Volume 286, number 1,2, 204–208 FEBS 09946 July 1991

ADONIS 001457939100655C

Cloning, overexpression, purification and crystallisation of ribosomal protein L9 from *Bacillus stearothermophilus*

Constantin E. Vorgias, Adrian J. Kingswell, Zbigniew Dauter and Keith S. Wilson

European Molecular Biology Laboratory (EMBL), c/o DESY, Notkestrasse 85, 2000 Hamburg 52, Germany

Received 17 April 1991; revised version received 16 May 1991

The cloning, sequencing and overexpression of the gene coding for *Bacillus stearothermophilus* ribosomal protein L9 is described. The sequence corresponds directly to that presented for the protein itself by classical methods, differing at only a few amino acid positions. The purification and crystallisation of the corresponding L9 protein is presented. The crystals are isomorphous to those described for L9 obtained by conventional methods.

Bacillus stearothermophilus; Ribosomal protein; Gene sequence; Crystallisation

1. INTRODUCTION

We are studying the non-sequence-specific nucleic acid binding proteins of *Bacillus stearothermophilus (Bst)*. During the course of purification of the DNA binding protein HU a major protein band was identified on a cation exchange column just after the HU peak, with molecular weight of 17 kDa. Our interest in nucleic acids binding proteins led us to investigate this protein. The apparent molecular weight as well as the basic character of this protein initially suggested that it was the *Bst* homologue of the 17K DNA-binding protein of *E. coli* [1,2]. We therefore proceeded to study this protein in more detail.

More recently, the determination of the first 25 N-terminal amino acid residues have shown that the protein is in reality ribosomal protein L9 from the large subunit of the *Bst* ribosomes. The protein was originally given the number L17 from its position on 2-dimensional electrophoresis gels [3] and its purification from *Bst* ribosomes and crystallisation were described [4]. At that time the sequence was shown not to be homologous to those of any of the *E. coli* proteins for which a complete sequence was available. Since then no further studies on this protein have been published. The sequence has been shown to be homologous to that subsequently determined for *E. coli* L9 and we refer to the protein as L9 throughout this paper.

During the past 20 years the procaryotic ribosome has been the subject of intensive study. To understand the function of the ribosomal machinery at the molecular level it is essential to gain a detailed information of the structure of its individual components. A variety of different techniques have been applied, including: circular dichroism [5]; proton magnetic resonance [6]; microcalorimetry [7,8]; immune electron microscopy [9]; chemical cross-linking [10]; fluorescence spectroscopy [11]; and reconstitution of ribosomal subunits from their components [12]. These methods have provided the bulk of the information about the relative topology of the ribosomal proteins in the organelle. The *E. coli* ribosomal proteins were studied first, and all 53 ribosomal proteins have been sequenced. However, the *E. coli* ribosomal proteins have in general not produced good crystals for crystallographic studies, in spite of intensive efforts in a number of laboratories.

The thermophilic bacterium *Bst* was early recognized as an excellent source of thermostable proteins. The fact that *Bst* proteins have been naturally selected to be stable at high temperature is reflected in a resistance to various denaturing conditions, and a resulting increase in the probability of crystallisation. This is particularly important for example with ribosomal proteins, which have been isolated from their normal environment in the cell where they probably rely on other proteins and rRNA for their stability.

The crystallisation of a protein can be a time consuming process. Moreover the crystallisation trials require a significant amount of highly purified protein. Therefore we proceeded to clone and overexpress the L9 protein from *Bst*. We used the in vitro gene amplification technique and cloned the *b19* gene which encodes the protein. The *E. coli* BL21(DE3) transformant harbouring the *b19* gene on the pET-3a plasmid was shown to overproduce the target gene at a very high level. The recombinant BL9 protein was subsequently purified and good quality stable crystals were obtained. Preliminary crystallographic data are presented.

Correspondence address: C.E. Vorgias, European Molecular Biology Laboratory (EMBL), c/o DESY, Notkestrasse 85, 2000 Hamburg 52, Germany. Fax: (49) (40) 89080149.

