Volume 293, number 1,2, 106–110 FEBS 10397 November 1991

Purification and characterization of human T-cell leukemia virus type I protease produced in *Escherichia coli*

Makoto Kobayashi[1], Yumiko Ohi[1], Tsuneo Asano[1], Takaki Hayakawa[1], Koichi Kato[1], Atsushi Kakinuma[1] and Masakazu Hatanaka[2]

[1]*Biotechnology Research Laboratories, Takeda Chemical Industries, Ltd., Yodogawa-ku, Osaka 532, Japan* and [2]*Institute for Virus Research, Kyoto University, Shogoin, Kyoto 606, Japan*

Received 2 September 1991; revised version received 24 September 1991

Human T-cell leukemia virus type I (HTLV-I) protease has been purified to homogeneity from a strain of recombinant *Escherichia coli*. The protease was expressed as a larger precursor, which was autoprocessed to form a mature protease. Protein chemical analyses revealed the coding sequence of mature protease, which agreed with the putative sequence predicted from the sequence of bovine leukemia virus protease. The purified protease processed the natural substrate *gag* precursor (p53) to form *gag* p19 and *gag* p24. The protease activity was inhibited by pepstatin A. These results provide direct evidence that this protease belongs to the aspartic protease family and has an activity consistent with the protease in HTLV-I virion.

HTLV-I protease; Aspartic protease; *gag* precursor

1. INTRODUCTION

Human T-cell leukemia virus type I (HTLV-I) is a retrovirus causing adult T-cell leukemia (ATL) and HTLV-I-associated myelopathy (HAM) [1–3]. The HTLV-I genome has been cloned from the peripheral blood leukocytes of an ATL patient, and subsequent characterization of the viral genome has revealed that it is composed of *gag*, *pol*, *env*, and *pX* genes [4]. HTLV-I proteins, like other retroviral proteins, are initially translated as the large precursor polyproteins, which are proteolytically processed to generate structural proteins (p19, p24, p15), enzymes (protease, reverse transcriptase, and integrase), and envelope proteins [5–11]. Processing of the *gag* and *pol* polyproteins is catalyzed by a viral-encoded specific protease [5–12]. Nam and Hatanaka [13] first demonstrated the nucleotide sequence of the protease from replication-competent provirus HTLV1C [14]. The open reading frame (ORF) for the protease (*prt*) is located between *gag* and *pol* and is different from those for *gag* and *pol*. Therefore, the occurrence of two frameshift events is believed to be necessary for the replication of the virus [15].

The putative amino acid sequence deduced from the ORF for the HTLV-I protease has suggested that it belongs to the family of aspartic proteases, like other retroviral proteases [13,15,16]. The active sites of cellular aspartic proteases contain two aspartic residues, both of which occur in a sequence-DTG-. The retroviral proteases contain only one such-DTG-sequence and their molecular weights are slightly less than half the size of typical aspartic proteases [17]. Recently, recombinant human immunodeficiency virus 1 (HIV-I) protease was obtained using a genetic engineering technique [18–21], which made it possible to investigate the biochemical properties of the protease and its inhibitors. Furthermore, the X-ray crystallographic analyses of HIV-1 protease [22] and avian retroviral protease [23] have revealed that these proteases exist as a dimer of identical subunits.

In this paper, we describe the purification of enzymatically active HTLV-I protease expressed in recombinant *Escheria coli* cells and the characterization of the purified enzyme.

Correspondence address: M. Kobayashi, Biotechnology Research Laboratories, Research and Development Division, Takeda Chemical Industries Ltd., Jusohonmachi 2-chome, Yodogawa-ku, Osaka 532, Japan. Fax: (81) (6) 300 6306.

2. MATERIALS AND METHODS

2.1. *Source of recombinant HTLV-I protease*

The expression of HTLV-I protease in *E. coli* will be published elsewhere. In brief, 750 bp *Mbo*II-*Eco*RI fragment of p7.5gagfs19 [15] containing the 3'portion of *gag* and the entire *prt* was ligated with an ATG-linker and the resulting 760 bp fragment was expressed in *E. coli* under the control of a T7 promoter. A transformant, *E. coli* MM294[DE3]/pHT17707, was cultured at 37°C in LB medium and HTLV-I protease was produced after induction with 0.1 mM isopropyl β-D-thiogalactoside.

