FEBS Letters 525 (2002) 88–92 FEBS 26350

Ribosomal protein S1 from *Thermus thermophilus*: its detection, identification and overproduction[1]

Vyacheslav M. Shiryaev[a], Olga M. Selivanova[a], Thomas Hartsch[b], Igor V. Nazimov[c], Alexander S. Spirin[a,*]

[a] *Institute of Protein Research, Russian Academy of Sciences, 142290 Pushchino, Moscow Region, Russia*
[b] *Göttingen Genomics Laboratory, Institut für Mikrobiologie und Genetik, Georg-August-Universität Göttingen, Grisebachstr. 8, D-37077 Göttingen, Germany*
[c] *M.M. Shemyakin and Yu.A. Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, 16/10 Miklukho-Maklay Str., 17871 Moscow, Russia*

Received 20 June 2002; accepted 7 July 2002

First published online 18 July 2002

Edited by Horst Feldmann

Abstract Ribosomal protein S1 has been identified in *Thermus thermophilus* ribosomes. The gene of ribosomal protein S1 from *Thermus thermophilus* has been cloned and overexpressed in *Escherichia coli*. A procedure for purification of the protein has been developed.

Key words: Ribosome; Ribosomal protein S1; Gene overexpression; *Thermus thermophilus*

1. Introduction

Until recently direct data on the presence of ribosomal protein S1 in *Thermus thermophilus* ribosomes have been absent. Moreover, its existence was repeatedly called in question [1,2]. By the end of 2001, the protein was claimed to be present in *T. thermophilus* ribosomes, but the results presented did not allow identifying it among proteins of the ribosomal 30S subunit [3].

At the same time, there are a large number of reports on the ribosomal protein S1 from *Escherichia coli*. Protein S1 from *E. coli* is the largest ribosomal protein with a molecular mass of 61 kDa and a length of 557 amino acid residues. The protein is weakly retained within the 30S ribosomal subunit. Using immunoelectron microscopy, cryoelectron microscopy and small-angle X-ray scattering analysis, protein S1 was located between the head of the 30S ribosomal subunit and its side bulge [4–6].

Functionally, protein S1 is subdivided into two domains. The N-terminal domain (amino acid residues 1–195) seems to be responsible mainly for protein–protein interactions and binds to the ribosomal subunit. The C-terminal domain (amino acid residues 196–557) is responsible for RNA–protein interactions and capable of binding to mRNA [7]. Protein S1 plays an important role in initiation of translation and is required for the efficient in vivo translation of most natural mRNAs in *E. coli* [7,8]. Like some other ribosomal proteins, protein S1 regulates its own synthesis as an autogenous repressor [7–11]. Protein S1 binds to tmRNA (an RNA molecule that combines the properties of a transfer and a messenger RNA) and, promoting its interaction with the 70S ribosome, participates in the *trans*-translation [2,12]. Some extraribosomal functions of the protein are also known. It is included in the Qβ phage replicase as one of its four subunits and is necessary for the replication of the plus strand of Qβ RNA [13,14]. Protein S1 is also required for the function of the RNA-specific phage T4 endonuclease [15].

The most specific feature of the primary structure of protein S1 is the presence of six homologous amino acid repeats of about 70 amino acids long [7,16]. This repeat was first identified in the *E. coli* protein S1 and named S1 RNA-binding motif or S1 domain [7,17]. The S1 motif has been found in some other RNA-binding proteins. One copy of the S1 motif is present in translation initiation factor IF1 of bacteria and chloroplasts, as well as in bacterial polynucleotide phosphorylase. Yeast protein S57596, whose function is not yet known, contains 12 copies of the S1 domain. Proteins containing nine to 12 copies of this domain have been found in *Caenorhabditis elegans* and humans [17].

There is uncertainty as to the conformational properties of protein S1. When isolated from the ribosomes of *E. coli*, the protein seems to have no compact conformation. The results of small-angle X-ray scattering analysis have demonstrated that *E. coli* protein S1 in solution has a highly asymmetric shape, its length (23 nm) being comparable with the dimension of the ribosome [7]. At the same time, the maximal dimension of the full Qβ replicase, which contains protein S1 as a subunit, is only 10 nm [18]. This indicates that protein S1 acquires a compact conformation within the Qβ replicase. It can be assumed that being within the ribosomal subunit protein S1 is also more compact than in solution [18].