Published by Elsevier Science Publishers B.V.

2. EXPERIMENTAL PROCEDURES

2.1. *Materials*

The bacterial strains used are the *E. coli* K12 strain HMS174 (F^- *recA* $r^-_{k12}m^+_{k12}Rif^R$) and the B strain BL21 (F^- *ompT* $r^-_Bm^-_B$). The set of pET-3 expression vectors, as well as the host cells, were kindly provided by Dr F.W. Studier (Brookhaven National Laboratory, Upton, NY) [14,15].

The enzymes used for the cloning procedures and the pUC cloning system were from Boehringer Mannheim. The sequencing kit Sequenase was purchased from United States Biochemical (USB). The gene amplification kit was supplied by Perkin-Elmer (Cetus). The column chromatography media were from Pharmacia/LKB and all other chemicals were either from Sigma or from Merck, in the highest analytical grade. Synthetic oligonucleotides were prepared at the central EMBL DNA synthesis laboratory and purified by HPLC (Synchropak 300 C_{18} 6.5 μm, MZ Analysentechnik).

2.2. *In vitro amplification of the* b19 *gene from* Bst *total chromosomal DNA*

Total *Bst* chromosomal DNA was purified as described in [13]. In a typical 50 μl reaction the following components were mixed at the final concentration: *Bst* template DNA 20 ng/μl; primers #N, #C, 1 μM of each; 200 μM of each dNTP; Amplitaq DNA polymerase 0.025 U/μl, in reaction buffer consisting of 10 mM Tris-HCl, pH 8.3 (at 25°C), 50 mM KCl, 0.01% (w/v) gelatin, 0.5-4 mM total $MgCl_2$. The gene amplification reaction was performed in the following steps per cycle: denaturation at 94°C for 1 min; annealing at 50°C for 1 min; and polymerization at 72°C for 2 min; for 30 cycles, using a Hybaid Biometra device. One-tenth of the amplification mixture was normally analyzed by 7% acrylamide/0.18% bis-acrylamide gel electrophoresis in TBE buffer (45 mM Tris-borate, pH 8.0, 1 mM ethylenediamine tetra-acetic acid, EDTA) at 100-200 V constant voltage and stained with ethidium bromide.

2.3. *Cloning of the* b19 *gene*

All further cloning procedures were carried out essentially as described in [13] or according to the manufacturers. The plasmids were purified using the Qiagen plasmid preparation kit (Diagen) and DNA fragments were extracted from agarose gels using the Mermaid kit from Dianova.

In brief, the amplified *b19* gene was inserted into the *Hin*dII linearized pUC19 vector by blunt-end ligation. Four individual positive clones were sequenced (direct and reverse) using the chain termination method [16]. The ca 470 bp *Bam*HI/*Nde*I DNA fragment, which contains the entire *b19* gene, was directly cloned into the pET-3a vector (amp^r and unique *Bam*HI,*Nde*I restriction sites) [15], resulting in a pCV*b19* expression plasmid. The pCV*b19* plasmid was used to transform the BL21(DE3). The transformation was carried out with the $CaCl_2$ method [17].

The BL21(DE3) *E. coli* cells, harbouring the pCV*b19* plasmid, were grown in LB medium overnight in the presence of 25 μg/ml ampicillin. Next day, the cells were diluted to a final $OD_{600nm} = 0.3$ and the overexpression of the target gene was induced by adding isopropyl-β-D-thiogalactoside (IPTG) to 0.4 mM, final concentration. At 0, 1 and 3 h after induction, samples were withdrawn and analyzed by the 12.5% sodium dodecylsulfate polyacrylamide gel electrophoresis (SDS-PAGE) Laemmli system [18] and stained with Coomassie blue G-250.

