Thermostable aminopeptidase from *Pyrococcus horikoshii*

Susumu Ando[a], Kazuhiko Ishikawa[a], Hiroyasu Ishida[a], Yutaka Kawarabayasi[a,b], Hisasi Kikuchi[b], Yoshitsugu Kosugi[a,*]

[a]*National Institute of Bioscience and Human-Technology, 1-1 Higashi, Tsukuba, Ibaraki 305-8566, Japan*
[b]*National Institute of Technology and Evaluation, Nishihara, Shibuyaku, Tokyo 151-0066, Japan*

Received 18 December 1998; received in revised form 12 February 1999

Abstract From the genome sequence data of the thermophilic archaeon *Pyrococcus horikoshii*, an open reading frame was found which encodes a protein (332 amino acids) homologous with an endoglucanase from *Clostridium thermocellum* (42% identity), deblocking aminopeptidase from *Pyrococcus furiosus* (42% identity) and an aminopeptidase from *Aeromonas proteolytica* (18% identity). This gene was cloned and expressed in *Escherichia coli*, and the characteristics of the expressed protein were examined. Although endoglucanase activity was not detected, this protein was found to have aminopeptidase activity to cleave the N-terminal amino acid from a variety of substrates including both *N*-blocked and non-blocked peptides. The enzyme was stable at 90°C, with the optimum temperature over 90°C. The metal ion bound to this enzyme was calcium, but it was not essential for the aminopeptidase activity. Instead, this enzyme required the cobalt ion for activity. This enzyme is expected to be useful for the removal of N^{α}-acylated residues in short peptide sequence analysis at high temperatures.

Key words: Aminopeptidase; Thermostability; Archaeon; *Pyrococcus horikoshii*

1. Introduction

Pyrococcus horikoshii strain OT3 is a hyperthermophilic archaeon with the optimum growth temperature at 98°C [1]. The acylamino acid releasing enzyme from this organism was shown to be highly thermophilic [2] and the other enzymes are also expected to be highly thermophilic. The sequencing of the genome of this organism was completed at the National Institute of Technology and Evaluation (Tokyo, Japan) [3,4]. During a search for endoglucanase genes from the genome sequence data of *P. horikoshii*, we found a gene to encode a protein homologous with an endoglucanase from *Clostridium thermocellum* (42% identity), deblocking aminopeptidase from *Pyrococcus furiosus* (42% identity) and an aminopeptidase from *Aeromonas proteolytica* [5] (18% identity). As yet there have been no reports concerning this protein. Therefore we cloned and expressed this gene from *P. horikoshii* in *Escherichia coli* and examined the characteristics of the expressed protein.

*Corresponding author. Fax: (81) (298) 54-6414.
E-mail: kosugi@nibh.go.jp

Abbreviations: Ac-, N^{α}-acetyl; f-, N^{α}-formyl; pNA, *p*-nitroanilide; NEM, *N*-ethylmorpholine

2. Materials and methods

2.1. Materials

Avicel SF was purchased from Asahi Kasei (Tokyo, Japan). Carboxymethyl cellulose and xylan were purchased from Nacalai Tesque (Kyoto, Japan). The peptides, acylpeptides and acylamino acid *p*-nitroanilides were purchased from Sigma (St. Louis, MO, USA), Bachem (Bubendorf, Switzerland) and Peptide Institute, (Minoh, Osaka, Japan). One letter symbols are used for the two peptides AcLEHD-4-methylcoumarin-7-amide (AcLEHD-MCA) and ARGIKGIRGFSG. The synthesis of DNA primers was performed by Nihon Seifun (Atsugi, Kanagawa, Japan). All the chemicals and reagents used were of the highest reagent grade commercially available.

