Volume 299, number 1, 80-84 FEBS 10790 March 1992

Nucleotide sequence of the *Staphylococcus aureus* signal peptidase II (*lsp*) gene

Xiao-Jiong Zhao and Henry C. Wu

Department of Microbiology, Uniformed Services University of the Health Sciences, Bethesda, MD 20899-4799, USA

Received 24 December 1991; revised version received 14 January 1992

The *lsp* gene encoding prolipoprotein signal peptidase (signal peptidase II) of *Staphylococcus aureus* was cloned by screening a genomic library for plasmid clones capable of complementing a conditionally lethal *lsp* allele of *Escherichia coli*. *E. coli* cells carrying one of five overlapping clones exhibited increased resistance to globomycin. The nucleotide sequence of the *S. aureus lsp* gene was determined. The deduced amino acid sequence of the signal peptidase II of *S. aureus* suggests that this enzyme has a hydropathy profile very similar to those of *E. coli*, *Enterobacter aerogenes* and *Pseudomonas fluorescens*. Comparison of the primary structures of this enzyme from these four distinct bacterial species reveals three highly conserved domains in proteins which have a low degree of overall sequence homology. Unlike the *lsp* genes from the Gram-negative bacteria, the *lsp* gene in *S. aureus* is not flanked by *x–ileS* and *orf149–orf316* as found in *E. coli*, *Ent. aerogenes*, and *P. fluorescens*.

Signal peptidase II; Prolipoprotein signal peptidase; *Staphylococcus aureus*; *lsp* gene

1. INTRODUCTION

The *slp* genes encoding prolipoprotein signal peptidase or signal peptidase II (SPase II) of *Escherichia coli*, *Enterobacter aerogenes* and *Pseudomonas fluorescens* have been cloned and sequenced [1–4]. The *lsp* genes of these three Gram-negative bacteria are organized into operons consisting of an unknown gene *x*, *ileS* encoding the isoleucyl-tRNA synthetase, *lsp*, and two open reading frames, designated *orf149* and *orf316* [2,3,5]. Lipid-modified proteins which result from the processing of lipoprotein precursors by SPase II have been found in many bacterial species including Gram-positive bacteria such as *Bacillus licheniformis*, *B. cereus* and *Staphylococcus aureus* (reviewed in [6]). In this paper, we report the cloning and sequence determination of the *lsp* gene of *S. aureus*, and a comparison of the deduced amino acid sequences of the SPase II in these four bacterial species.

2. MATERIALS

2.1. *Bacterial strains and medium*

E. coli strains DH5α *supE44 ΔlacU169 (φ80 dlacZΔM15) hsdR17 recA1 gyrA96 thi-1 relA1*, 331c⁻ *ileS*(Ts) [7], Y815 *lsp* [8] and JE5505 *lsp* (a *lsp* transductant of JE5505 constructed by P1 transduction using Y815 as the donor and carrying the same *lsp* allele as that in strain Y815) were used. LB medium, supplemented with ampicillin (10 μg/ml), tetracycline (10 μg/ml), IPTG (6 mM) or X-Gal (40 μg/ml) when appropriate, was ued for the selection or identification of plasmid-containing strains.

A genomic library of *S. aureus* was obtained from Dr. M. Sugai who purchased it from Clontech Laboratories Inc. The cloning vector of the library was pGEM-1, and the host strain was DH5α; partial *Sau*3A digested *S. aureus* DNA was cloned into the *Bam*H1 site of the vector.

2.2. *DNA manipulation and determination and analysis of the DNA sequence*

Plasmid DNA preparation, restriction endonuclease digestion, agarose gel electrophoresis and transformation were performed according to methods described in [9]. The Cyclone I Biosystem (IBI), a system for producing a series of overlapping clones for use in DNA sequencing, was employed. DNA was sequenced with the dideoxy-chain termination method for single-stranded template sequencing [9]. The DNA data were analyzed on a VAX computer with the University of Wisconsin Genetics Computer Group Program Package, version 5 [10], as well as on a personal computer using the Protylyze software program of D. Ward.

2.4. *Assay of SPase II activity*

The assay for SPase II activity was carried out as described by Tokunaga et al. [11].