2.4. *Overexpression and purification of the BL9 protein*

In a routine protein preparation a 2 litre cell culture induced for 3 h was used. The induced bacteria were collected by low speed centrifugation and washed once with cold buffer W (50 mM Tris-HCl, pH 7.5, 100 mM NaCl, 0.1 mM phenylmethylsulphonyl fluoride, PMSF). All further procedures were carried out at 0-4°C, except where otherwise specified. The bacterial paste was resuspended in 5 ml/g buffer A (10 mM Na-phosphate, pH 7.0, 1 mM EDTA, 0.1 mM PMSF) containing 0.1% (w/v) Triton X-100. The cells were disrupted by sonication for 10 min. The total extract was clarified by centrifugation in an SS-34 rotor (Sorval) at 20000 rpm, for 20 min. The supernatant was adjusted to 30% saturation in ammonium sulfate. After 30 min stirring, the non-precipitated material was separated by centrifugation as above. The soluble supernatant was dialysed against buffer A overnight. The dialysed BL9 protein fraction was applied to a 20 ml CM-Sepharose CL-6B column. Bound proteins were eluted with a 400 ml NaCl (0-0.5 M) linear gradient. The collected fractions were analysed by 12.5% SDS-PAGE. The BL9 protein fractions were combined and diluted 1:2 with buffer B (10 mM Na-phosphate, pH 6.0, 1 mM EDTA, 0.1 mM PMSF). The pH was adjusted to 6.0 and the protein solution was applied, portionwise, to a Mono-S FPLC column. The BL9 protein was eluted with a 200-250 mM NaCl linear short gradient. The final purification step gave very pure and concentrated protein material ready for crystallisation. The protein concentration was determined by the Bradford method as described in [19].

2.5. *Crystallisation*

Crystals of BL9 were grown by 'hanging-drop' vapor diffusion [20]. 10 mg/ml solution of BL9 protein was dialysed against 1.5 M ammonium sulfate, 50 mM sodium phosphate, pH 7.5. 10 μl hanging drops of this solution were suspended above reservoirs containing 1.75-2.0 M ammonium sulfate. Crystals appeared after 1 week at 5°C. The conditions described are similar to those in [9]. We also used 60-80% saturated sodium citrate, pH 6.4 as precipitant.

3. RESULTS AND DISCUSSION

3.1. *Cloning and overexpression of the* b19 *gene in E. coli*

The known primary structure of the BL9 protein [4] was used to design both primers in order to copy and amplify the corresponding gene from the total chromosomal DNA of *Bacillus stearothermophilus*.