2.2. Cloning and expression of the gene

The gene of *P. horikoshii* homologous to the *C. thermocellum* gene for an endoglucanase, the *P. furiosus* gene for the deblocking aminopeptidase and the *A. proteolytica* gene for an aminopeptidase was found using a BLAST search [6]. The gene was amplified using the polymerase chain reaction (PCR) with primers having *Nde*I and *Bam*HI restriction sites according to the method reported previously [2]. The sequences of the primers were: 5′-TTTTGAATTCTTGCATATGATGTCAATATAGAGAAG-3′ (upper primer, containing an *Nde*I cutting site as underlined) and 5′-TTTTGGTACCTTTGGATCCTTATCCCTCCTAGAGCTCAAATGCTAA (lower primer, containing a *Bam*HI cutting site as underlined). Since there was an *Nde*I cutting site in the gene sequence (positions 273–278), the T at position 275 was replaced by a C prior to gene amplification so that the restriction site was eliminated without altering the amino acid sequence of the encoded protein. This was done according to the overlap extension PCR method [7]. The amplified gene was hydrolyzed using *Nde*I and *Bam*HI and inserted in pET11a cut with the same restriction enzymes. The nucleotide sequence of the inserted gene was determined using an LI-COR Model LIC-4200L(S)-2 sequencer (Aloka, Mitaka, Tokyo, Japan) to verify the identity with the anticipated sequence. The amplified gene was expressed using the pET11a vector system in the host *E. coli* BL21(DE3) according to the instructions by the manufacturer. The host *E. coli* BL21(DE3) was first transformed using the constructed plasmid, after which the production of the protein was performed according to the method described previously [2]. The concentration of the expressed protein was determined using a Coomassie protein assay reagent (Pierce Chemical Company, Rockford, IL, USA) and utilizing bovine serum albumin as the standard protein.

2.3. Purification of the enzyme

The crude enzyme solution was prepared from the transformant *E. coli* and the enzyme was purified by chromatography using a HiTrap Q column (Pharmacia, Uppsala, Sweden) and a HiLoad Superdex 200 column (Pharmacia). The details of these procedures are as described previously [2]. The purified protein demonstrated a single band with a molecular weight of 36.9 kDa measured on SDS-PAGE followed by Coomasie blue staining.

The molecular weight was also determined using high performance liquid chromatography (HPLC) on a Superdex 200 column (Pharmacia) according to the method described previously [2].

2.4. Enzyme assay

The endoglucanase activity of this protein was examined using Avicel SF, carboxymethyl cellulose and xylan as substrates. The enzyme was incubated at 85°C with the substrates in 100 mM phosphate

