적으로 단순정제만으로 충분히 Edman degradation이 가능한 단백질을 코딩할 수 있는 *gag-pro* 중첩영역을 포함한 hybrid 유전자를 구축하였고 구축된 hybrid 유전자로부터 *in vitro* transcription 및 cell-free translation을 통하여 *in vivo*에서 정제할가능한 Gag-Pro transframe을 *in vitro*에서 비교적 다량으로 생산할 수 있는 단백질합성계의 확립을 시도하였다.

재료 및 방법

재조합 DNA 실험재료

재조합 DNA의 recipient cell로 *E. coli* JM109를 사용하였다. 플라스미드의 구축 및 검색에 사용된

DNA 제한효소와 Klenow fragment, T4 DNA ligase 및 8 nucleotide의 인산화된 SstI linker는 New England Biolabs의 제품을 사용하였다. 각종 효소 반응은 제조회사에서 지시한 조건에 따라 실시하였다. *Staphylococcus aureus*의 protein A 유전자는 Pharmacia LKB Biotec.이 시판하고 있는 플라스미드인 pRIT2T에서 분리하여 사용하였다(14, 15). Synthetic 유전자의 합성을 위한 oligonucleotide는 Applied Biosystem사의 automatic DNA synthesizer로 합성하였다.

DNA 염기서열의 측정

합성유전자와 gag-pro 유전자 중첩영역을 포함하는 HindIII-HincII DNA단편을 M13 mp18RF DNA에 삽입시킨 후 Sequenase kit(United States Biochemical Incorporation)를 이용하여 dideoxy chain termination법으로 염기서열을 결정하였다(10).

In vitro에서 mRNA의 합성

In vitro transcription 반응에는 Triton X-100으로 용균시킨 E. coli 세포추출물에서 Polyethylene glycol 6000-NaCl 침전에 의해 얻어진 ccDNA의 플라스미드만을 RNA 합성의 주형으로 사용하였다(17). mRNA는 Melton 등이 확립한 방법을 기초로 Stratagene의 실험조건에 따라 합성하였으며 반응에는 Stratagene에서 시판하는 mRNA capping kit를 이용하였다(6). 다만 SP6 RNA polymerase와 반응 buffer는 Promega에서 구입하였다.

In vitro translation

Cell-free translation에 사용된 rabbit reticulocyte lysate는 Promega에서 구입하였고 제조회사에서 권장하는 실험조건에 따라 단백질을 합성하였다. 반응후 반응혼합물의 일부를 Sodium Dodecyl Sulfate-12% polyacrylamide gel 전기영동과 20% PPO를 이용한 fluorography로 합성된 단백질의 분자량을 확인하였다(5).

Gag-Pro transframe 단백질의 정제

In vitro에서 합성된 transframe 단백질을 rabbit IgG-Sepharose(Pharmacia LKB Biotec.)를 사용하여 Pharmacia의 manual에 따라 정제하였다. 간단히 설명하면 50 mM Tris(pH 7.6), 150 mM NaCl, 0.05% Tween 20의 washing buffer로 충분히 beads를 세정한 다음 반응이 끝난 rabbit reticulocyte lysate를 직접 column에 주입하였다. 그 다음 washing buffer로 beads에 결합하지 않은 단백질을 씻어 버린 후 0.5 M 초산용액(pH 3.4)으로 IgG와 특이적으로 결합한 단백질을 용출시킨 다음 Amicon membrane concentrator로 용출된 단백질분획을 농축하였다.

