Over-expression of *Escherichia coli* F_1F_o–ATPase subunit a is inhibited by instability of the *uncB* gene transcript

Ignacio Arechaga[1], Bruno Miroux[1,2], Mike J. Runswick, John E. Walker*

The Medical Research Council Dunn Human Nutrition Unit, Wellcome Trust/MRC Building, Hills Road, Cambridge CB2 2XY, UK

Received 21 May 2003; revised 5 June 2003; accepted 5 June 2003

First published online 18 June 2003

Edited by Stuart Ferguson

Abstract **Little is known about the stability of transcripts encoding membrane proteins in strong expression systems and its effect on membrane protein over-production. We have expressed all the genes encoding subunits of the membrane domain F_o of the ATP synthase in a T7 RNA polymerase-based system. All of them but *uncB* (subunit a) were expressed separately at very high levels in the bacterial hosts *Escherichia coli* C41(DE3) and C43(DE3). However, expression of *uncB* was extremely toxic to the bacteria. Northern blot analysis showed that the level of accumulation of the mRNA from *uncB* was very low. Deletion of *uncB* in combination with gene fusion experiments demonstrated that the middle region of the gene, encoding amino acids 92–171, exhibited a dominant toxic phenotype associated with a very poor level of expression. Green fluorescent protein fusions with N- and C-ends of *uncB* helped to stabilize the mRNA and to obtain high yields of protein.**

Key words: Membrane protein; Over-expression; ATP synthase; mRNA degradation

1. Introduction

F_1F_o–ATP synthase is a membrane-bound complex that couples proton translocation across the membrane to ATP synthesis. It consists of a membrane domain, F_o, where proton translocation occurs, and a soluble, catalytic F_1 domain linked together by central and peripheral stalks. The central stalk, together with a ring of c subunits in F_o, forms an ensemble, which rotates. The rotation is driven by proton translocation across the membrane domain and it induces conformational changes in F_1 that result in ATP synthesis [1,2].

In *Escherichia coli* all eight subunits are encoded by the polycistronic 7.3-kb *unc* operon. The order of the genes is *uncIBEFHAGCD* [3]. The promoter proximal gene, *uncI*, encodes a 14-kDa, membrane-embedded polypeptide of unknown function that may play a role in the assembly of the complex [4,5]. The next three genes encode the F_o subunits a, c and b, respectively, and the remaining five genes encode the F_1 subunits δ, α, γ, β and ε, respectively [6]. The *unc* operon is an example of an uncoordinate operon [7] where the rates of synthesis of the individual subunits are adjusted to accommodate the stoichiometry found in the ATP complex ($\alpha_3\beta_3\gamma\delta\varepsilon ab c_{10-12}$) [7–9]. The subunits are produced from a single polycistronic mRNA bearing all nine cistrons [3,10]. Expression of the individual genes is regulated post-transcriptionally [11] and is achieved by the control of the efficiency of translational initiation [12] and by mRNA degradation [13,14].

The first gene of the *unc* operon to encode a mature ATP synthase subunit is *uncB*. Its product is the polytopic membrane-spanning subunit a [3]. Degradation and instability of the *uncB* mRNA plays an important role in the control of its expression [13,14]. Over-expression of *uncB* causes growth inhibition of *E. coli* [15–17]. In the current work, we have investigated whether *uncB* also down-regulates the expression of the other F_o subunits, c and b, in the context of the strong T7-based expression system. Previously, we have reported high levels of over-production of c (*uncE*) and b (*uncF*) subunits, expressed separately in the bacterial host *E. coli* C41(DE3) and C43(DE3) [18–20]. However, expression of *uncB* (subunit a) in these host systems was extremely toxic for the cells. Northern blot analysis confirmed the instability of the *uncB* mRNA. We have identified the most unstable region of the *uncB* mRNA, and we have stabilized the other regions by fusions with the green fluorescent protein (GFP).

The results are discussed in the context of the over-production of membrane proteins for structural studies, and the influence that mRNA instability may have in the failure to over-express some membrane proteins at high levels.

2. Materials and methods

2.1. Plasmid construction

All constructs were made by ligating a PCR product into the *Nde*I and *Hin*dIII or *Eco*RI sites of the expression vector pMW172 [21]. The GFP gene was cloned into the *Bam*HI and *Hin*dIII site of pMW172 to allow the production of fusions by inserting the gene, or fragment of gene, of interest into the *Hin*dIII and *Eco*RI sites. When over-expressed, these constructs gave fusion proteins with the GFP (*Aequoria victoria*) at the N-terminus.