PII: S0014-5793(99)00257-4

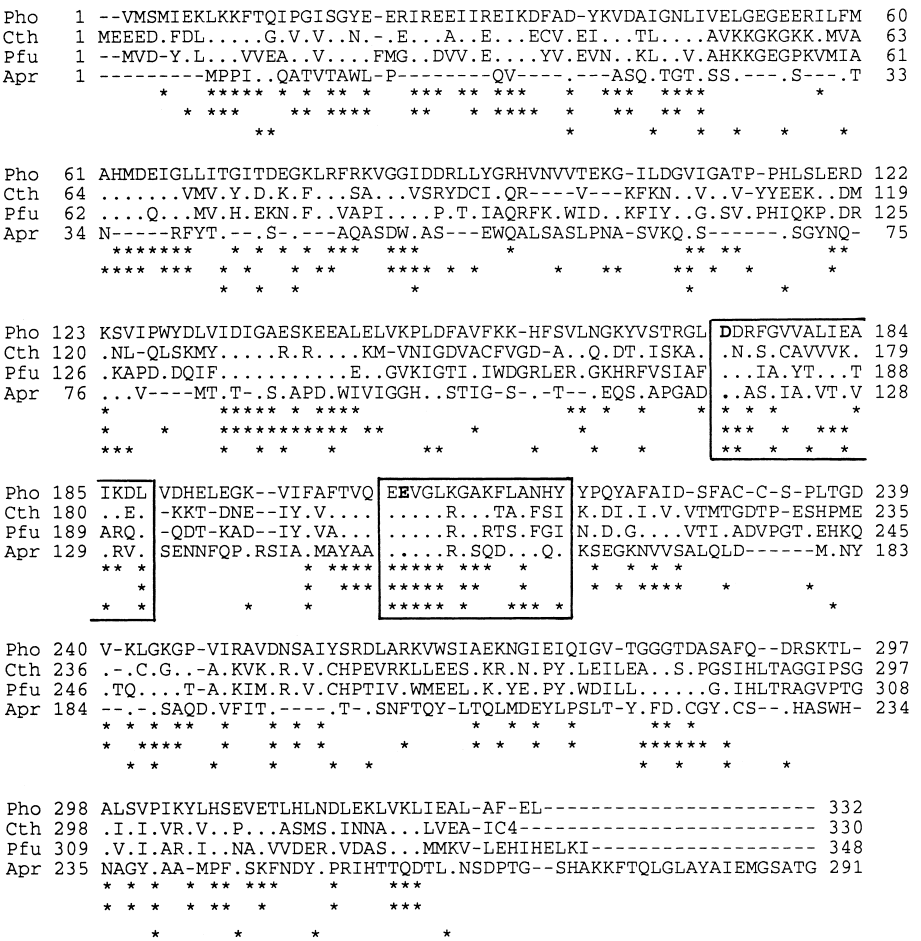

Fig. 1. Comparison of the amino acid sequences of aminopeptidase from *P. horikoshii*, endoglucanase from *C. thermocellum*, deblocking aminopeptidase from *Pyrococcus furiosus* (42% identity) and aminopeptidase from *A. proteolytica*. Sequence alignment was performed by a GenetyxMac program (Software Development, Tokyo, Japan). The sequences have been aligned with dashes indicating gaps. Dots(·) in the Cth, Pfu and Apr sequences indicate the same amino acid as that in the Pho sequence. The residues shown in the bold face indicate the active residues. The regions conserved between the Pho and Apr sequences are indicatted by boxes. The asterisks (*) on the fifth line indicate the residues conserved between the Pho and Cth sequences. The asterisks (*) on the sixth line indicate the residues conserved between the Pho and Pfu sequences. The asterisks (*) on the seventh line indicate the residues conserved between the Pho and Apr sequences. The Cth and Pfu sequences were cited from Swiss-Prot (accession number P55742) and PIR (accession number JC5634), respectively. Abbreviations: Pho, Aminopeptidase from *P. horikoshii*; Cth, endoglucanase from *C. thermocellum*; Pfu, deblocking aminopeptidase from *P. furiosus*; Apr, aminopeptidase from *A. proteolytica*.

buffer (pH 5.8) and the reducing sugar content was assayed by the modified Nelson-Somogyi [8] method using glucose as the standard.

The aminopeptidase activity of this protein was determined using peptides, acylpeptides and acetyl-amino acid-*p*-nitroanilides (Ac-amino acid-pNAs) as substrates. The enzyme was incubated at 85°C with the substrates in 50 mM *N*-ethylmorpholine (NEM) buffer (pH 7.5) containing 100 mM NaCl and 0.5 mM cobalt chloride. The released products were then measured. The activity toward peptides and acylpeptides was measured by detecting the exposed α-amino group using the cadmium-ninhydrin colorimetric method [9]. The activity towards the pNA derivatives was calculated using the absorption coefficient $\varepsilon_{406} = 9.91\ mM^{-1}$ of nitroaniline released [10]. The hydrolysis products were analyzed by paper chromatography [11] and subsequent ninhydrin reaction.

2.5. Analysis of the metal ion bound to the enzyme

The metal ion bound to the enzyme was analyzed using inductively coupled plasma atomic emission spectrometry (ICP-AES) (model SPS1200VR, Seiko Instrument, Tokyo, Japan). Purified enzyme was dialyzed against 50 mM sodium phosphate buffer (pH 7.5). The enzyme sample at the concentration of 0.5 mg/ml was used. The amount of calcium ion was calculated using a calcium standard.

Table 1
Substrate specificity of the aminopeptidase

Substrate (2 mM)	Activity (k_{cat}) (per s^{-1})
Ac-Leu-pNA	0.116 ± 0.002
Ac-Ala-pNA	0.0 22 ± 0.003
Ac-Ala*	$(1.12 \pm 0.09) \times 10^{-4}$
Pyr-Ala*	$(7.96 \pm 0.47) \times 10^{-5}$
Ac-Ala-Ala-Ala	0.034 ± 0.006
f-Ala-Ala-Ala	0.054 ± 0.007
Ala-Ala	0.18 ± 0.03
Ala-Ala-Ala	8.71 ± 0.36
Ala-Pro-Ala	ND

The hydrolytic reaction was measured at 85°C in 50 mM NEM acetate buffer (pH 7.5) containing 0.1 M NaCl and 0.5 mM cobalt chloride. ND, not detected. *, Ala was identified by paper chromatography.