3. RESULTS AND DISCUSSION

3.1. *Cloning of the* S. aureus lsp *gene*

E. coli strain Y815, a *lsp* (Ts) mutant harboring a plasmid with an IPTG-inducible *lpp* gene, is IPTG-sensitive due to an accumulation of lipid-modified prolipoprotein in the presence of IPTG [8]. To select the *lsp*-containing clone in the genomic library of *S. aureus* the plasmids of the library were transformed into strain Y815; among 80,000 transformants about 80 clones were found to confer an IPTG-resistant phenotype. Six IPTG-resistant clones were chosen for further analysis. These six plasmids which carried the putative *S. aureus*

Correspondence address: X.J. Zhao, Department of Microbiology, Uniformed Services University of the Health Sciences, Bethesda, MD 20899-4799, USA.

 Published by Elsevier Science Publishers B.V.

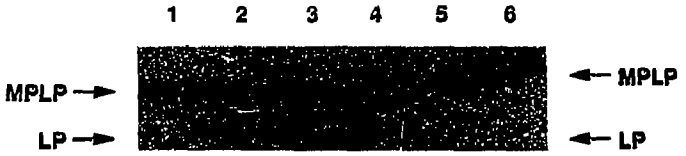

Fig. 1. In vitro assay of SPase II activity in *E. coli lsp* cells harboring plasmid pS4-2.0. JE5505 *lsp* and JE5505 *lsp* (pS4-2.0) cells grown in 0.5 ml LB broth were harvested and resuspended in 50 μl of lysis buffer (0.1% lysozyme, 10 mM EDTA, 20% sucrose, 50 mM TRIS-HCl buffer, pH 8.0, and 10 μg each of RNase and DNAse). After a 15 min incubation at 37°C, 200 μl of 50 mM TRIS-HCl buffer (pH 7.4) containing 0.25% Nikkol was added. 5 μl of this cell lysate was added to 5 μl of [^{35}S]methionine-labeled glyceride-modified prolipoprotein (20,000 cpm) and 15 μl of 50 mM TRIS-HCl buffer (pH 7.4) containing 0.2% Nikkol and 10 mM DTT. The reaction mixtures were incubated for 90 min at 37°C, and analyzed by PAGE using the Ito's gel system [16]. (Lane 1 and 5) Cell extract of JE5505 *lsp*; (lanes 2–4) cell extract of JE5505 *lsp* (pS4-2.0); (lane 6) glyceride-modified prolipoprotein substrate control. MPLP, lipid-modified prolipoprotein; LP, mature lipoprotein.

lsp gene were designated pS1 to pS6, and the sizes of DNA inserts were 3.2, 2.6, 5.9, 3.6, 3.6 and 6.7 kb for pS1 through pS6, respectively. Mapping with various restriction enzymes revealed that pS4 and pS5 were identical. A 2.8 kb *Hin*dIII fragment in the insert of pS4 was isolated and used as a probe for Southern blot analysis of the other four plasmids (pS1–3 and pS6); all these four plasmids hybridized with the 2.8 kb fragment from pS4/pS5, indicating that these pladmids were related and contained common restriction fragments.

To further confirm the presence of the *S. aureus* gene in these plasmids they were transformed into DH5α strain and the resistance of these transformants towards globomycin [12] was determined. All six clones were found to be resistant to globomycin at concentrations exceeding 100 μg/ml, a greater than five-fold increase in globomycin resistance as compared to the control cells. These results suggest that the *lsp* gene of *S. aureus* is expressed in *E. coli* to a significant level so as to confer increased resistance to globomycin. Since the *lsp* and *ileS* genes are linked and presumably co-expressed in three Gram-negative bacterial species so far examined, *E. coli*, *Ent. aerogenes*, and *P. fluorescens*, these six plasmids were transformed into *E. coli* strain 331c$^-$ to ascertain whether they would complement the *ileS*(Ts) mutation in this strain. No complementation of the temperature-sensitive mutation in strain 331c$^-$ was observed with any of these six plasmids. These results indicate that none of these six clonex contained sufficient *ileS* sequence, or the *ileS* gene was not linked to the *lsp* gene in *S. aureus*. It is possible that the *ileS* gene is linked to the *lsp* gene in *S. aureus* as in *E. coli* and other Gram-negative bacteria, but fails to be expressed in *E. coli*; this possibility was ruled out by the sequence data shown below (section 2 and Fig. 2).