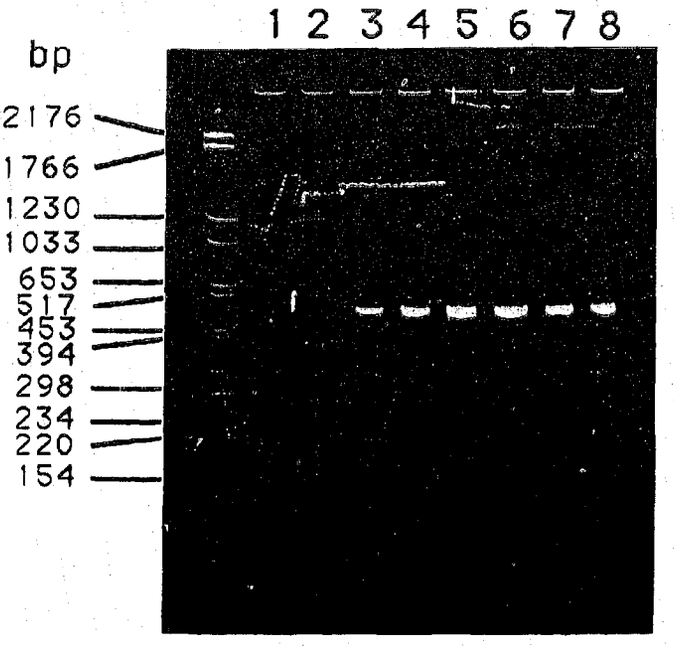

Fig. 1. Amplification of *b19* target gene using the Taq polymerase reaction (PCR) from total chromosomal DNA of *Bst*. Slots 1-8 correspond to 0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0 mM total Mg^{2+} in the reaction mixture. One-tenth of the reaction products was examined by 7% PAGE in TBE buffer and stained with ethidium bromide (for experimental details see under methods).

```
1/1                                                    31/11
ATG AAG GTT ATT TTT TTA AAG GAT GTC AAA GGA AAA GGG AAA AAA GGG GAA ATC AAA AAT
Met lys val ile phe leu lys asp val lys gly lys gly lys lys gly glu ile lys asn
61/21                                                  91/31
GTC GCC GAC GGC TAT GCC AAC AAC TTC TTA TTT AAA CAA GGG CTG GCC ATT GAA GCG ACG
val ala asp gly tyr ala asn asn phe leu phe lys gln gly leu ala ile glu ala thr
121/41                                                 151/51
CCA GCC AAT TTA AAA GCG CTC GAG GCG CAA AAA CAA AAA GAG CAG CGC CAG GCG GCC GAG
pro ala asn leu lys ala leu glu ala gln lys gln lys glu gln arg gln ala ala glu
181/61                                                 211/71
GAG CTG GCG AAT GCG AAA AAA TTG AAA GAA CAG CTT GAG AAG CTG ACA GTA ACC ATT CCA
glu leu ala asn ala lys lys leu lys glu gln leu glu lys leu thr val thr ile pro
241/81                                                 271/91
GCA AAG GCA GGC GAA GGC GGC CGT TTG TTC GGC TCG ATT ACG AGC AAG CAA ATC GCC GAG
ala lys ala gly glu gly gly arg leu phe gly ser ile thr ser lys gln ile ala glu
301/101                                                331/111
TCG CTG CAA GCC CAA CAC GGC TTG AAG CTC GAC AAG CGC AAA ATT GAG CTT GCC GAT GCC
ser leu gln ala gln his gly leu lys leu asp lys arg lys ile glu leu ala asp ala
361/121                                                391/131
ATT CGC GCG CTT GGA TAC ACG AAC GTG CCG GTG AAA CTC CAC CCA GAG GTA ACG GCG ACG
ile arg ala leu gly tyr thr asn val pro val lys leu his pro glu val thr ala thr
421/141
CTG AAA GTC CAT GTG ACC GAG CAG
leu lys val his val thr glu gln
```

Fig. 2. Nucleotide sequence and the deduced amino acid sequence of the *bl9* gene amplified from the *Bst* total chromosomal DNA and cloned in the pET-3a vector.

The amplification reaction was carried out as above and the amplified gene was detected by gel electrophoresis. As shown in Fig. 1 the best amplification results were found between 2.5 and 3 mM total Mg^{2+} in the reaction mixture. The amount of the produced (copied and amplified) gene, as well as the specificity of the amplification product provided an excellent starting material for further cloning procedures. Both primers were designed to build an *Nde*I restriction site at the 5′-end as well as a *Bam*HI site next after the stop codon at the 3′-end of the copied gene. The amplified gene was then introduced into the *Hin*dII site of the pUC19 plasmid by blunt-end ligation and verified by DNA sequencing. The sequence of the *bl9* gene and that deduced by amino acid sequence of the protein are shown in Fig. 2. There are three differences from the primary structure of the BL9 protein determined by classical peptide sequencing [3]. The sequence Asn^{27}-Asn^{28}-Ser^{29} in the gene sequence is Asn-Phe in the peptide sequence, i.e. one Asn is missing from the peptide sequence and the Ser is replaced by a Phe. The peptide sequence corresponding to Glu^{60}-Glu^{61}-Leu^{62} in the gene is Glu-Leu, i.e. one Glu is missing from the peptide sequence. Finally at the C-terminus of the protein we could not read the Lys residue, although is has been designed to be there on the 3′-primer. In order to avoid ambiguities from the amplification reaction we have cloned and sequenced the *bl9* gene from two individual amplification reactions with identical results.