This important ribosomal protein has been found neither in *T. thermophilus*, nor in some other microorganisms. It was proposed that other proteins could play the role of protein S1 in such microorganisms [1,2]. In this paper direct evidence for the presence of protein S1 in *T. thermophilus* ribosomes is given.

*Correspondence author. Fax: (7)-095-924 04 93.
E-mail address: spirin@vega.protres.ru (A.S. Spirin).

[1] The nucleotide sequence for *Thermus thermophilus* ribosomal protein S1 has been deposited to EMBL data bank under accession number AJ458461.

0014-5793/02/$22.00
PII: S0014-5793(02)03092-2

2. Materials and methods

2.1. Materials, database and programs

All chemicals were from Sigma (USA) and Serva (Germany). Restriction enzymes *Eco*RI and *Nco*I, and T4 DNA ligase were from Promega (USA), Pfu DNA polymerase was from Stratagene (USA). DEAE-Sepharose was from Pharmacia Biotech (Sweden).

In this work the pET21d vector and *E. coli* strains XL1-Blue and BL21(DE3) were used.

The amino acid sequence of the *E. coli* protein S1 was taken from the SwissProt Data Bank (http://www.expasy.ch/sprot/, accession number P02349). The 'Gene Runner' program (Hastings Software) was used to analyze the nucleotide sequences. Alignment of the protein sequences was done with the 'Clustal W' program [19].

2.2. Preparation of ribosomes, ribosomal subunits and total ribosomal protein

The growth of *T. thermophilus* HB8 cells, purification of ribosomes and their separation into subunits were done as described earlier [20]. *E. coli* MRE-600 ribosomes were prepared according to [21] and separated into subunits as described in [22]. The total preparation of ribosomal proteins of the 30S ribosomal subunit was obtained using the standard procedure of acetic acid extraction of proteins [23] in the presence of 100 mM $MgCl_2$ at 4°C overnight.

2.3. Two-dimensional gel electrophoresis of the 30S ribosomal proteins

To separate individual ribosomal proteins, the two-dimensional electrophoresis method according to Agafonov et al. [24] with minor modifications was used. The protein samples were not treated with iodoacetamide prior to the electrophoretic separation. In the second dimension, a non-gradient separating gel was used (16.5% T, 6% C).

2.4. Determination of the N-terminal amino acid sequence

The N-terminal amino acid sequence analysis of the protein was performed in an automatic gas phase sequencer (Applied BioSystems model 447) equipped with HPLC system 120A from the same manufacturer. Prior to this, the protein was transferred onto Immobilon-P^{sq} transfer membrane according to Matsudaira [25] in transfer buffer containing 10 mM Na_3BO_3 (pH 11) and 0.02% (v/v) β-mercaptoethanol.

2.5. Cloning and overexpression

The Studier system [26] including the host strain *E. coli* BL21(DE3) and vector pET21d was used. Chromosomal DNA from *T. thermophilus* was isolated by a modified Marmur procedure [27] and used as a template for PCR amplification of the protein S1 gene. To perform cloning, sites of *Nco*I and *Eco*RI (underlined) were introduced into the oligonucleotide primers: N-terminus: 5′-GGCGCCATGGAAGACAAGGCGACCC-3′; C-terminus: 5′-GGAAGAATTCTTAGCCCCGGCTCTTCTCCTC-3′.

After purification of the PCR product in a melting agarose gel and cutting it with *Nco*I and *Eco*RI it was cloned into vector pET21d in these sites. The DNA fragments were ligated by the T4 DNA ligase, and the recombinant plasmid called pET21d-tthS1, containing the protein S1 gene under the T7 promoter, was used for transformation of *E. coli* BL21(DE3) cells.

For expression of the protein gene, the BL21(DE3)/pET21d-tthS1 cells were grown in the LB medium containing 100 μg/ml ampicillin at 37°C to an $A_{600} = 0.8–1.0$. The synthesis of *T. thermophilus* ribosomal protein S1 was induced by the addition of IPTG to a concentration of 1 mM, and cell growth was continued for 3 h under the same conditions. The level of protein S1 synthesis in the obtained clones was analyzed using SDS–PAGE according to Laemmli [28]. The protein content in the inclusion bodies was determined as described [29].