2.2. *Purification of HTLV-I protease*

All the solutions used contained 0.1% nonidet P40 (NP40), 10% glycerol, 5 mM EDTA, 1 mM phenylmethylsulfonyl fluoride (PMSF), 0.1 mM (*p*-amidino-phenyl)methanesulfonyl fluoride (APMSF), 0.02% 2-mercaptoethanol (2-ME). *E. coli* MM294[DE3]/pHT117707 cells were suspended in 50 mM Tris-HCl (pH 7.6), 1% NP40, 100 μg/ml egg white lysozyme (Seikagaku Kogyo Co.). The cells were

 Published by Elsevier Science Publishers B.V.

disrupted by sonication. The crude extract was obtained by centrifugation, diluted with 4 vols. of 10 mM sodium phosphate buffer (pH 6.0), and loaded onto an S-Sepharose fast-flow (Pharmacia) column (2.5 × 25 cm) equilibrated with the same buffer. HTLV-I protease was eluted with a gradient of NaCl concentrations. The crude protease from the S-Sepharose column was concentrated with a YM 5 membrane, and loaded onto a Sephacryl S-100HR (Pharmacia) column (5 × 82 cm) equilibrated with 50 mM sodium phosphate buffer (pH 6.0) and 0.1 M Na_2SO_4. The fractions containing the active protease were pooled, diluted with 20 mM CH_3COONH_4 buffer (pH 5.5), and loaded onto a mono S column (Pharmacia; 0.5 × 5 cm). HTLV-I protease was eluted with a gradient of NaCl concentrations. The fractions containing the active protease were pooled and re-chromatographed on the same column. The fractions containing the purified protease were collected and stored at −80°C before use.

2.3. *Preparation of recombinant HTLV-I* gag *p53*

The expression and purification of recombinant HTLV-I *gag* p53 in *Saccharomyces cerevisiae* (NA74-3Aρ^-/pHTI511) will be described elsewhere. In brief, recombinant yeast cells expressing HTLV-I *gag* p53 were disrupted by BEAD-BEATER (Biospec Products) and centrifuged. The microsomal fraction was obtained by ultracentrifugation and *gag* p53 was purified by Butyl-Toyopearl 650M (Tosoh) and Protein PAK G-CM (Waters) column chromatography.

2.4. *Determination of protease activity*

The activity of the HTLV-I protease was determined by measuring the rate of hydrolysis of a synthetic peptide, gag-1 (YVEPTAPQVLPVMHP). Gag-1 (12 μg) was dissolved in 30 μl of 100 mM sodium citrate buffer (pH 5.0). The reaction was initiated by the addition of the protease, and the reaction mixture was incubated at 37°C for 1 h. The reaction was terminated by adding trichloroacetic acid and the quantity of a product peptide, YVEPTAPQVL, was determined with reverse-phase HPLC using a YMC-PACK FL-ODS column (Yamamura Kagaku; 4.6 × 30 mm). Pepstatin A used for inhibition studies was purchased from Sigma and dissolved in dimethyl sulfoxide at 10 mg/ml before use.

2.5. *Protein chemical analyses*

The N-terminal amino acid sequence was determined with a gas-phase protein sequencer model 470A and a model 120A analyzer (Applied Biosystems Inc.). Amino acids at the C-terminal region were determined by hydrazinolysis [24] and carboxy peptidase Y digestion. Amino acid composition was determined with a PICO-TAG amino acid analysis system (Waters).

3. RESULTS

3.1. *Purification of HTLV-I protease*

Western blotting analysis of a cell free extract of *E. coli* MM294[DE3]/pHTI7707 gave a band with an M_r of 14 000 which was reactive to anti-HTLV-I protease antibodies (Fig. 1A). Purification of the recombinant HTLV-I protease was monitored by the peptidolytic assay using YVEPTAPQVLPVMHP (gag-1) which has the *gag* p19/p24 cleavage site between Leu^{130} and Pro^{131} in *gag* p53. Treatment of gag-1 with the active protease resulted in a time-dependent decrease of the substrate on high performance liquid chromatography analysis and a concomitant increase of 2 peaks, which comigrated with the chemically synthesized anticipated products, YVEPTAPQVL (gag-3) and PVMHP (gag-2), confirming the position of the cleavage site within the peptide (data not shown). A summary of the purification is shown in Table I. The SDS-polyacrylamide gel electrophoresis (SDS-PAGE) pattern of the purified protease under reducing conditions is shown in Fig. 1B. Major and minor bands with M_rs of 14 400 and 28 600, respectively, were detected. Each band was reactive to anti-protease antibodies (data not shown), so these 2 bands were judged to be monomer and dimer forms of the mature protease. The purity was estimated to be greater