결과 및 고찰

Gag-Pro transframe polyprotein을 코드하는 hybrid 유전자의 구축

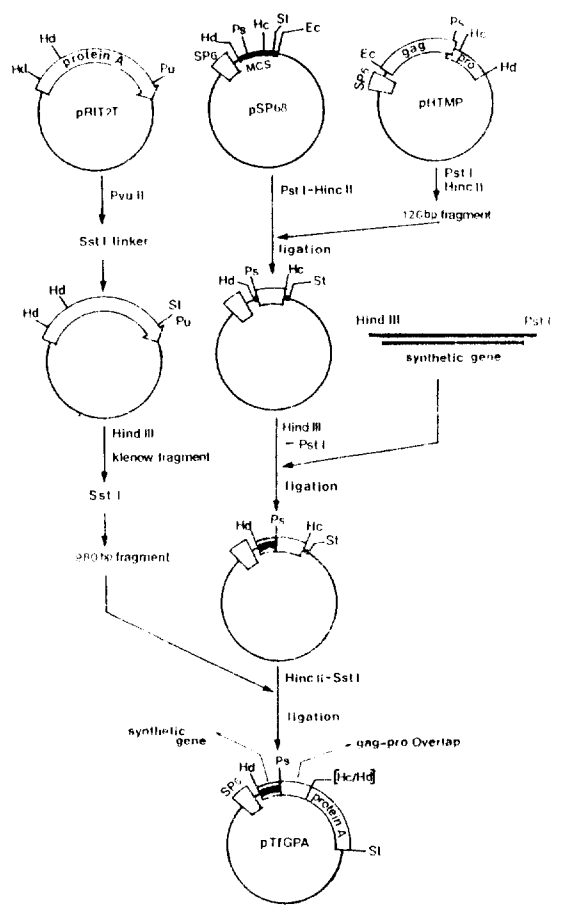


Fig. 1. Schematic diagram for construction of pTIGPA. Plasmid pRIT2T harbors *Staphylococcus aureus* protein A gene fragment. Abbreviations: MCS; multiple cloning site, Ec; EcoRI, Hc; HincII, Hd; HindIII, Ps; PstI, Pu; PvuII, St; StuI.

HTLV-1의 Gag-Pro polyprotein을 virion에서 정제하는 것이 가장 자연스럽고 바람직한 방법이나 바이러스를 대량으로 생산하는 배양세포주가 확립되지 못한 현 시점에서 이 방법에 의한 정제는 거의 불가능에 가깝다. 따라서 Gag-Pro polyprotein을 유전공학적으로 대량합성해야할 필요가 있지만 대부분의 경우 단백질의 순수정제에 어려움이 있다. 따라서 필요한 부분의 아미노산서열만을 신속하게 밝힐 수 있도록 in vitro에서 단백질을 합성하고 또 손쉽게 정제할 수 있는 방법을 HTLV-1의 경우에도 적용시키고자 다음의 실험을 실시하였다(4, Fig. 1). 우선 in vitro transcription을 위한 재조합 플라스미드를 구축하기 위하여 pHTMP를 PstI과 HincII로 절단하여 얻어진 124 bp의 gag-pro 중첩영역을 pSP68의 multiple cloning site에 삽입시켰다. pHTMP는 바이러스

스 genome의 nt.2074 부근의 염기서열인 CTCCAT가 site-directed mutagenesis에 의하여 *Pst*I site인 CAGCTG로 변경된 일종의 mutant로서 hybrid 유전자의 구축시 설계상의 문제점 때문에 mutant 클론을 사용할 수 밖에 없었다(9, 12). 실제로 두 염기의 치환이 *in vitro*에서 -1 ribosomal frameshifting의 효율에는 아무런 영향을 주지 않았으므로 mutant의 사용으로 야기되는 실험상의 문제점은 배제할 수 있을 것이다(9). Subclone된 124 bp의 단편에는 frameshifting site로 생각되는 adenine의 homopolymeric sequence(AAAAAAC)가 없을 뿐 아니라, translation이 시작되는 methionine codon도 포함되어 있지 않다. 이와 같은 제반 신호서열을 도입하기 위하여 5'-AGCTTATGAGGTCCCTACCACCC-CAAAAACTGCA-3'와 5'-GTTTTTGGGTGTGGTAGGGACCTCATA-3'의 상보적인 염기서열을 가진 36 nucleotide와 28 nucleotide의 oligonucleotide를 합성하였다. 두 oligonucleotides의 annealing으로 만든 synthetic gene을 124 bp의 DNA가 삽입된 pSP68의 *Hind*III와 *Pst*I site에 도입하여 일단 initiator methionine과 frameshifting site를 완비한 *gag-pro* 중첩영역을 구축하였다(Fig. 2). 그 다음 transframe 단백질을 신속 간단히 정제할 목적으로 *gag-pro* 중첩유전자단편의 3'말단에 다시 *Staphylococcus aureus*의 protein A 유전자를 도입하였다. protein A 유전자의 발현벡터인 pRIT2T의 *Pvu*II site에 8 nucleotide의 *Sst*I linker(5'-CGAGCTCG-3')를 부착시켜 *Sst*I site로 변환시킨 플라스미드를 *Hind*III 및 Klenow fragment의 처리로 *Hind*III 절단면을 blunt로 만든 다음 *Sst*I으로 재차 절단하여 protein A 유전자영역의 일부를 분리하였다(14, 15). 이 DNA 단편은 Protein A가 immunoglobulin의 constant영역을 인식하는데 필요한 기능 domain중 B domain과 A와 C domain의 일부에 해당하는데 immunoglobulin과의 결합에는 근개의 domain만으로도 충분하다고 보고 있다(12). 분리한 protein A 유전자단편을 *gag-pro* 중첩영역 하류의 *Hinc*II와 *Sst*I site사이에 삽입시켜서 결국 SP6 promoter의 하류에 5'-synthetic gene-*gag-pro* overlap-protein A-3'의 순서로 조립된 hybrid 유전자를 완성할 수 있었으며 구축된 플라스미드를 pTiGPA로 명명하였다. pTiGPA가 코드하는 polypeptide에서 N말단의 5 아미노산은 chick prelysozyme의 N말단 아미노산서열을 그대로 도입한 것이다(Fig. 4). *In vivo* 또는 *in vitro*에서 합성된 chick prelysozyme은 N말단에 지방산이 수식되지 않는다고 알려져 있으므로 hybrid 유전자에서 발현한 단백질은 Edman degradation의 좋은 기질이 될 것이다(13). 또한 ribosomal frameshift가 일어난다고 생각되는 homopolymeric sequence도 5'말단으로부터 27 nucleotide 떨어져 있으므로 10 cycle의 Edman degradation으로 reading frame의 switch에 의해 *pro* frame에서 최초로 코드된