The coding sequencing for the BL9 protein was introduced into the expression vector pET-3a deriving the pCV*bl9* plasmid [15]. The pET-3a expression vector has the *Nde*I unique site located at the translation start site and therefore can direct the production of native proteins. The expression vector contains a strong T7 promoter, the T7 gene *10* ribosome binding site and ϕ terminator (Tϕ) which may make the transcripts more resistant to exonucleolytic degradation [14]. The pCV*bl9* plasmid was used to transform the *E. coli* BL21(DE3) strain which is a lysogen bearing the T7 RNA polymerase gene on the host chromosome under the control of the *lac*UV5 promoter. The transformants were induced with IPTG for up to three hours. Fig. 3 shows a 12.5% SDS-PAGE of the induction protein products. The overproduction of the BL9 protein is clearly shown since this protein is about 20-25% of the total host protein.

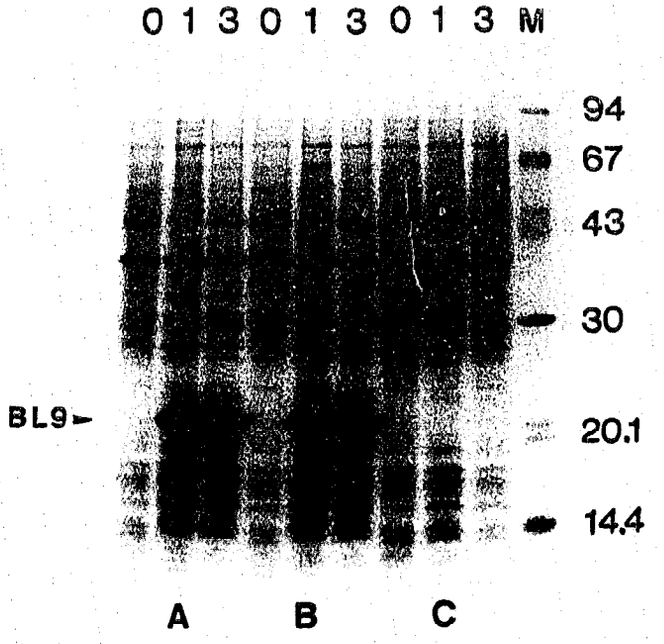

Fig. 3. 12.5% SDS-PAGE of the total bacterial proteins of the BL21(DE3) cell line harbouring the pCV*bl* 17 plasmid. 0, 1, 3 are the time intervals, in hours, after adding IPTG (0.4 mM final concentration) to induce the overexpression of the *bl9* target gene under the control of the T7 RNA polymerase ø *10* promoter. A and B are two different individual clones expressing the *bl9* target gene and C is the control (no target gene).

Fig. 4. Crystals of BL9 protein obtained at 5°C, over one week from 1.75–2.0 M ammonium sulfate. The dimensions of the crystal are 2 mm × 0.5 mm × 1.2 mm. (Bar length: 0.25 mm)

3.2. *Purification of the overexpressed BL9 protein*

The extraction of the BL9 protein overproduced in *E. coli* was made in a low ionic strength Tris buffer at neutral pH by physical disruption of the cells. An ammonium sulfate precipitation up to 30% was used to remove subcellular debris and gave a significant enrichment (40% of the total extractable proteins) in BL9 protein. BL9 was further purified by cation-exchange column chromatography on a CM-Sepharose CL-6B column at pH 7.0. BL9 was bound to the CM column and eluted as a symmetrical peak between 170 and 240 mM NaCl. The first chromatographic step increased the purity of BL9 protein to about 90%. The final purification and simultaneous concentration step was made on a Mono-S FPLC column at pH 6.0. A routine preparation of 5 g bacteria (2 litre culture) takes 2–3 days and gives at least 50 mg BL9. The BL9 protein is more than 95% pure, and no other protein bands can be detected in a heavily overloaded acrylamide gel stained with Coomassie blue.