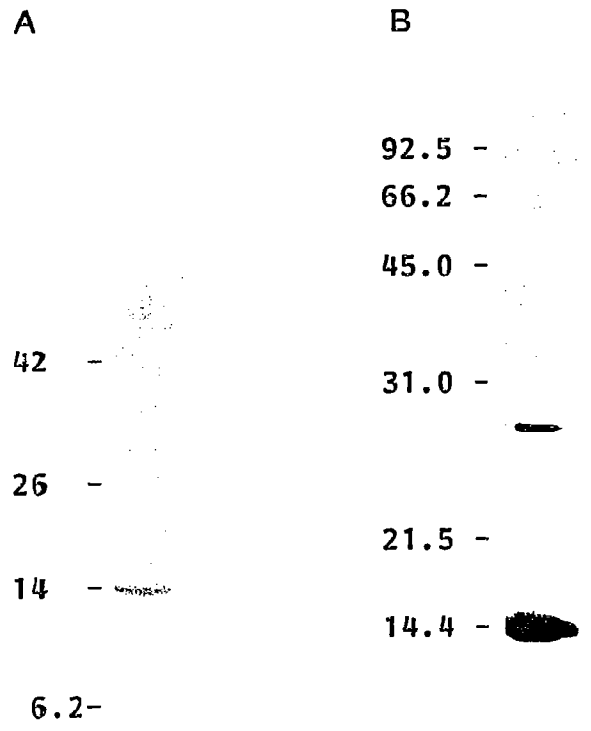

Fig. 1. (A) Immunoblotting of a cell free extract of *E. coli* expressing HTLV-I protease; (B) SDS-PAGE of purified HTLV-I protease. SDS-PAGE was carried out according to the method of Laemmli [25] using 18% (A) and 15% (B) acrylamide gels. (A) HTLV-I protease was developed with the affinity-purified anti-protease peptide (DPARRP-VIKAQVDTQTS, I-7) rabbit IgG and HRP-conjugated goat anti-rabbit IgG. The marker proteins used were prestained standards from BRL. (B) Protein bands were stained with silver staining reagent. The marker proteins used were low mol. wt. standards from Bio-Rad.

Table I

Summary of the purification of HTLV-I protease

Step	Volume (ml)	Total protein[1] (mg)	Total activity[2] (mU)	Specific activity (mU/mg)	Yield (%)
Extract	2500	13500	7800	0.58	100
S-Sepharose	700	161	4259	26.5	55
Sephacryl S-100HR	305	33	2517	76.0	32
Mono S 1st	320	16	1556	96.6	20
Mono S 2nd	375	7	826	111.0	11

[1]Protein content was determined by micro BCA protein assay (Pierce) using bovine serum albumin as standard

[2]One unit of the enzyme activity was defined as the amount of enzyme that formed 1 μmol of gag-3/min in the standard assay

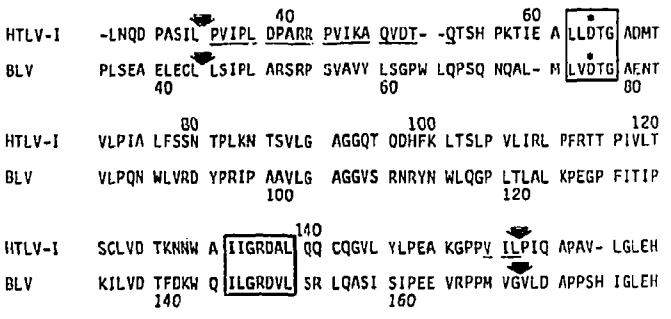

Fig. 2. Putative amino acid sequences of *prt* of HTLV-I [13] and BLV [6,11]. Arrows indicate putative autoprocessing sites; boxes indicate conserved regions in aspartic proteases; *, putative active site of aspartic proteases.

than 95% by densitometric analysis. The amount of the purified HTLV-I protease obtained was 7 mg starting from 20 liter of culture. This preparation was used as HTLV-I protease throughout the present work.

3.2. *Protein chemical properties*

To determine the region which codes for the active protease, the N- and C-terminal amino acid sequences were analyzed. The N-terminal amino acid sequence of the purified protease was found to be PVIPLDPARRP-VIKAQVDT-. The C-terminal amino acid was found to be Leu by the hydrazinolysis method. Since carboxy peptidase Y released Leu, Ile and Val from the purified protease the C-terminal amino acid sequence was determined to be -VIL. The N- and C-terminal sequences were identical to those in the putative HTLV-I protease sequences predicted from that of bovine leukemia virus (BLV) protease (Fig. 2). Amino acid composition of the purified protease also agreed with that calculated from the deduced amino acid sequence consisting of 125 residues.