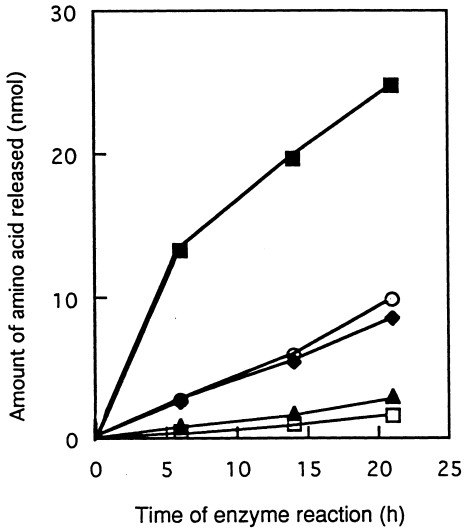

Fig. 2. Amino acid release from the peptide ARGIKGIRGFSG by the aminopeptidase. The reaction was performed at 85°C and [E]/[S]=1/1700. The slightly larger amount of G than that of R may be due to the larger number of residues of G than that of R and the overlapping of the peak of G with the much larger peak of A. Symbols: ■ A, ◆ R, ○ G, ▲ I and □ K.

3. Results and discussion

3.1. Identification of the P. horikoshii gene and expression of the enzyme

In the genome sequence data of *P. horikoshii* [3,4], we found that the open reading frame PH0519, composed of 996 base pairs, encodes a protein homologous both with the endoglucanase from *C. thermocellum* (42% identity), the deblocking aminopeptidase from *P. furiosus* (42% identity) and the aminopeptidase from *A. proteolytica* [5] (18% identity) (Fig. 1). The protein encoded, consisting of 332 amino acids, had two regions conserved between this and the other three enzymes (Fig. 1). After comparison with the crystal structure of the aminopeptidase from *A. proteolytica* [5], we estimated the active sites of the *P. horikoshii* enzyme. Two of the active residues of the former (117D and 152E), which are located in the two conserved regions, correspond to 174D and 206E of the latter (Fig. 1).

The induction of the recombinant protein was performed by cultivating *E. coli* for 5 h at 37°C following the addition of IPTG. Approximately 26 mg of a thermostable 36.9 kDa protein (as determined by SDS-PAGE) was purified from 8 l of culture medium. The molecular weight is consistent with that calculated on the basis of the amino acid sequence. The molecular weight of the protein determined by HPLC was approximately 440 kDa (data not shown). Therefore, this protein is likely to exist as an oligomer with a fairly large number of subunits. The absorption coefficient (A_{280nm}) of the protein at 1% was determined to be 12.43 ± 0.08.

3.2. Substrate specificity of the enzyme

The endoglucanase activity of this protein was examined using Avicel SF, carboxymethylcellulose and xylan as substrates. No activity was detected for these substrates.

To examine the aminopeptidase activity of this protein, we used Ala oligomers (dimer, trimer, tetramer and pentamer), Ala-Pro-Ala, N^{α}-acetyl (Ac-)Leu-pNA, Ac-Ala-pNA, Ac-Ala-Ala-Ala, N^{α}-formyl (f-)Ala-Ala-Ala, Ac-Ala and Pyr-Ala as substrates. The hydrolytic activity of this enzyme in regards to these substrates are summarized in Table 1. At 85°C and pH 7.5, the protein exhibited some hydrolytic activity toward Ac-amino acid pNA derivatives, Ac-Ala-Ala-Ala, f-Ala-Ala-Ala and non-blocked Ala oligomers. Ac-Ala and Pyr-Ala were hydrolyzed very weakly. Ala-Pro-Ala was not hydrolyzed to any appreciable extent. The cleavage pattern of the hydrolysis by this enzyme was confirmed using Ac-LEHD-MCA and ARGIKGIRGFSG as substrates. Hydrolysis of the former (14 h, [E]/[S] = 1/1700) resulted in the most intense spot of Glu followed by a less intense spot of His and a faint spot of Leu detected by paper chromatography and subsequent ninhydrin reaction. The time course of hydrolysis product from the latter substrate was measured by amino acid analysis (Fig. 2). From these results we concluded that this protein is an aminopeptidase with hydrolytic activity to cleave both blocked and non-blocked N-terminal amino acid from pepti-