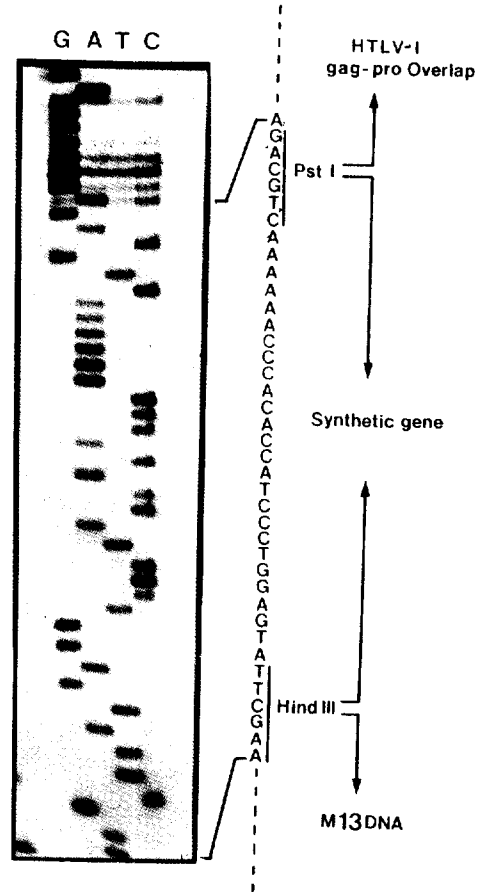


Fig. 2. Autoradiogram showing nucleotide sequence of the synthetic gene and its junction regions.

아미노산을 충분히 동정할 수 있을 것이다(4). hybrid 유전자의 또다른 장점은 pTiGPA가 발현하는 polypeptide중 -1 ribosomal frameshifting을 통하여 생성된 transframe 단백질만이 Protein A와 융합단백질의 형태로 합성되기 때문에 IgG와의 친화력을 이용하여 조단백질상태의 시료에서 직접 transframe 단백질을 정제할 수 있다는 것이다.

*In vitro*에서 mRNA의 합성

위에서 설명한 chimera 유전자로부터 *in vitro*에서 C말단에 Protein A영역을 보유하게 된 Gag-Pro transframe 단백질을 합성하기 위하여 우선 pTiGPA를 DNA 제한효소인 *Sst*I 또는 *Bam*HI으로 완전히 절단하였다. *Sst*I은 protein A 유전자의 3' untranslated region을 인식하지만 *Bam*HI은 *Sst*I site보다 620 bp 상류영역에 존재하는 pRIT2T 유래의 multiple cloning site 내부를 인식한다. 어떤 DNA 제한효소로 처리해도 pTiGPA의 전체크기를 나타내는 4.1 kb의 직선상 DNA가 관찰된다(Fig. 3, A). 그러나 각각의 DNA를 주형으로 SP6 RNA polymerase에