Analytical gel filtration of the purified HTLV-I protease demonstrated that the activity was eluted with an M_r of 27 000 at a position just prior to the protein standard chymotrypsinogen A (M_r 25 700, Fig. 3). The M_r of native protease obtained was nearly twice that calculated from the amino acid sequence consisting of 125 residues (M_r 13 459, Fig. 2) and that determined by SDS-PAGE (M_r 14 400, Fig. 1B).

3.3. *Enzymatic properties*

On incubation with *gag* p53, the protease generated 2 new bands with M_rs of 19 000 and 24 000 which were detected with anti-HTLV-I *gag* monoclonal antibodies, GIN-7 (anti-p19) and NOR-1 (anti-p24), respectively (Fig. 4). These results suggested that the HTLV-I protease expressed in the bacterial cells had the ability to cleave the natural substrate, *gag* precursor, as well as the synthetic peptide.

The optimum pH for hydrolytic activity of the synthetic peptide was 5. A similar pH-dependence was observed for the cleavage of natural substrate *gag* p53. The

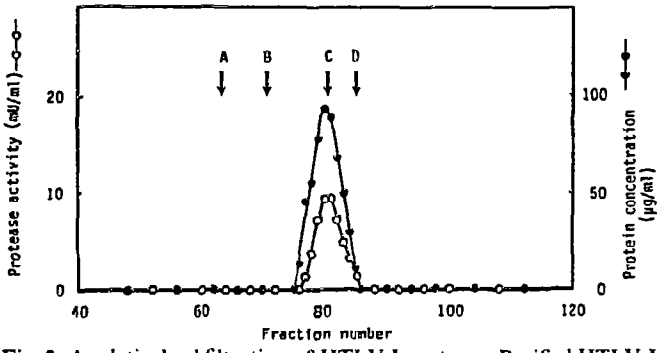

Fig. 3. Analytical gel filtration of HTLV-I protease. Purified HTLV-I protease (1 mg/0.5 ml) was loaded on a Hi-Load 16/60 Sephacryl S-200HR column (Pharmacia; 1.6 × 60 cm) equilibrated with 50 mM sodium phosphate buffer (pH 6.0), 0.1 M Na_2SO_4, 0.1% NP40, 10% glycerol, 5 mM EDTA, 1 mM PMSF, 0.1 mM APMSF, 0.02% 2-ME at a flow rate of 30 ml/h. (○), protease activity; (●), protein concentration; arrows indicate peaks of marker proteins: A, bovine serum albumin (66 200); B, ovalbumin (45 000); C, chymotrypsinogen A (25 700); D, ribonuclease A (13 700).

peptide hydrolytic activity obeyed simple Michaelis–Menten kinetics: a K_m value of 0.50 mM and a k_{cat} value of 0.085 sec^{-1} were obtained at pH 5.0 by Woolf plot. The K_m value for gag-1 changed remarkably depending on the concentration of $(NH_4)_2SO_4$. In the presence of 0.5 M $(NH_4)_2SO_4$, the K_m value decreased by 20-fold, whereas the k_{cat} value was not affected.

Pepstatin A, a typical inhibitor of aspartic proteases, inhibited the proteolysis of gag-1 by HTLV-I protease.

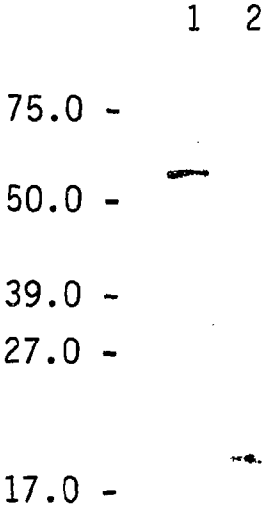

Fig. 4. Processing of *gag* precursor p53. HTLV-I *gag* precursor p53 was incubated with HTLV-I protease for 3 h at 37°C. The reaction was terminated by adding SDS-PAGE sample buffer and heating at 100°C for 5 min. SDS-PAGE and immunoblotting were carried out by using anti-p19 mouse monoclonal IgG, GIN-7, and the anti-p24 mouse monoclonal IgG, NOR-1; (Lanes 1–3) treated with buffer only; (lanes 2–4) treated with HTLV-I protease. Lanes 1 and 2 were developed with GIN-7 and lanes 3 and 4 with NOR-1. The marker proteins used were prestained standards from Bio-Rad.