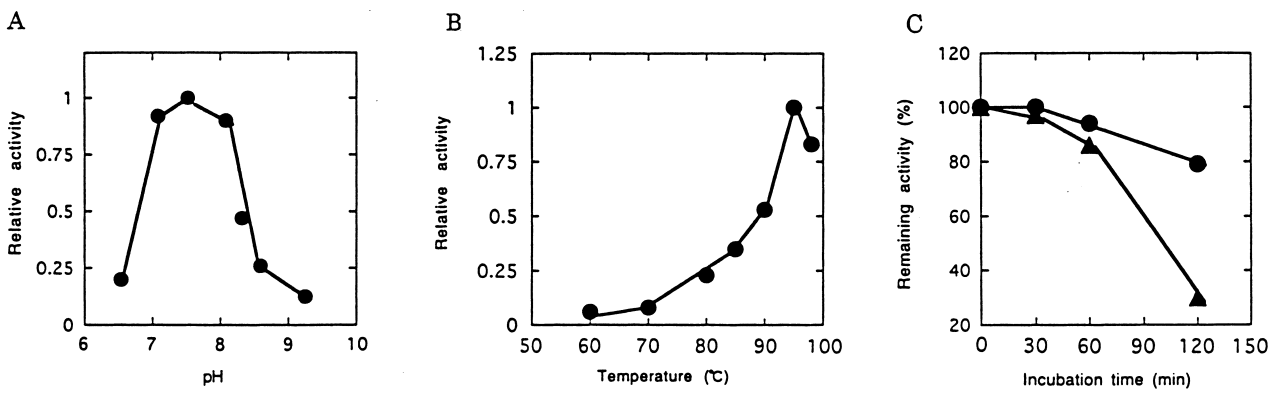

Fig. 3. (A) Effect of pH on the hydrolytic activity of the aminopeptidase on Ala-Ala-Ala. The hydrolytic activity was measured at 85°C in 10 mM sodium phosphate buffer (pH 6.54–7.52) and 10 mM NaH_2PO_4-$Na_2B_4O_7$ buffer (pH 8.08–9.25), containing 0.1 M NaCl and 0.5 mM cobalt chloride. The assay was measured for 10 min. (B) Effect of the temperature on the hydrolytic activity of the aminopeptidase on Ala-Ala-Ala. The hydrolytic activity was measured in 50 mM NEM acetate buffer (pH 7.5) containing 0.1 M NaCl and 0.5% cobalt chloride. The assay was measured for 10 min. (C) Effect of heating on the enzyme activity. One sample of the enzyme (0.1 mg/ml) was incubated at 95°C and a second at 90°C in 50 mM NEM acetate buffer (pH 7.5) containing 0.1 M NaCl. At the time shown, aliquots were taken out and the activity was measured in the same buffer at 85°C using Ala-Ala-Ala as substrate. Symbols: ● and ▲ for the remaining activity of the sample incubated at 90°C and of that incubated at 95°C, respectively.

Table 2
Effect of the substrate length on the aminopeptidase activity

Substrate	K_m (mM)	V_{max} (per s^{-1})
Ala-Ala	39.9 ± 7.3	0.54 ± 004
Ala-Ala-Ala	18.2 ± 1.4	43.2 ± 6.6
Ala-Ala-Ala-Ala	13.2 ± 0.4	89.7 ± 6.6
Ala-Ala-Ala-Ala-Ala	10.8 ± 8.9	160 ± 106

The hydrolytic reaction was measured at 85°C in 50 mM NEM acetate buffer (pH 7.5) containing 0.1 M NaCl and 0.5 mM cobalt chloride.

des, as well as weak activity to cleave the blocking group from the blocked amino acid. The results shown in Fig. 2 also indicate that no significant multiple attacks [12] occur in the hydrolysis by this enzyme. The optimum reaction conditions are approximately pH 7.5 and 95°C (Fig. 3A and B). The deblocking activity was also observed in the aminopeptidase found in *P. furiosus* [13,14].

K_m and V_{max} values were calculated from the measured activity toward Ala oligomers of various lengths (dimer, trimer, tetramer and pentamer), to examine the effect of the length of the substrate on the enzyme activity. The values of K_m decreased, whereas the values of V_{max} increased with the increase in the length of the substrate (Table 2). These results indicate that the active site is composed of at least five subsites.

3.3. Thermostability of the enzyme

After incubating one sample at 95°C and another at 90°C, the relative activity of the enzyme of both samples was measured at 85°C to examine the effect of heating. The enzyme remained fairly stable after incubating at 90°C for at least 2 h as shown in Fig. 3C. The results described above indicate that this enzyme is a hyperthermostable aminopeptidase.