2.6. Isolation and purification of the recombinant protein S1

Cells obtained from 1 l culture (3 g) were suspended in 40 ml of 100 mM Tris–HCl buffer pH 8.0, containing 100 mM $MgCl_2$, 800 mM NaCl, 1 mM EDTA and 0.3 mM phenylmethylsulfonyl fluoride, and disrupted by sonication. After removal of cell debris the supernatant was heated at 65°C for 20 min, and the denatured *E. coli* proteins were removed by centrifugation (12 000×*g*, 30 min) [30]. The supernatant was concentrated, dialyzed against 50 mM Tris–HCl buffer pH 8.0, containing 1 mM EDTA, and loaded on 40 ml DEAE-Sepharose column equilibrated against the same buffer. After washing, the recombinant protein was eluted with a linear 2×200 ml gradient of 0–0.2 M NaCl in the same buffer. The elution rate was 50 ml/h with a fraction volume of 10 ml. The gradient fractions were analyzed by SDS–PAGE and those containing the pure protein were pooled. The protein was precipitated with 2 M ammonium sulfate. The protein concentration was determined by the standard Bradford method [31].

3. Results and discussion

The analysis of the protein composition of ribosomes and their subunits is usually done with two-dimensional electrophoresis according to Madjar et al. [32]. This method of electrophoresis, however, is not quite suitable for separation of relatively high molecular mass proteins, such as protein S1, in the second dimension. In order to analyze the protein composition of the 30S ribosomal subunit from *T. thermophilus*, we used a modification of the two-dimensional gel electrophoresis recently developed by Agafonov et al. [24]. The main difference from the standard Madjar two-dimensional electrophoresis of proteins is that the modified system of SDS–PAGE according to Schägger and von Jagow [33] is used in the second dimension, instead of electrophoresis in 6 M urea and acid pH. The system allows separating the ribosomal

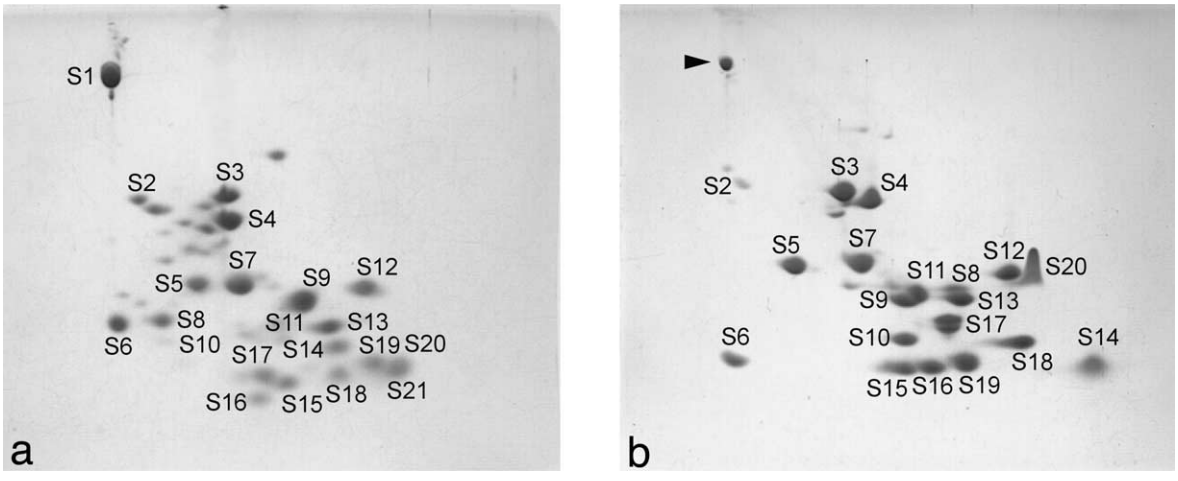

Fig. 1. Electrophoretic separation of proteins of 30S ribosomal subunits from *E. coli* (a) and *T. thermophilus* (b). Two-dimensional gel electrophoresis was performed as described in [24]. The arrow indicates the protein found as a candidate for S1.