Plasmid pS4 DNA was digested with *Pvu*II and re-ligated to yield the plasmid pS4-2.5, which contained a 2.5 kb insert. pS4-2.5 was digested with *Hin*dIII and re-ligated to generate pS4-2.0 which contained a 2.0 kb insert. *E. coli* strain DH5α containing pS4-2.0 showed the expected phenotypes, Amp^r, and increased globomycin resistance. Further evidence for the expression of the *S. aureus lsp* gene on plasmid pS4-2.0 in *E. coli* strain JE5505 *lsp* lacking this enzyme was provided by an in vitro assay of SPase II as shown in Fig. 1.

3.2. *Nucleotide sequence analyses of pS4-2.0 subclones*

The 2.0 kb fragment was excised from plasmid pS4-2.0 and cloned into both M13mp18 and M13mp19 to generate pSM18-2.0 and pSM19-2.0 in order to sequence the 2.0 kb insert in both directions. Overlapping clones were obtained using the Cyclone I Biosystem; fifteen M13mp18 and twelve M13mp19 overlapping clones were sequenced which covered the entire 2.0 kb sequence in both directions. The insert in pS4-2.0 was found to consist of 2,002 bp (Fig. 2). Two open reading frames were present which could potentially encode polypeptides comprising 163 and 94 amino acids, respectively. Based on the deduced amino acid sequence, we designated the first open reading frame (nucleotide 1213–1701) the *lsp* gene, the gene product of which is similar in size and overall structure (especially the hydropathy profile) to the *E. coli* enzyme. The deduced amino acid sequence of the putative *S. aureus* SPase II indicates that this enzyme has a calculated molecular mass of 18,331, and has a low overall homology with the *E. coli* enzyme. The % identity (and similarity) of amino acid residues of SPase II's from *E. coli*, *Ent. aerogenes*, *P. fluorescens* and *S. aureus* were found to be 100 (100), 91 (96), 41 (63) and 34% (63%), respectively, and the corresponding nucleotide sequence homologies for these four *lsp* genes were 100, 88, 53 and 41%, respectively. The second open reading frame (not translated in Fig. 2) immediately follows the *lsp* gene, and its putatitve initiation codon overlaps with the termination codon of the *lsp* gene; the deduced amino acid sequence of this hydrophilic protein lacks sequence and topological homology with the *E. coli lsp* gene product. A closer examination of the nucleotide sequence of the 2.0 kb insert failed to reveal the presence of genes allelic to genes *x*, *ileS*, *orf149* and *orf316*. According to the sequence analysis, the *lsp* gene of *S. aureus* was cloned in the orientation opposite to the SP6 promoter on pGEM-1 vector and the T7 promoter was deleted during the subcloning. We inver that the promoter(s) for the transcription of the *lsp* gene is present in the insert

```
   1  CTGTAAAATG TCCTTGAATA GCATGGAAAG TTTCTTTAAA TCTTTCAGTA ACCTCTTGAT
  61  CGTCAAGATT TCACACAGTT ACAGCTATTA GCGCGTGAAC AGGTTCTTTA TATCGAAGAG
 121  CCATTTAAAG ATATCTCCAT GCTTGATGAA GTAGTGGATG GAACAATACC TCCAATAGCC
 181  CTTGATGAAA AGGCAACATC GTTGTTGGAC ATAATTAATT TGATAGAATT ATACAATGTG
 241  AAAGTTGTTG TATTAAAACC ATTTCGTCTA GGTGGCATTG ATAAGTGCAA ACGGCAATTG
 301  ATACTTTGAA AAGTCATGGT GTAAAAGTGG TTATTGGCGG CATGTACGAT ATGGTTTAAG