2.2. Bacterial growth and RNA extraction

Cells were grown at 37°C in 2× TY medium. After 5–6 h of growth and once an O.D. (optical density) at 600 nm had reached a value 0.5–0.7, IPTG (isopropyl-thiogalactoside) was added to a concentration of 0.7 mM. After 0, 1, 2 and 3 h of induction, samples were removed and cells were lysed by addition of 1/5 volumes of lysis buffer (1.5 M sodium acetate, 12 mM EDTA, 5% sodium dodecyl sulfate (SDS)) followed by boiling for 2 min. RNA was extracted as follows: aqueous

*Corresponding author. Fax: (44)-1223-410506.
E-mail address: walker@mrc-dunn.cam.ac.uk (J.E. Walker).

[1] These authors contributed equally to this article.
[2] Present address: Faculté de Médecine Necker-Enfants-Malades CNRS-UPR 9078, 156 rue de Vaugirard, 75730 Paris, France.

0014-5793/03/$22.00
doi:10.1016/S0014-5793(03)00677-X

phenol (1/1 v/v, saturated with 10 mM Tris–HCl, pH 7.5) was added to the lysed cells and the sample was centrifuged (5 min, 2000×*g*). A mixture phenol:chloroform:water was added to the supernatant (v/v/v) and centrifuged at 2000×*g* for 5 min. This process was repeated three times and to the final supernatant 2 vols. of ethanol (−20°C) were added. After precipitation overnight at −20°C, samples were centrifuged at 10 000×*g* and pellets were washed with 75% ethanol, dried and dissolved in H_2O–DEPC (diethyl pyrocarbonate). The concentration of RNA was estimated by the A_{260}/A_{280} ratio. Samples were checked for correct size and concentration in agarose gels.

2.3. Northern blot analysis

RNA samples (20 μg) were denatured by boiling (3 min). Dye was added and they were loaded on agarose gels (1% w/v dissolved in a buffer containing 20 mM MOPS, pH 6.8, 2 mM EDTA, 3 mM sodium acetate, and 2% formaldehyde) and run in the same MOPS buffer at 80 V for 4–6 h. Samples were transferred onto Hybond membranes (Amersham Biosciences, Uppsala, Sweden) in the presence of 20× SSC buffer (3M sodium chloride, 0.3 M sodium citrate, pH 6.5) for 9 h and cross-linked with UV light (2× 1 min).

Membranes were treated with prehybridation mix [40 mM Tris–HCl, pH 7.5, 0.3 M NaCl, 2 mM sodium pyrophosphate, 1% SDS, 41.6% formamide, 16.8% Denhardt solution (3% polyvinylpyrolidone, 3% bovine serum albumin, 3% Ficoll 400)]. To this mix, 50% dextran and 0.2% salmon sperm DNA (10 mg/ml, pre-boiled) were added. Membranes were incubated at 42°C for 6–12 h.

2.4. Labeling of hybridization probes

Probes (3 μl, 20 ng) for full-length and truncated forms of *uncB* and for *uncF* and *GFP* were made by PCR reaction using the oligonucleotides used in the plasmid construction, and labeled with 5 μl [α-^{32}P]-dCTP (3000 Ci/μmol, 50 mCi), 10 μl reagent mix, 31 μl H_2O, 1 μl Klenow mix (1 unit Klenow fragment, 5 mM Tris–HCl, 5 mM $MgCl_2$, 50 mM NaCl, 50% glycerol) for 30–60 min at 37°C. Labeled probes were purified with QIAquick nucleotide removal kit (Qiagen, Hilden, Germany) and boiled.

2.5. Hybridization

Membranes were treated with hybridization mix (60 mM Tris–HCl, pH 7.5, 2.6 mM sodium pyrophosphate, 60% formamide, 24% Denhardt solution, 12% SDS), 8% salmon sperm DNA (10 mg/ml) and the respective labeled probe. Membranes were incubated for 16 h at 42°C, and then washed with 2× SSC buffer/0.1% SDS for 10–20 min (twice) at 42°C, 0.2× SSC buffer/0.1% SDS for 10–20 min (twice) at 55°C and 0.1× SSC buffer/0.1% SDS for 10 min at 65°C. Membranes were exposed to Fuji film for 2–12 h and scanned in a Molecular Dynamics Phosphorimager (Amersham Bioscience).