proteins in the second dimension in a wide range of molecular masses and correlating the results with those of SDS–PAGE. While performing electrophoresis of the proteins of *T. thermophilus* 30S ribosomal subunits in the new system, an additional spot of a relatively high molecular mass protein was detected. The comparison of the results of the electrophoretic separation of 30S subunit proteins from *E. coli* and *T. thermophilus* in this system has shown that the position of this spot corresponds to that of protein S1 from *E. coli* (Fig. 1). SDS–PAGE of proteins of ribosomes and ribosomal subunits from *E. coli* and *T. thermophilus* has demonstrated that this high molecular mass protein, similar to protein S1 from *E. coli*, is present in the 70S ribosome and the 30S subunit of *T. thermophilus* but is absent in the 50S subunit (Fig. 2). It should be noted that, according to SDS–PAGE and two-dimensional electrophoresis maps, the content of the S1-like protein in the samples of *T. thermophilus* ribosomes is typically lower than that of protein S1 in the standard preparations of *E. coli* ribosomes. This can be explained by the more drastic procedure used for the isolation of *T. thermophilus* ribosomes, as compared with that in the case of *E. coli* ribosomes.

In order to identify the S1-like protein of *T. thermophilus* more accurately, we performed N-terminal amino acid analysis and determined 13 N-terminal amino acids. It turned out that the N-terminal sequence of this protein is not homologous to that of the *E. coli* protein S1. At the same time this amino acid sequence allowed us to identify the corresponding DNA sequence in the *T. thermophilus* genome that is under investigation in the Georg-August-Universität Göttingen. The analysis of the sequence has led to the detection of the open reading frame, i.e. the gene, encoding the polypeptide of 536 amino acid residues in length that contains six homologous amino acid repeats corresponding to the S1 motifs. In addition, a high homology of the deduced amino acid sequence of

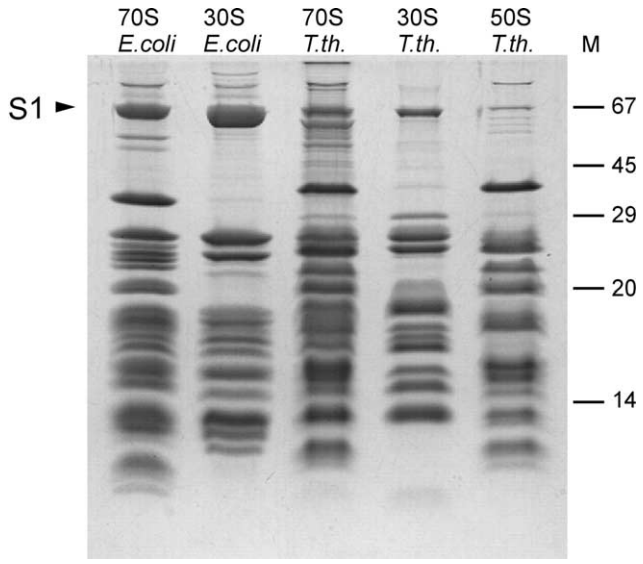

Fig. 2. SDS–PAGE analysis of proteins of ribosomes and ribosomal subunits from *E. coli* and *T. thermophilus* according to Laemmli [28]. M, molecular weight markers, kDa.