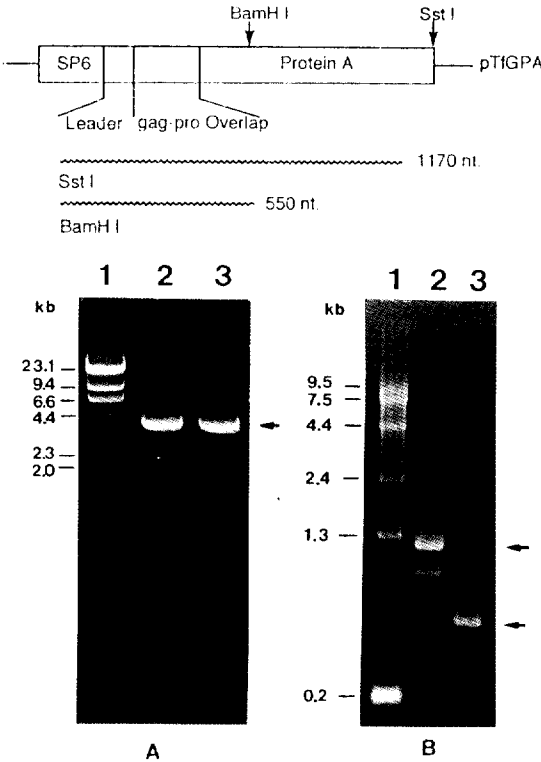


Fig. 3. *In vitro* synthesis of mRNAs (A) Linearization of pTfGPA with either SstI (lane 2) and BamHI (lane 3). DNAs were fractionated by 1% agarose gel electrophoresis. HindIII-digested λ DNA was used as molecular size standards (lane 1). (B) 1.5% agarose gel pattern of mRNAs transcribed *in vitro* from SstI-digested pTfGPA (lane 2) and from BamHI-digested pTfGPA (lane 3). 0.24~9.5 kb RNA ladder (lane 1) are used as molecular size standards.

의해 mRNA를 합성하면 SstI으로 절단한 DNA 주형에서는 1170 nucleotides의 RNA가 합성되는 반면 BamHI 처리한 DNA 주형으로부터는 550 nucleotides의 RNA가 합성되는 것을 볼 수 있다(Fig. 3, B). 이 결과는 pTfGPA가 보유한 hybrid gene의 구조와 잘 일치하고 있다.

Cell-free translation과 transframe 단백질의 정제

SstI으로 절단한 pTfGPA로부터 합성된 RNA를 *in vitro*의 transframe 단백질합성을 위한 message로 사용하였다. [³⁵S] methionine으로 표식한 전체 단백질의 일부를 SDS-polyacrylamide gel 전기영동을 통하여 분자량에 의해 분획한 결과 hybrid 유전자가 코드할 수 있는 단백질보다 큰 분자량의 비특이적 단백질이 소량 검출되었지만 21 kDa의 단백질이 *in vitro* translation에 의하여 생성된 주된 polypeptide임을 알 수 있었다(Fig. 4). 이 단백질이 gag-pro 중첩영역내에서 ribosomal frameshifting에 의하여

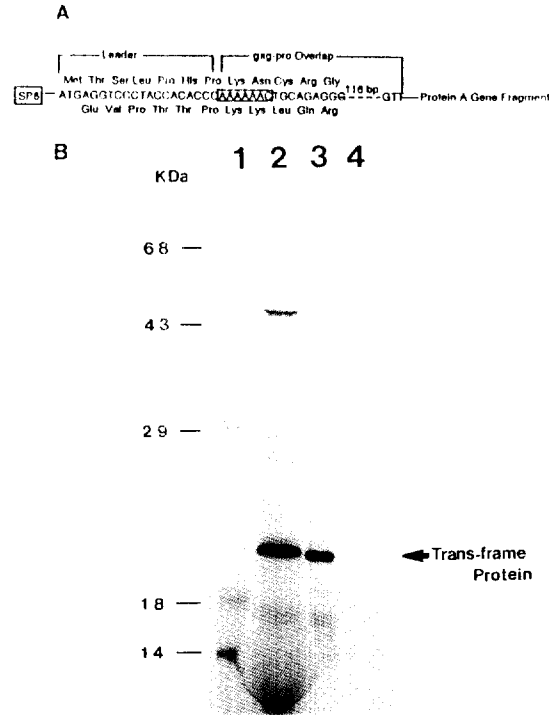


Fig. 4. *In vitro* translation and purification of the Gag-Pro trans-frame protein. (A) The nucleotide sequence of a portion of the protein sequencing vector pTfGPA is shown along with the translation in the gag frame (above the nucleotide sequence) and pro frame (below the nucleotide sequence). The box indicates the putative ribosomal frameshifting site in gag-pro overlap. (B) Fluorogram of [³⁵S] methionine-labeled proteins synthesized in response to RNA transcribed from SstI-digested pTfGPA (lane 2, 3) or no RNA (lane 4). The translation products were applied to SDS-polyacrylamide gel with (lane 3) or without purification (lane 2) through IgG-Sepharose column.