The K_i value for pepstatin A was determined to be 17 μM, which was far greater than those for pepsin (0.045 nM), cathepsin D (0.1 nM) and HIV protease (0.5 μM). The K_i value was reduced to 2.2 μM in the presence of 0.5 M $(NH_4)_2SO_4$.

4. DISCUSSION

It was postulated that the HTLV-I genome encoded a protease essential for the virus' replication based on analogy with other retroviruses [5,6,9–11,21]. We have already reported the ORF encoding HTLV-I protease [13], but the structure of the mature protease and its biochemical properties have remained to be clarified. Here, we have reported the production of the active HTLV-I protease in *E. coli*, which made it possible to investigate its biochemical properties. As there was no direct information for the region encoding the mature protease, the protease was designed to be expressed as a larger precursor in *E. coli* cells. Protein chemical analyses of the purified protease have revealed that both N- and C-terminal peptides are released on cleavage of the peptide bonds between Leu^{32} and Pro^{33} and between Leu^{157} and Pro^{158}, respectively, to form the mature protease composed of 125 amino acid residues (Fig. 2). Therefore, this protease is thought to recognize its own N- and C-terminal sequences and to autoprocess itself exactly. This is the first report that directly shows the structure of mature HTLV-I protease. The analytical gel filtration revealed that the M_r of the native protease was almost twice that calculated from the amino acid sequence, suggesting that the protease existed as a dimer like other retroviral proteases such as HIV-1 and avian retroviral proteases [21,26].

We reported that replacing Asp^{64} in the putative active site with Gly resulted in the abolishment of processing of the *gag* precursor in cells infected by recombinant vaccinia virus [15]. We have also found the loss of autoprocessing activity in *E. coli* cells by the same replacement (unpublished). These results clearly suggest that the protease belongs to aspartic protease. Biochemical characterization of the purified HTLV-I protease demonstrated that the activity was inhibited by pepstatin A, a typical inhibitor of aspartic proteases, but the K_i value was far greater than those for other aspartic proteases, such as pepsin, cathepsin D, and even HIV-1 protease. Therefore, HTLV-I protease may have an active site somewhat different from those of other aspartic proteases. HTLV-I protease seems to be less similar to cellular aspartic proteases than HIV-1 protease as judged by pepstatin A inhibition. There are some differences in the predicted amino acid sequences of the protease between HTLV-I and HIV-1 [27]. The differences in the inhibitory effect of pepstatin A might reflect the evolutional distance of these two viruses.

The activity of the HTLV-I protease changed remarkably in the presence of $(NH_4)_2SO_4$. The K_m for gag-1 and K_i for pepstatin A decreased depending on the concentration of $(NH_4)_2SO_4$, but little change of k_{cat} was observed, suggesting that $(NH_4)_2SO_4$ affects the interaction between the enzyme and the substrate rather than the peptide hydrolytic activity. Similar results were reported for the effects of salts on the K_m values of HIV-1 protease [28]. This property may be common to retroviral proteases.

Pathogenic human retroviruses, HTLV-I and HIV-1, are classified as belonging to different groups of retroviruses based on the differences in their genomic structures and pathogenicities. Recently, the HIV-1 protease was obtained using a genetic engineering technique [18–21]. Here, we have obtained HTLV-I protease using the same technique and it has become possible to compare the biochemical properties of these human retroviral proteases with each other and with retroviral proteases of different origin. Such information will facilitate the investigation of virion maturation of pathogenic retroviruses and may lead to the discovery of effective agents for the treatment of diseases caused by retroviruses.

Acknowledgements: We are grateful to Drs. K. Igarashi, Y. Fujisawa and K. Tsukamoto for helpful discussion. We thank Dr. O. Nishimura and T. Kurokawa for antibodies against anti-protease peptides, and to Dr. T. Fukuda and Mrs. M. Kusunoki for providing the synthetic peptides. Thanks are due to Dr. K. Kitano for cultivating the recombinant *E. coli* cells, and to J.A. Hogan for reading the manuscript.