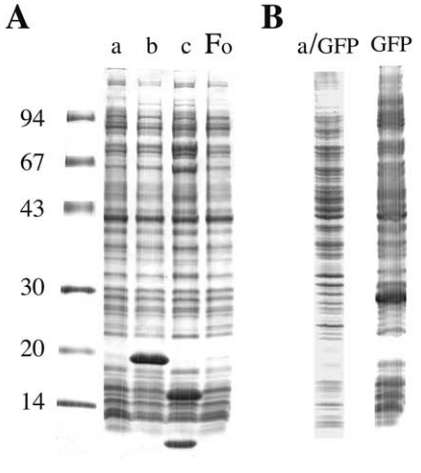

Fig. 1. A: Protein contents of total extracts from *E. coli* C41(DE3) cells over-expressing constructs consisting of *uncB* (subunit a), *uncF* (subunit b), *uncE* (subunit c) and the mini-operon *uncBEF* (F_o). Portions (5 μl, ca. 10^7 cells) were analyzed by SDS–PAGE. B: Protein contents of total cell extracts from *E. coli* C41(DE3) cells over-expressing the fusion of *uncB* with the GFP or the GFP alone.

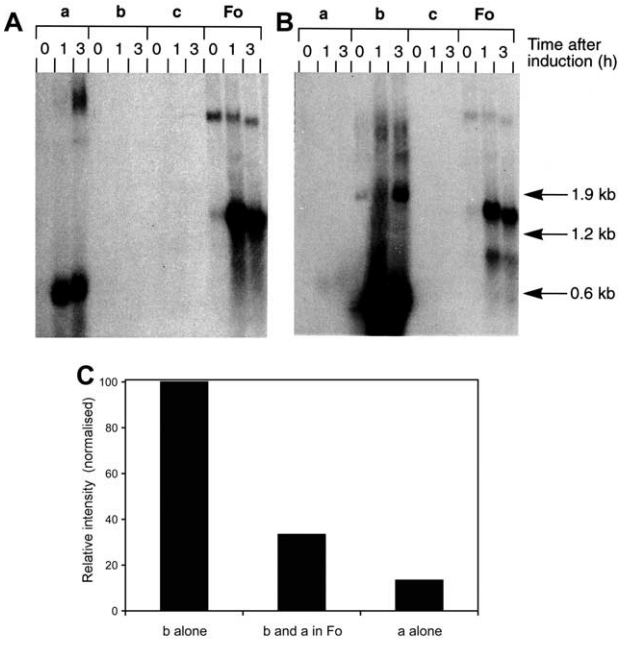

Fig. 2. Northern blot analysis of the expression of *uncB* (subunit a). *UncB* was cloned into a pET vector, separately or in the presence of other subunits of F_o, in the same order as in the *unc* operon. RNA from cells expressing these constructs was extracted after 1, 2 and 3 h of induction. RNA was separated electrophoretically and transferred onto HyBond membranes. A: Membranes hybridized with a radioactive probe that recognizes *uncB* sequence only, exposed to a photographic film and scanned. B: Membranes hybridized with a radioactive probe that recognizes *uncF* sequence only. C: The normalized values for the different constructs after scanning.

3. Results

3.1. Instability of uncB

In contrast to subunits b and c, expression of subunit a was toxic to strains C41(DE3) and C43(DE3), resulting in growth inhibition in 2× TY medium. Subunit a was undetectable by SDS–PAGE analysis of total cell extracts (Fig. 1A). Co-expression of a subunit as a fusion or separately in mini-operons completely abolished the expression of genes that are normally highly expressed in the same expression system, such as the GFP (Fig. 1B), and subunits c and b (as in F_o) (Fig. 1A). In order to understand the source of the toxicity and the absence of the desired proteins, the levels of mRNA of the target genes were analyzed after induction of the expression. Northern blot analysis showed the levels of accumulation of *uncB* mRNA expressed separately or in a mini-operon containing the genes encoding all three membrane subunits of the F_o domain (Fig. 2A). The levels of *uncB* mRNA expressed separately were approximately half the mRNA level from those observed in F_o (after 3 h of induction). In contrast, *uncF* mRNA (subunit b) expressed in the context of F_o was three-fold lower than the level observed when the *uncF* is expressed separately (Fig. 2B). Comparison between the levels of mRNA accumulation of *uncB* and *uncF* expressed separately suggested a source of instability inherent to *uncB* transcript, which prevents high levels of *uncB* mRNA accumulation, despite the strength of the T7 promoter.