합성된 transframe 단백질이라는 것을 밝히기 위하여 IgG-Sepharose를 이용한 affinity chromatography를 실시하였다. *In vitro* translation으로 합성된 단백질 중에서 IgG와 특이적으로 결합한 분획을 산성 조건으로 분리시킨 다음 SDS-polyacrylamide gel 전기영동 및 fluorography로 용출된 단백질의 분자량을 검정하였는데 용출된 단백질의 분자량은 21 kDa로서 조단백질에서 관찰된 주반응생성물과 분자량이 일치하였다(Fig. 4, Lane 3). 또한 21 kDa 단백질은 IgG를 인식하는 성질로 미루어 이 단백질은 ribosomal frameshifting을 통하여 hybrid 유전자의 3'영역에 위치하는 protein A 유전자가 발현된 Gag-Pro transframe 단백질이라고 결론지을 수 있다. gag-pro 중첩영역에 존재하는 adenine의 homopolymeric

sequence의 일부를 다른 염기로 치환시키면 *pro* 유전자가 발현되지 않았다는 연구결과는 homopolymeric sequence로 구성된 염기서열인 AAAAAAC가 frameshifting site라는 사실을 시사한다(8, 9). pTfGPA는 frameshifting site로 추정되는 AAAAAAC가 hybrid 유전자의 5'말단에서 27 bp 떨어져 있도록 설계되었기 때문에 transframe 단백질의 N말단으로부터 10 아미노산 이내에 0 frame에서 -1 frame으로 reading frame의 switch로 인하여 -1 frame이 코딩하는 최초의 아미노산이 위치하고 있을 것으로 보인다. 그러나 *in vitro* translation으로는 합성가능한 단백질의 양에 제한이 있기 때문에 방사성 동위원소를 사용하는 microsequencing으로 아미노산서열을 결정하는 것이 최선의 방법으로 생각된다(18). 즉 [³⁵S] methionine과 함께 [³H] leucine이나 [³H] proline 또는 [¹⁴H] asparagine으로 단백질을 표식한 뒤 Edman degradation을 실시하여 합성된 transframe 단백질의 N말단부터 순차적으로 용출된 분획에 포함된 방사성동위원소의 에너지 profile과 hybrid gene의 0 또는 -1 frame에서 추정되는 아미노산서열중 leucine, proline 및 asparagine의 위치를 비교하면 손쉽게 reading frame의 switch가 일어난 부분을 동정할 수 있다(4). 현재 이 부분에 대한 연구를 계속 진행 중이며 본 연구에서 사용한 실험방법은 다수의 재료를 대상으로 간편하고 신속하게 ribosomal frameshifting을 위시한 기타 translational suppression 발생기작에 관계하는 *cis*- 및 *trans*-acting factor에 대한 지식을 넓혀가는데 크게 기여할 것으로 생각된다.

사 사

이 연구는 1990년~1991년도 한국과학재단 학술연구구조성비의 일부에 의하여 이루어졌음

참 고 문 헌

- Hiramatsu, K., J. Nishida, A. Naito and H. Yoshikura, 1987. Molecular cloning of the closed circular provirus of human T cell leukemia virus type I: a new open reading frame in the *gag-pol* region. *J. Gen. Virol.* **68**: 213-218.
- Jacks, T. and H.E. Varmus, 1985. Expression of the Rous sarcoma virus *pol* gene by ribosomal frameshifting. *Science*. **230**: 1237-1242.
- Jacks, T., K. Townsley, H.E. Varmus and J. Majors, 1987. Two efficient ribosomal frameshift events are required for synthesis of mouse mammary tumor virus *gag*-related polypeptides. *Proc. Natl. Acad. Sci. USA*. **84**: 4298-4302.
- Jacks, T., M.D. Power, F.R. Masiarz, P.A. Luciw, P.J. Barr and H.E. Varmus, 1988. Characterization of ribosomal frameshifting in HIV-1 *gag-pol* expression. *Nature*. **331**: 280-283.
- Laemmlli, U.K., 1970. Cleavage of structural