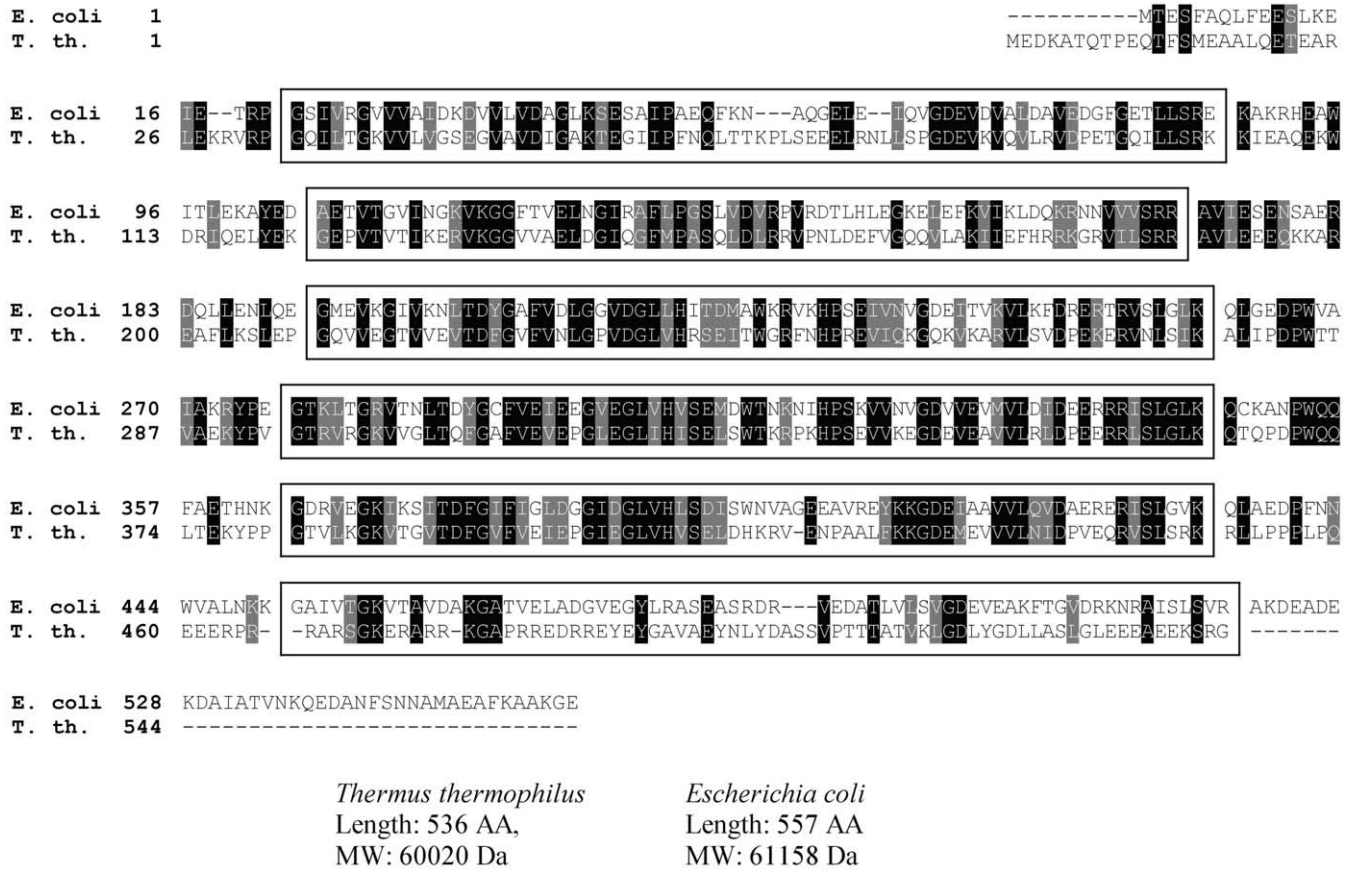

Fig. 3. Alignment of the amino acid sequences of ribosomal proteins S1 from *E. coli* and *T. thermophilus*. Six homologous S1 RNA-binding motifs are given in frames.

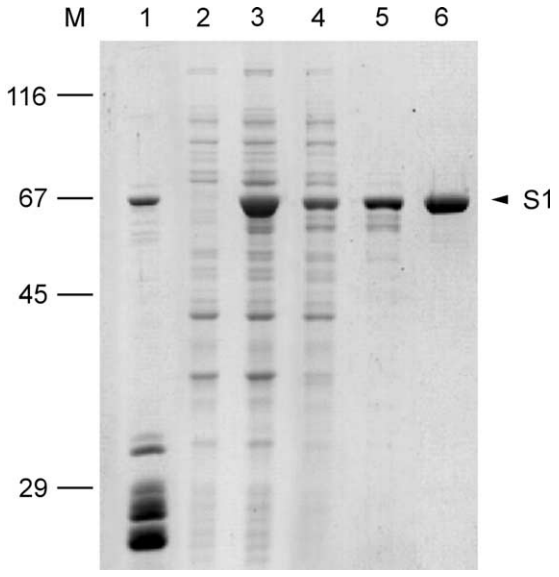

Fig. 4. Expression and purification of *T. thermophilus* ribosomal protein S1 overproduced in *E. coli*. SDS–PAGE analysis was performed as described in [28]. Lane 1, proteins of *T. thermophilus* 30S ribosomal subunit; lanes 2 and 3, cell extract of *E. coli* before and after induction of the synthesis of the recombinant protein S1; lanes 4 and 5, cell lysate of *E. coli* before and after heating at 65°C; lane 6, recombinant protein S1 after DEAE-Sepharose chromatography. M, molecular weight markers, kDa.

this polypeptide with the amino acid sequence of the protein S1 from *E. coli* is observed (Fig. 3). Therefore, the protein found in the 30S ribosomal subunit of *T. thermophilus* can be definitely identified as the ribosomal protein S1.