3.2. uncB deletions and GFP fusions

In order to identify the unstable regions of *uncB*, we created

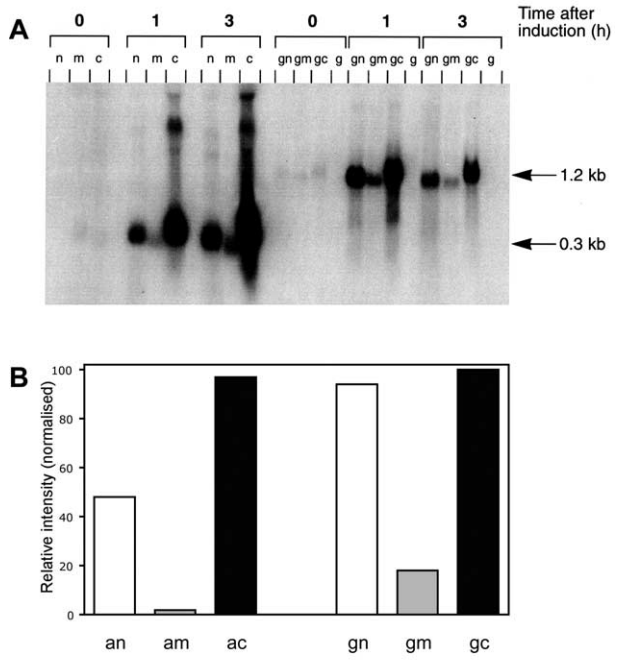

Fig. 3. Northern blot analysis of truncated forms of *uncB*. Constructs consisting of *uncB* regions encoding subunit a domains 1–92 (n), 93–171 (m) and 172–271 (c), separately and with GFP fusions (gn, gm and gc, respectively), were expressed in *E. coli* C41(DE3) cells. RNA was extracted after 1, 2 and 3 h of induction and analyzed by Northern blotting. Panel B shows normalized values (to the highest intensity, in this case gc fusion) after scanning.

constructs corresponding to truncated forms of subunit a consisting of regions 1–92, 93–171 and 172–271 (numbers refer to the amino acid sequence) that we named n, m and c (N-terminal, middle and C-terminal), respectively. Expression of these truncated forms separately showed that the middle region of the transcript was the most unstable (Fig. 3). In no case was protein production detectable by SDS–PAGE (Fig. 4A), and the toxicity associated with the expression of these constructs was very high, as they caused total growth inhibition after induction. In contrast, the toxicity associated with the expression of the C-terminal region was less severe. These results suggest that, in addition to mRNA instability, truncated forms of subunit a were also subjected to proteolytic degradation.

We investigated the possibility of stabilizing these truncated constructs by creating fusions with the GFP. Gene fusion with GFP failed to stabilize the middle part of *uncB* mRNA (Fig. 3) and toxicity was still associated with it. However, fusion of GFP with the C-terminal region of the *uncB* gene, construct GFP–a (residues 172–271 of subunit a), resulted in a high level of protein production (Fig. 4B). The over-produced protein was accumulated as inclusion bodies. However, a similar result was not obtained with the fusion GFP–a(1–92), and despite its being a high level of mRNA, very low levels of protein were detected in SDS–PAGE (Fig. 4B), which suggested that the protein was being removed by proteolytic degradation.

4. Discussion

One of the most fascinating properties of the expression of the genes in the *unc* operon is that it is uncoordinate [7]. Since

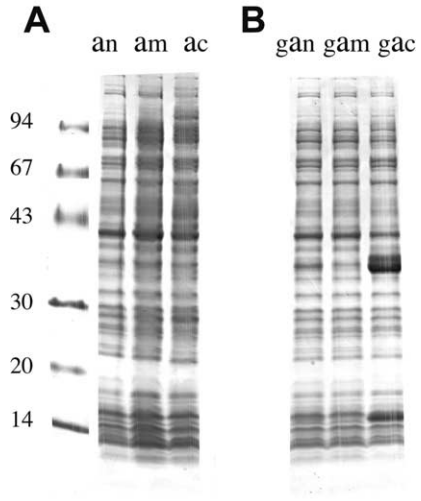

Fig. 4. Protein contents of total extracts from *E. coli* C41(DE3) cells over-expressing constructs consisting of *uncB* regions encoding subunit a domains 1–92 (an), 93–171 (am) and 172–271 (ac), separately and (B) with GFP fusions (gan, gam and gac, respectively). Cells were harvested 3 h after induction with IPTG, centrifuged, resuspended in water and concentrated 10 times. Portions (5 μl, ca. 10^7 cells) were analyzed by SDS–PAGE.