REFERENCES

[1] Hinuma, Y., Nagata, K., Hanaoka, M., Nakai, M., Matsumoto, T., Kinoshita, K., Shirakawa, S. and Miyoshi, I. (1981) Proc. Natl. Acad. Sci. USA 78, 6476–6480.
[2] Osame, M., Usui, K., Izumo, S., Ijichi, N., Amitani, H., Igata, A., Matsumoto, M. and Tara, M. (1986) Lancet i, 1031–1032.
[3] Poiesz, B.J., Ruscetti, F.W., Reitz, M.S., Kalyanaraman, V.S. and Gallo, R.C. (1981) Nature 294, 268–271.
[4] Hattori, S., Kiyokawa, T., Inakawa, K., Shimizu, F., Hashimura, E., Seiki, M. and Yoshida, M. (1984) Virology 136, 338–347.
[5] Katoh, I., Yoshinaka, Y., Rein, A., Shibuya, M., Okada, T. and Oroszlan, S. (1985) Virology 145, 280–292.
[6] Sagata, N., Yasunaga, T. and Ikawa, Y. (1984) FEBS Lett. 178, 79–82.
[7] Schwartz, D.E., Tizard, R. and Gilbert, W. (1983) Cell 32, 853–869.
[8] von der Helm, K. (1977) Proc. Natl. Acad. Sci. USA 74, 911–915.
[9] Yoshinaka, Y., Katoh, I., Copeland, T.D. and Oroszlan, S. (1985) J. Virol. 55, 870–873.
[10] Yoshinaka, Y., Katoh, I., Copeland, T.D. and Oroszlan, S. (1985) Proc. Natl. Acad. Sci. USA 82, 1618–1622.
[11] Yoshinaka, Y., Katoh, I., Copeland, T.D., Smythers, G.W. and Oroszlan, S. (1986) J. Virol. 57, 826–832.
[12] Dittmar, K.J. and Moelling, K. (1978) J. Virol. 28, 106–118.
[13] Nam, S.H. and Hatanaka, M. (1986) Biochem. Biophys. Res. Commun. 139, 129–135.
[14] Mori, K., Sabe, H., Siomi, H., Iino, T., Tanaka, A., Takeuchi, K., Hirayoshi, K. and Hatanaka, M. (1987) J. Gen. Virol. 68, 499–506.
[15] Nam, S.H., Kidokoro, M., Shida, H. and Hatanaka, M. (1988) J. Virology 62, 3718–3728.
[16] Katoh, I., Yasunaga, T., Ikawa, Y. and Yoshinaka, Y. (1987) Nature 329, 654–656.
[17] Pearl, L.H. and Taylor, W.R. (1987) Nature 329, 351–354.

[18] Darke, P.L., Leu, C.-T., Davis, L.J., Heimbach, J.C., Diehl, R.E., Hill, W.S., Dixon, R.A.F. and Sigal, I.S. (1989) J. Biol. Chem. 264, 2307–2312.
[19] Hansen, J., Billich, S., Schulze, T., Sukrow, S. and Moelling, K. (1988) EMBO J. 7, 1785–1791.
[20] Krausslich, H.-G., Ingraham, R.H., Skoog, M.T., Wimmer, E., Pallai, P.V. and Carter, C.A. (1989) Proc. Natl. Acad. Sci. USA 86, 807–811.
[21] Meek, T.D., Dayton, B.D., Metcalf, B.W., Dreyer, G.B., Strickler, J.E., Gorniak, J.G., Rosenberg, M., Moore, M.L., Magaard, V.W. and Debouck, C. (1989) Proc. Natl. Acad. Sci. USA 86, 1841–1845.
[22] Navia, M.A., Fitzgerald, P.M.D., McKeever, B.M., Leu, C.-T., Heimbach, J.C., Herber, W.K., Sigal, I.S., Darke, P.L. and Springer, J.P. (1989) Nature 337, 615–620.
[23] Miller, M., Jakolski, M., Rao, J.K.M., Leis, J. and Wlodawer, A. (1989) Nature 337, 576–579.
[24] Narita, K., Murakami, H. and Ikehara, T. (1966) J. Biochem. 59, 170–175.
[25] Laemmli, U.K. (1970) Nature 227, 680–685.
[26] Kotler, M., Danho, W., Katz, R.A., Leis, J. and Skalka, A.M. (1989) J. Biol. Chem. 264, 3428–3435.
[27] Yasunaga, T., Sagata, N. and Ikawa, Y. (1986) FEBS Lett. 199, 145–150.
[28] Wondrac, E.M., Louis, J.M. and Oroszlan, S. (1991) FEBS Lett. 280, 344–346.