For the purpose of preparative production of *T. thermophilus* protein S1, the gene of the thermophilic protein was amplified, cloned in expression vector pET21d under the promoter of T7 RNA polymerase and overexpressed in *E. coli* BL21(DE3) cells. The expression was performed as described in Section 2. According to the data from SDS–PAGE, the production of the recombinant protein in *E. coli* cells was about 20% of the total cellular protein (Fig. 4). The overproduced protein was in a soluble form. Heating of the cell extract at 65°C for 20 min made it possible to remove the greater part of *E. coli* proteins as precipitates at this initial stage of isolation of the recombinant protein (Fig. 4). The high NaCl concentration at this stage of purification made it possible to avoid co-precipitation of the *T. thermophilus* protein S1 with *E. coli* proteins.

The final purification of the protein was done with ion exchange DEAE-Sepharose chromatography. The protein was eluted from the column at 0.13 M NaCl. According to SDS–PAGE analysis, the protein purity was higher than 95% (Fig. 4). The final yield of the recombinant protein was more than 3.5 mg from 1 g of the *E. coli* cell biomass.

Recently the three-dimensional structure of bacterial ribosomes and their 30S and 50S ribosomal subunits has been determined with the use of X-ray crystallography [34–36]. A high resolution (up to 2.4–3.0 Å) has been achieved for such complex ribonucleoprotein complexes as the ribosome and its subunits. However, the largest and functionally important ribosomal protein S1 was absent in the atomic structure of the *T. thermophilus* 30S ribosomal subunit [35], probably due to its loss during ribosome and subunit purification. At the same time, the interaction of the protein S1 with the 30S subunit may contribute to its overall structure. Due to the interdomain localization of protein S1 on the 30S subunit [4–6] it can affect the domain disposition in the functioning ribosome. Thus, the solution of the atomic structure of the ribosome seems to be incomplete without protein S1.

Besides, the identification of *T. thermophilus* ribosomal protein S1 and its preparative production makes it possible to study the conformation of this protein in the isolated state. It is known that proteins of thermophilic organisms are more stable in solution than their homologs from mesophilic ones, this permitting the preservation of their native structure after isolation, their easier crystallization and the use of other physical approaches to their conformational analyses. There is high hope that the uncertainty about the folding state of ribosomal protein S1, both within the ribosome and in the isolated form, will be solved with the use of the *T. thermophilus* protein as an object of physical studies.

Acknowledgements: The authors are grateful to Dr. A.B. Chetverin for discussion and valuable remarks, Prof. N.I. Matvienko for the gift of the pET21d plasmid, *E. coli* XL1-Blue and BL21(DE3) strains and discussion on the genetic engineering part of the work, Dr. V.I. Ugarov for helpful recommendations at the initial stage of our study and Dr. D.E. Agafonov for consultations during carrying out the electrophoresis experiments.