3.3. *Crystallisation and data collection*

Initial crystallisation conditions were identical to those described in [4]. However, recently we found that the use of 60–80% saturated sodium citrate pH 6.4, as precipitant significantly decreases the time for full crystal formation and gave an increased number of useful crystals. The crystals obtained were up to 2 mm long as shown in Fig. 4. The crystals have been characterised using an imaging plate scanner mounted on a conventional sealed-tube Seifert X-ray source running at 2 kW. The space group and cell dimensions are shown in Table I, and are essentially identical to those reported for the non-cloned material [4].

The crystals are quite stable in the X-ray beam. They diffract to at least 3.0 Å and native data have been recorded to this resolution. A search for heavy-atom derivatives is in progress. We are working to introduce cysteine residues at several positions in the molecule by site-directed mutagenesis, to allow us to obtain mercury derivatives. The absence of sulfate ions in our crystallization buffer may well prove to be advantageous in the formation of heavy atom derivatives.

Subsequent to our cloning and crystallisation of the BL9 protein a publication appeared presenting cloning of the S5, L6, L9 (i.e. BL9) and L18 from *Bst* and crystallisation of S5 [21]. One amino acid difference was detected between our *bl9* clone and those described in [21]; an extra Lys at the C-terminus of the protein was found.

Table I

Properties of the BL9 ribosomal protein crystals

Space group:	$P2_12_12$
Lattice constants:	a = 135.3 Å
	b = 37.5 Å
	c = 48.9 Å
Volume of the unit cell:	V = 248 100 Å^3
Volume per unit protein mass:	V_M = 3.65 Å^3/Da
Limiting resolution:	3 Å