3.4. Analysis of the metal ion bound to the enzyme

The results of the ICP-AES showed that the enzyme is chelated with one calcium ion per monomer of the protein. However, removal of the calcium ion by dialysis for 12 h in 50 mM EDTA did not affect the enzyme activity (data not shown), suggesting that calcium has no effect on the activity. Replacement of calcium with zinc also showed no effect on the activity. The activity was solely dependent on the cobalt ion, although this was not detected by ICP-AES. The activity of this enzyme increases approximately 6-fold with the presence of 0.5 mM cobalt ion, whether the calcium ion is present or not. This suggests that this enzyme has a cobalt binding site which participates in the hydrolytic activity, in addition to the calcium binding site. The dissociation constant of the cobalt ion determined from activity measurement was 0.086 ± 0.004 mM.

The aminopeptidase we reported here has an unique substrate specificity and is expected to be useful in analysis of *N*-blocked peptides. There is no explanation at present why this protein has no endoglucanase activity in spite of its homology with the *Clostridium thermocellum* endoglucanase. Further studies are in progress to clarify the crystal structure, thermostability and hydrolytic mechanism of this enzyme. Structural studies of this aminopeptidase, which has a high homology with endoglucanase, will help the clarification of the molecular evolution of this enzyme.

Acknowledgements: The authors are deeply grateful to Dr Alain Roussel of AFMB-CNRS (Marseille, France) and Dr Susumu Maruyama of the National Institute of Bioscience and Human-Technology for their valuable suggestions and collaboration. Thanks are also due to Ms E. Yamamoto, Ms T. Katakura and Ms N. Nakayama for their technical assistance devoted to this work.

References

[1] Gonzalez, J.M., Masuchi, Y., Robb, F.T., Ammerman, J.W., Maeder, D.L., Yanagibayashi, M., Tamaoka, J. and Kato, C. (1998) Extremophiles 2, 123–130.
[2] Ishikawa, K., Ishida, H., Koyama, Y., Kawarabayasi, Y., Kawahara, J., Matsui, E. and Matsui, I. (1998) J. Biol. Chem. 273, 17726–17731.
[3] Kawarabayasi, Y., Sawada, M., Horikawa, H., Haikawa, Y., Hino, Y., Yamamoto, S., Sekine, M., Baba, S., Kosugi, H., Hosoyama, A., Nagai, Y., Sakai, M., Ogura, K., Otsuka, R., Nakazawa, H., Takamiya, M., Ohfuku, Y., Fuhnahashi, T., Tanaka, T., Kudoh, Y., Yamazaki, J., Kushida, N., Oguchi, A., Aoki, K., Yoshizawa, T., Nakamura, Y., Robb, F.T., Horikoshi, K., Masuchi, Y., Shizuya, H. and Kikuchi, H. (1998) DNA Res. 5, 55–76.
[4] Kawarabayasi, Y., Sawada, M., Horikawa, H., Haikawa, Y., Hino, Y., Yamamoto, S., Sekine, M., Baba, S.-I., Kosugi, H., Hosoyama, A., Nagai, Y., Sakai, M., Ogura, K., Otsuka, R., Nakazawa, H., Takamiya, M., Ohfuku, Y., Fuhnahashi, T., Tanaka, T., Kudoh, Y., Yamazaki, J., Kushida, N., Oguchi, A., Aoki, K., Yoshizawa, T., Nakamura, Y., Robb, F.T., Horikoshi, K., Masuchi, Y., Shizuya, H. and Kikuchi, H. (1998) DNA Res. 5, 147–155.
[5] Chevrier, B., Schalk, C., D'Orchymont, H., Rondeau, J.-M., Moras, D. and Tarnus, C. (1994) Structure 15, 283–291.
[6] Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) J. Mol. Biol. 215, 403–410.
[7] Ho, S.N., Hunt, H.D., Horton, R.M., Pullen, J.K. and Pease, L.R. (1989) Gene 77, 51–59.
[8] Hiromi, K., Takahashi, Y. and Ono, S. (1963) Bull. Chem. Soc. Jpn. 36, 563–569.
[9] Doi, E., Shibata, D. and Matoba, T. (1981) Anal. Biochem. 118, 173–184.
[10] Kobayashi, K. and Smith, J.A. (1987) J. Biol. Chem. 262, 11435–11445.
[11] Gade, W. and Brown, J.L. (1978) J. Biol. Chem. 253, 5012–5018.
[12] Bailey, J.M. and French, D. (1957) J. Biol. Chem. 226, 1–14.
[13] Kamp, R.M., Tsunasawa, S. and Hirano, H. (1998) J. Protein Chem. 17, 512–513.
[14] Tsunasawa, S. (1998) J. Protein Chem. 17, 521–522.