References

[1] Tsiboli, P., Herfurth, E. and Choli, T. (1994) Eur. J. Biochem. 226, 169–177.
[2] Wower, I.K., Zwieb, C.W., Guven, S.A. and Wower, J. (2000) EMBO J. 19, 6612–6621.
[3] Clemons Jr., W.M., Brodersen, D.E., McCutcheon, J.P., May, J.L.C., Carter, A.P., Morgan-Warren, R.J., Wimberly, B.T. and Ramakrishnan, V. (2001) J. Mol. Biol. 310, 827–843.
[4] Walleczek, J., Albrecht-Ehrlich, R., Stöffler, G. and Stöffler-Meilicke, M. (1990) J. Biol. Chem. 265, 11338–11344.
[5] Sengupta, J., Agrawal, R.K. and Frank, J. (2001) Proc. Natl. Acad. Sci. USA 98, 11991–11996.
[6] Sillers, I.-Yu. and Moore, P.B. (1981) J. Mol. Biol. 153, 761–780.
[7] Subramanian, A.R. (1983) Prog. Nucleic Acid Res. 28, 101–142.
[8] Sorensen, M.A., Fricke, J. and Pedersen, S. (1998) J. Mol. Biol. 280, 561–569.
[9] Skouv, J., Schnier, J., Rasmussen, M.D., Subramanian, A.R. and Pedersen, S. (1990) J. Biol. Chem. 265, 17044–17049.
[10] Boni, I.V., Issaeva, D.M., Musychenko, M.L. and Tzareva, N.V. (1991) Nucleic Acids Res. 19, 155–162.
[11] Boni, I.V., Artamonova, V.S. and Dreyfus, M. (2000) J. Bacteriol. 182, 5872–5879.
[12] Hanawa-Suetsugu, K., Bordeau, V., Himeno, H., Muto, A. and Felden, B. (2001) Nucleic Acids Res. 29, 4463–4673.
[13] Wahba, A.J., Miller, M.J., Niveleau, A., Landers, T.A., Carmichael, G., Weber, K., Hawley, D.A. and Slobin, L.I. (1974) J. Biol. Chem. 249, 3314–3316.
[14] Kamen, R., Kondo, M., Römmer, W. and Weissmann, C. (1972) Eur. J. Biochem. 31, 44–51.
[15] Ruckman, J., Ringquist, S., Brody, E. and Gold, L. (1994) J. Biol. Chem. 269, 26655–26662.
[16] Gribskov, M. (1992) Gene 119, 107–111.
[17] Bycroft, M., Hubbard, T.J.P., Proctor, M., Freund, S.M.V. and Murzin, A.G. (1997) Cell 88, 235–242.
[18] Berestowskaya, N.H., Vasiliev, V.D., Volkov, A.A. and Chetverin, A.B. (1988) FEBS Lett. 228, 263–267.
[19] Tompson, J.D., Higgins, D.G. and Gibson, T.G. (1994) Nucleic Acids Res. 22, 4673–4680.
[20] Gogia, Z.V., Yusupov, M.M. and Spirina, T.N. (1986) Mol. Biol. (Moscow) 20, 415–421.
[21] Agafonov, D.E., Kolb, V.A. and Spirin, A.S. (1997) Proc. Natl. Acad. Sci. USA 94, 12892–12897.

[22] Gavrilova, L.P., Kostiashkina, O.E., Koteliansky, V.E., Rutkevitch, N.M. and Spirin, A.S. (1976) J. Mol. Biol. 101, 537–552.
[23] Hardy, S.J.S., Kurland, C.G., Voynow, P. and Mora, G. (1969) Biochemistry 8, 2897–2905.
[24] Agafonov, D.E., Kolb, V.A., Nazimov, I.V. and Spirin, A.S. (1999) Proc. Natl. Acad. Sci. USA 96, 12345–12349.
[25] Matsudaira, P. (1987) J. Biol. Chem. 262, 10035–10038.
[26] Studier, F.W. (1990) Methods Enzymol. 185, 60–89.
[27] Marmur, J. (1961) J. Mol. Biol. 3, 208–218.
[28] Laemmli, U.K. (1970) Nature 227, 680–685.
[29] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd edn., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
[30] Khairullina, A.R., Scherbakov, D.V., Tischenko, S.V., Nikonov, S.V. and Garber, M.B. (1997) Biochemistry (Moscow) 62, 221–224.
[31] Bradford, M.M. (1976) Anal. Biochem. 72, 248–254.
[32] Madjar, J.-J., Michel, S., Cozzone, A.J. and Rebiund, J.-P. (1979) Anal. Biochem. 92, 174–182.
[33] Schägger, H. and von Jagow, G. (1987) Anal. Biochem. 166, 368–379.
[34] Cate, J.H., Yusupov, M.M., Yusupova, G.Zh., Earnest, T.N. and Noller, H.F. (1999) Science 285, 2095–2104.
[35] Wimberly, B.T., Brodersen, D.E., Clemons, W.M., Morgan-Warren, R.J., Carter, A.P., Vornhein, C., Hartsch, T. and Ramakrishnan, V. (2000) Nature 407, 327–339.
[36] Ban, N., Nissen, P., Moore, P.B. and Steitz, T.A. (2000) Science 289, 905–920.