 361  CCGTTATTTT ACAGCGATGC TTGCTCGTAA AGGTGATTAT CCAGGAGATG TTACACCAGC
 421  AGGTTATTAT TTTGATCTTT TTATACAAGT TATGAAGGAG GCTGGGACAT TAAGTTCTTA
 481  GGCAATGTAA AAAGCTGATT TCTATTAATT ATTTGATAGA AATCAGCTTT TTTGATATGT
 541  ATTTTATAAT GTACAGCTCG TTGAGCAGCT ATTTTCCTTA TATTAAGTGC CATTAATACA
 601  AAACCTAGCT CTCGTTTAAC TTTATTTATT CCTCGAACTG ACATTCGAGT GAAACTCAAA
 661  ATAGCCTACT TTCTTAAATT AACAATATCT ATTCTCATAG AATTTGTCCA ATTAAGTGTA
 721  GACGATTCAT GTCTTATTTT TTTAAAGTAT TTAAAAGTGA AATTACATGT TAATACGTAG
 781  TATTAATGGC GAGACTCCTG AGGGAGCAGT GCCAGTCGAA GACCGAGGCT GAGACGGCAC
 841  CCTAGGAAGG GACCCATCAT CAAAAATTCT ATTTATAGAA TTTTACAGTT TAGTGCCAGA
 901  TGGGCATAGC GAAGCCATTC AATACGAAGT ATTGTATAAA TAGAGAACAG CAGTAAGATA
 961  ATTTTTAATT AGAAAATATC TTACTACTGT CTTTTTAGGG ATTTATGTTC CAGCCTGCTT
1021  TCCTAATTTT TAATGTCATC TTAAATATGA TAAATGAATA ATTAAGTTCA TATTTAATGT
1081  CAAAACATAG TAGTTTATCA AGTATTGAGT GAGTAACATT AGATTTAATG TAATATCGTT

1141  ACTTTTTTTA TTAGCAGGTG TAAGCTATAA TATAAAGAGT TGTCTTATGG ACGATTGATT

1201  GGAGGAACGA AAATGCACAA AAAATATTTT ATTGGCACTT CCATTTTAAT AGCAGTATTT
                    MetHisLys LysTyrPhe IleGlyThr SerIleLeuIle AlaValPhe

1261  GTCGTTATAT TTGACCAAGT TACTAAATAT ATTATAGCTA CTACAATGAA AATTGGAGAT
      ValValIle PheAspGlnVal ThrLysTyr IleIleAla ThrThrMetLys IleGlyAsp

1321  TCATTTGAAG TGATACCGCA CTTTTTAAAC ATAACATCAC ATCGAAATAA TGGTGCTGCA
      SerPheGlu ValIleProHis PheLeuAsn IleThrSer HisArgAsnAsn GlyAlaAla

1381  TGGGGAATAT TGAGTGGAAA AATGACATTT TTCTTTATTA TTACCATTAT TATATTAATA
      TrpGlyIle LeuSerGlyLys MetThrPhe PhePheIle IleThrIleIle IleLeuIle

1441  GCCTTAGTAT ATTTCTTTAT TAAAGATGCT CAATATAATT TGTTTATGCA AGTTGCTATT
      AlaLeuVal TyrPhePheIle LysAspAla GlnTyrAsn LeuPheMetGln ValAlaIle

1501  AGTTTACTTT TTGCAGGTGC ACTTGGAAAC TTTATTGATA GAATTTTAAC AGGAGAAGTT
      SerLeuLeu PheAlaGlyAla LeuGlyAsn PheIleAsp ArgIleLeuThr GlyGluVal

1561  GTTGACTTTA TTGATACAAA TATTTTTGGT TATGATTTTC CAATATTTAA TATAGCAGAT
      ValAspPhe IleAspThrAsn IlePheGly TyrAspPhe ProIlePheAsn IleAlaAsp

1621  TCAAGTTTAA CAATTGGTGT AATATTAATT ATTATTGCCT TATTAAAGGA TACTTCCAAT
      SerSerLeu ThrIleGlyVal IleLeuIle IleIleAla LeuLeuLysAsp ThrSerAsn

1681  AAAAAGGATA AGGAGGTTAA GTAATGGAGA CTTATGAATT TAACATTACA GATAAAGAAC
      LysLysAsp LysGluValLys End

1741  AAACAGGTAT GCGTGTAGAT AAGTTGCTGC CTGAATTAAA TAATGATTGG TCTCGTAACC
1801  AGATACAAGA TTGGATTAAA GCAGGTTTAG TCGTTGCAAA CGATAAAGTT GTTAAATCTA
1861  ATTATAAAGT GAAACTTAAT GATCCATACT ACACGAATAA TTTTGAAGGT GTATACGACA
1921  TGGTATTATC ATCTTGTGAT AATTTAATAG ACTACATCGT AAAAGATGCA AATTTGAAAG
1981  AGGGGTAGTT TTTTATGGAA AA
```

Fig. 2. Nucleotide sequence of the *lsp* gene and its flanking sequences in *S. aureus*. The deduced amino acid sequence of the SPase II is shown. The putative Shine–Delgarno sequence for the translation of the *lsp* gene is underlined. The sequence data in this figure will appear in the EMBL/GenBank nucleotide sequence data base under the accession number M84707.

of pS4-2.0. A putative ribosome-binding site was found 5′ to the initiation codon of the *lsp* gene (Fig. 2), but the putative promoter(s) for the *lsp* gene remains to be identified.