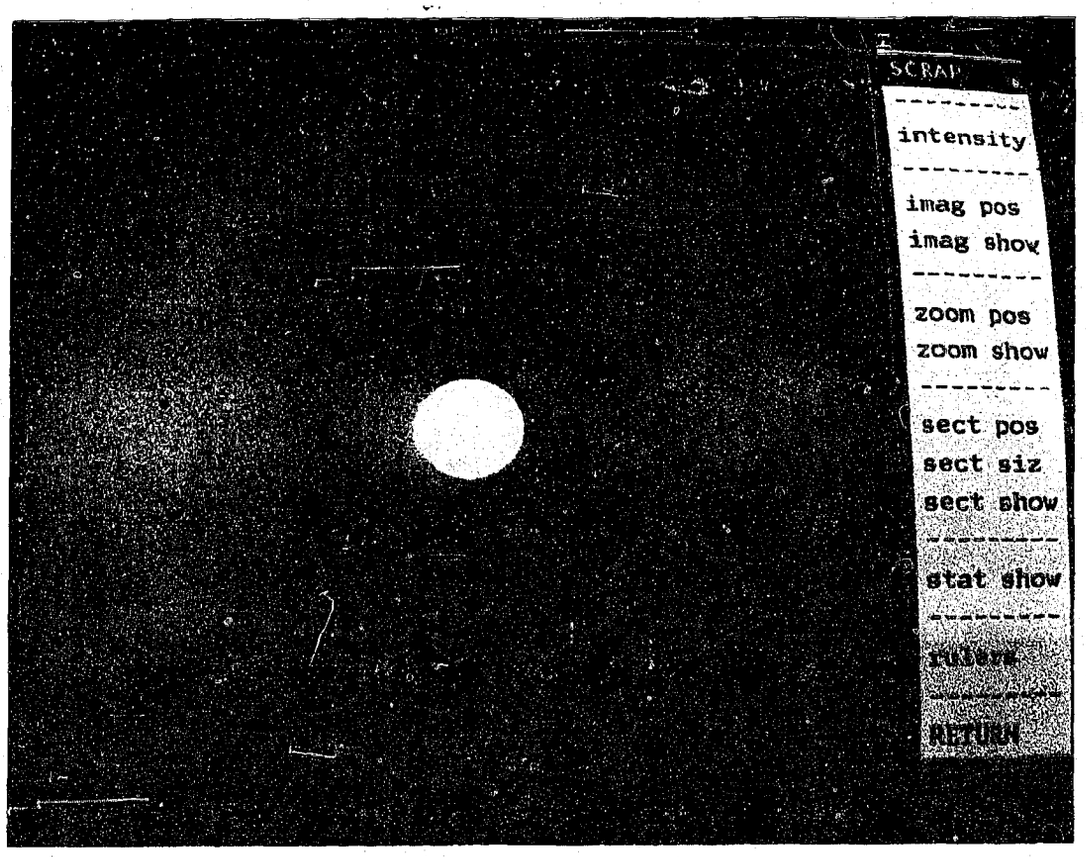

Fig. 5. A 1° rotation image from a BL9 crystal. The data extend to about 3.0 Å.

Acknowledgements: We thank Miroslawa Dauter for excellent technical assistance, Susie Weston for synthesizing the oligonucleotides, Panos Padas for performing the sequencing and Dr F.W. Studier for providing the complete set of vectors and host cells of the T7 RNA polymerase expression system.

REFERENCES

[1] Holck, A., Lossius, I., Aasland, R. and Kleppe, K. (1987) Biochim. Biophys. Acta 914, 49–54.
[2] Holck, A. and Kleppe, K. (1988) Gene 67, 117–124.
[3] Kimura, M., Dijk, J. and Heiland, I. (1980) FEBS Lett. 21, 323–326.
[4] Appelt, K., Dijk, J. and Epp, O. (1979) FEBS Lett. 103, 66–70.
[5] Brimacombe, R., Stöffler, G. and Witmann, H.G. (1978) Annu. Rev. Biochem. 47, 217–249.
[6] Morrison, C.A., Bradbury, E.M., Littlechild, J. and Dijk, J. (1977) FEBS Lett. 83, 348–352.
[7] Privalov, P.L. and Khechinashvili, N.N. (1974) J. Mol. Biol. 86, 665–684.
[8] Khechinashvili, N.N., Kotcliansky, V.E., Gogia, Z.V., Littlechild, J. and Dijk, J. (1978) FEBS Lett. 95, 270–272.
[9] Moore, P.B. (1979) in: Ribosomes: Structure, Function and Genetics (G. Chambliss et al., eds.), pp. 111–134, Univ. Park Press, Baltimore, MD.
[10] Traut, R.R., Lambert, J.M., Boileau, G. and Kenny, J.W. (1979) in: Ribosomes: Structure, Function and Genetics (G. Chambliss et al., eds.), pp. 89–110, Univ. Park Press, Baltimore, MD.
[11] Cantor, C.R. (1974) in: Ribosomes (M. Nomura et al., eds.), pp. 587–599, Cold Spring Harbor Lab., Long Island, NY.
[12] Nomura, M. and Held, W.A. (1974) in: Ribosomes (M. Nomura et al., eds.), pp. 193–223, Cold Spring Harbor Lab., Long Island, NY.
[13] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning. A Laboratory Manual (2nd ed.), Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
[14] Studier, F.W. and Moffatt, B.A. (1986) J. Mol. Biol. 189, 113–130.
[15] Studier, F.W., Rosenberg, A.H., Dunn, J.J. and Dubendorff, J.W. (1990) Methods Enzymol. 185, 60–89.
[16] Sanger, F., Nicklen, S. and Coulson, A.R. (1979) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[17] Mandel, M. and Higa, A. (1970) J. Mol. Biol. 53, 159–162.
[18] Laemmli, U.K. and Favre, M. (1973) J. Mol. Biol. 80, 575–599.
[19] Bradford, M. (1976) Anal. Biochem. 72, 248–254.
[20] Davies, D.R. and Segal, D.M. (1971) Methods Enzymol. 22, 266–269.
[21] Ramakrishnan, V. and Gerchman, S.E. (1991) J. Biol. Chem. 266, 880–885.