- proteins during the assembly of the head of bacteriophage T4. *Nature*. **227**: 680-685.
- Melton, D.A., P.A. Krieg, M.R. Rebagliati, T. Maniatis and M.R. Green. 1984. Efficient *in vitro* synthesis of biologically active RNA and RNA hybridization probes from plasmids containing SP6 promoter. *Nucleic Acids Res.* **12**: 7035-7056.
- Nam, S.H. and M. Hatanaka, 1986. Identification of protease gene of human T-cell leukemia virus type I (HTLV-I) and its structural comparison. *Biochem. Biophys. Res. Commun.* **139**: 129-135.
- Nam, S.H., M. Kidokoro, H. Shida and M. Hatanaka. 1988. Processing of *gag* precursor polyprotein of human T-cell leukemia virus type I by virus-encoded protease. *J. Virol.* **62**: 3718-3728.
- Nam, S.H. and M. Hatanaka. 1992. Ribosomal frameshifting signals required for the expression of *pro* gene of human T-cell leukemia virus type I (HTLV-I). (manuscript in preparation).
- Oroszlan, S. and R.B. Luftig. 1990. Retroviral proteinases. *Curr. Topic. Microbiol. Immunol.* **157**: 153-185.
- Sanger, F., S. Nicklen and A.R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA*. **74**: 5463-5467.
- Seiki, M., S. Hattori, Y. Hirayama and M. Yoshida. 1983. Human adult T-cell leukemia virus: complete nucleotide sequence of the provirus genome integrated in leukemia cell DNA. *Proc. Natl. Acad. Sci. USA*. **85**: 3618-3622.
- Palmiter, R.D., J. Gagnon, L.H. Ericsson and K.A. Walsh. 1977. Precursor of egg white lysozyme. *J. Biol. Chem.* **252**: 6386-6393.
- Uhlen, M., B. Guss, B. Nilsson, S. Gatenbeck, L. Philipson and M. Lindberg. 1984. Complete sequence of Staphylococcus gene encoding Protein A. *J. Biol. Chem.* **259**: 1695-1702.
- Uhlen, M., M. Lindberg and L. Philipson. 1984. The gene for staphylococcal protein A. *Immunol. Today*. **5**: 244-248.
- Weiss, R., N. Teih, H.E. Varmus and J. Coffin. (eds) 1984. RNA tumor viruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
- Wu, R., L. Grossman and K. Moldave. (eds) 1983. Methods in enzymology. 100. Academic Press, New York.
- Yoshinaka, Y. and S. Oroszlan. 1985. Bovine leukemia virus post-envelope gene coded protein: evidence for expression in natural infection. *Biochem. Biophys. Res. Commun.* **131**: 347-354.
- Yoshinaka, Y., I. Katoh, T.D. Copeland and S. Oroszlan. 1985. Murine leukemia virus protease is encoded by the *gag-pol* gene and is synthesized through suppression of an amber termination codon. *Proc. Natl. Acad. Sci. USA*. **82**: 1618-1622.

(Received September 9, 1992)

(Accepted September 23, 1992)

ABSTRACT: Construction of Recombinant DNA for Purification of the Gag-Pro Transframe Protein of Human T-cell Leukemia Virus Type I (HTLV-I)

Seok Hyun Nam (Department of Biological Science, College of Natural Science, Ajou University, Suwon 441-749, Korea)

To determine the site at which -1 ribosomal frameshifting occurs within the *gag-pro* overlap of HTLV-I, DNA fragment corresponding to a portion of the gene overlap was cloned into a SP6 vector. The resultant plasmid harbors the hybrid gene consisting of a synthetic gene encoding 5 amino acids derived from chick prelysozyme including the initiator methionine plus 141 nucleotides of *gag-pro* overlapping region followed by *Staphylococcus aureus* protein A gene fragment. *In vitro* transcription by SP6 RNA polymerase with this DNA template made an abundant amount of single species mRNA. Cell-free translation programmed with the RNA transcribed *in vitro* yielded a polypeptide of 21 kDal in size, which could be purified into homogeneity by IgG-Sepharose affinity chromatography. *In vitro* system described in this study must be useful for rapid purification and sequencing of the Gag-Pro transframe protein, allowing to determine the exact frameshift site on mRNA and to identify the tRNA involved in frameshifting event for the expression of *pro* gene.