The hydropathy profile of *S. aureus* SPase II as examined by the Kyte and Doolittle algorithm [13] contains four hydrophobic segments (A–D in Fig. 3). These regions, which are comprised mainly of hydrophobic residues, presumably represent the transmembrane regions similar to those postulated for the *E. coli* enzyme [14]. Three regions of highly conserved amino acid sequences were identified (I, II and III in Fig. 3), which are further refinements of those identified previously based on an comparison of the primary structures of SPase II's from *E. coli, Ent. aerogenes* and *P. fluorescens*. The existence of such highly conserved domains in protein sequences of otherwise relatively low overall homology suggests that they may be essential for the catalytic activity of this enzyme. In addition, these conserved sequences might be exploited for the purpose of cloning the *lsp* gene from a diverse group of bacteria using the polymerase chain reaction.

```
          1                A                                      50
Stapep    MHKKYFI.GT SI.LIAVFVV ..IFDQVTKY IIATTMKIGD SFEVIPHFLN
Psfpep    MPNADSRFGR LGWLVLSLLV LVI.DQVSKA HFEGSLEMFQ QIVVIPDYFS
Entpep    MSKSICSTG. LRWLWVVVAV LII.DLGSKF LILQNFALGE TVSLFPS.LN
Ecopep    MSQSICSTG. LRWLWLVVVV LII.DLGSKY LILQNFALGD TVPLFPS.LN
Lineup    MsksicstG. lrwLwlvvvV lii.D..sKy lilqnfalgd tv...Ps.ln

          51                              B                100
Stapep    ITSHRNNGAA WGILSGKMTF FFIITIIILI ALVYFFIKDA QYNLFMQV..
Psfpep    WTLAYNTGAA FSFLADGGGW QRWLFAVIAV VVSAVLVVWL KRLGRDDTWL
Entpep    LHYARNYGAA FSFLADSGGW QRWFFAGIAV GICVVLAVLM YRSKATQKLN
Ecopep    LHYARNYGAA FSFLADSGGW QRWFFAGIAI GISVILAVMM YRSKATQKLN
Lineup    l.yarNyGAA fsfLadsggw qrwffagIa. gisvvlav.m yrskatqkln
                    I
          101     C                                        150
Stapep    AISL.LFA.G ALGNFIDRIL TGEVVDFIDT NIFGYD.FPI FNIADSSLTI
Psfpep    AIALALVLGG ALGNLYDRIA LGHVIDFILV HWQNRHYFPA FNFADSAITV
Entpep    NIAYALIIGG ALGNLFDRLW HGFVVDMIDF YV.GDWHFAT FNLADSAICI
Ecopep    NIAYALIIGG ALGNLFDRLW HGFVVDMIDF YV.GDWHFAT FNLADTAICV
Lineup    .Ia.aLiigG AlGNlfDR.w hGfVvD.Idf yv.gdwhF.t FNlADsai..
                      II                                  III
          151  D                  174
Stapep    GVILIIIALL KDTSNKKDKE VK
Psfpep    GA..IMLALD MFKS.KKTGE TVND
Entpep    GAALIVLEGF LPSSDKKTS
Ecopep    GAALIVLEGF LPSRAKKQ
Lineup    GaaLIvl..f lpss.KKt.e ....
```

Fig. 3. Comparison of the deduced amino acid sequences of SPase II from *S. aureus* (Stapep), *P. fluorescens* (Psfpep), *Ent. aerogenes* (Entpepl), and *E. coli* (Ecopep). In the consensus (lineup) rows, periods indicate lack of identity, lowercase letters indicate identity in at least one of the four sequences, and uppercase letters indeiate identity in all four sequences. The postulated transmembrane domains for *S. aureus* SPase II (A–D) are overlined. The three highly conserved domains (I–III) are underlined.

A proposed model for the topology of *S. aureus* SPase II is depicted in Fig. 4. While it differs significantly from that of the *E. coli* enzyme in both the lengths and amino acid sequences of the hydrophilic, presumptive cytosolic and periplasmic loops, it shares with the other three enzymes from the Gram-negative bacteria the preservation of a basic topological structure, i.e. four hydrophobic membrane-spanning segments and highly positively charged N- and C-terminal segments. It is highly likely that these basic features contribute to the topology of this enzyme in the membrane, and allow the highly conserved regions in this enzyme to participate in the unique catalysis of proteolytic processing of lipid-modified prolipoproteins by this enzyme [15]. The importance of the C-terminal tail of this enzyme is supported by the observation that mutations affecting this region significantly reduce the activity of this enzyme [1].

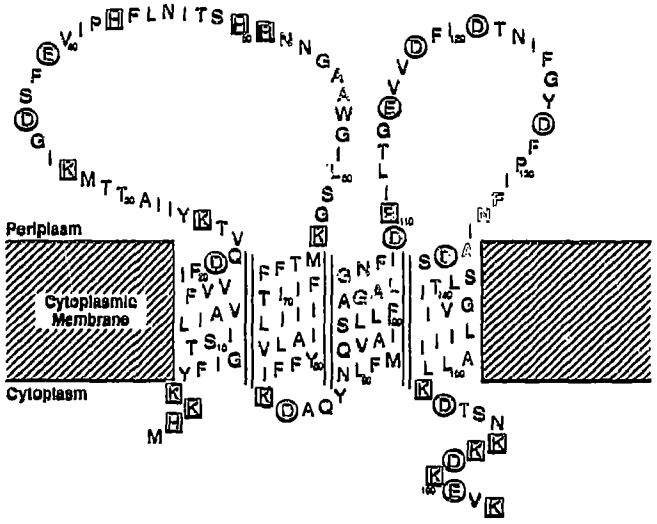

Fig. 4. A proposed model for the topology of *S. aureus* SPase II in the cytoplasmic membrane. The charged amino acid residues (R, K, H, D and E) are highlighted. The three regions of high sequence homology are presented as hollow letters.

Acknowledgements: This work was supported by NIH Grant GM28811. We thank Dr. M. Sugai for making the genomic library of *S. aureus* available to us.

REFERENCES

[1] Innis, M.A., Tokunaga, M., Williams, M.E., Loranger, J.M., Chang, S.Y., Chang, S. and Wu, H.C. (1984) Proc. Natl. Acad. Sci. USA 81, 3708–3712.
[2] Isaki, L., Kawakami, M., Beers, R., Hom, R. and Wu, H.C. (1990) J. Bacteriol. 172, 469–472.
[3] Isaki, L., Beers, R. and Wu, H.C. (1990) J. Bacteriol. 172, 6512–6517.
[4] Yu, F., Yamada, H., Daishima, K. and Mizushima, S. (1984) FEBS Lett. 173, 264–268.
[5] Miller, K.W., Bouvier, J., Stragier, P. and Wu, H.C. (1987) J. Biol. Chem. 262, 7391–7397.
[6] Hayashi, S. and Wu, H.C. (1990) J. Bioenerget. Biomemb. 22, 451–471.
[7] Isaksson, L., Slold, S.-E., Skjoldebrand, J. and Takata, R. (1977) Mol. Gen. Genet. 156, 233–237.
[8] Yamagata, H., Ipplito, C., Inukai, M. and Inouye, M. (1982) J. Bacteriol. 152, 1163–1168.

[9] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning. A Laboratory Manual, Cold Spring Harbor Laboratory Press, New York.
[10] Devereux, J., Haeberli, P., Smithies, O. (1984) Nucleic Acids Res. 12, 387–395.
[11] Tokunaga, M., Loranger, J.M. and Wu, H.C. (1983) J. Biol. Chem. 258, 12102–12105.
[12] Inukai, M., Takeuchi, M., Shimizu, K. and Arai, M. (1978) J. Antibiot. 31, 1203–1205.
[13] Kyte, J. and Doolittle, R.F. (1982) J. Mol. Biol. 157, 105–132.
[14] Muñoa, F., Miller, K.M., Beers, R., Graham, M. and Wu, H.C. (1991) J. Biol. Chem. 266, 17667–17672.
[15] Hussain, M., Ichihara, S. and Mizushima, S. (1980) J. Biol. Chem. 255, 3707–3712.
[16] Ito, K., Date, T. and Wickner, W. (1980) J. Biol. Chem. 255, 2123–2130.