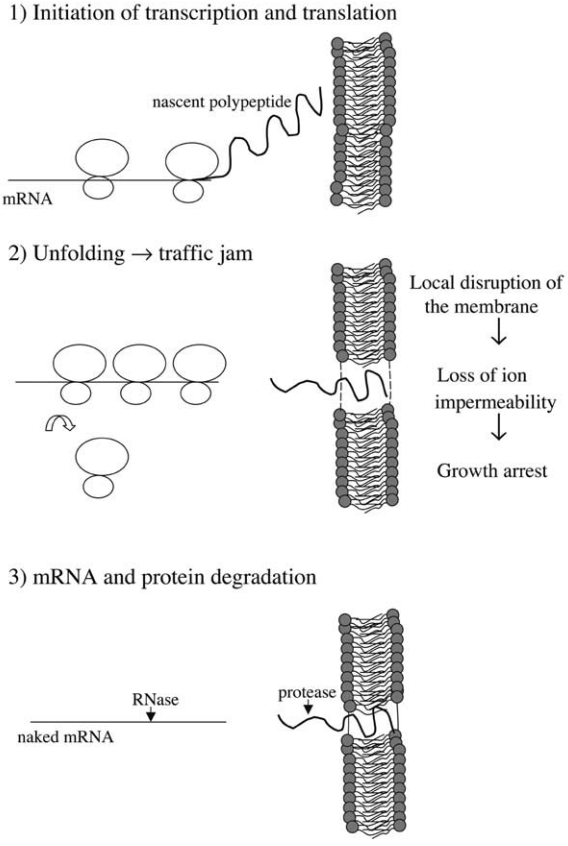

Fig. 5. Membrane protein and mRNA degradation in bacterial protein over-expression systems. 1: Recombinant transcript competes with housekeeping and other host transcripts for ribosome binding sites. 2: Recombinant nascent protein competes for folding and translocation machinery of host cells. Recombinant protein may mis-fold into inclusion bodies or alternatively mis-insert into the membrane. Ribosomes get dissociated from mRNA. 3: Unprocessed mRNA and mis-inserted proteins are pruned to degradation.

the sequence of the entire *unc* operon was published [3], a number of studies described the post-transcriptional control of the expression of the *unc* genes. This post-translational control is exercised by the efficiency of translational initiation [11,12] and by mRNA degradation and instability [13,14]. In other *E. coli* operons, such as *lac* [22], *pap* [23], *mal* [24] and the *puf* operon of *Rhodobacter capsulatus* [25], differential gene expression is controlled by the efficiency of translational initiation and by degradation and segmental instability of the mRNA.

The *uncB* gene encodes the a subunit, a multi-spanning membrane polypeptide which is present in a single copy in the mature ATPase complex. Over-expression of *uncB* has been shown to cause growth inhibition of *E. coli* [15–17]. The inhibition was attributed to various factors, such as increased protonophoric activity of the membranes [16] and unbalanced synthesis of membrane proteins and lipids [17]. None of these studies took into account the effect of mRNA accumulation and stability in the inhibition of growth after over-expression of subunit a. In the present study, we have observed that the growth inhibition associated with the over-expression of subunit a in *E. coli* C41(DE3) cells is related to the stability of the *uncB* gene, and we have found that the a (92–171) domain carries the dominant toxic phenotype associated with poor expression, which is in agreement with previous studies [17].

The *uncB* mRNA is more susceptible to cleavage and degradation than other transcripts from the *unc* operon [13,14], with a number of endonucleolytic cleavage sites found in the 3′ end corresponding to the C-terminal region of *UncB* [14]. However, in the T7-based expression system used in the present study the level of mRNA of the C-terminal region of subunit a (amino acid residues 172–271) is very high compared with the middle region (residues 92–171). In addition, the GFP–a(172–271) fusion protein is expressed at high level. Therefore, the cleavage sites in the transcript encoding the C-terminal region are not relevant and only those cleavage sites found in the middle region may be responsible for mRNA degradation. It has also been shown that the presence of an intragenic ribosome-binding site in the region 246–333 bp (encoding amino acid residues 82–112) may reduce the expression of the *uncB* gene, probably due to a frame-shifting that introduces a translational false start, thus affecting the number of ribosomes that complete translation of *uncB*, which will, in turn, increase the sensitivity to mRNA degradation [26]. In the present work we have shown that the *uncB* region that encodes the subunit a middle domain (residues 93–171) is the most unstable, probably because it is more sensitive to RNA degradation. One of the most important features of the present findings is the possibility of removing the toxicity associated with the expression of some subunit a domains by stabilizing *uncB* mRNA fragments with GFP fusions. These GFP fusions led to high yields of protein production. In contrast, previous attempts to stabilize the *uncB* transcript with *lacZ* fusions [26–28] failed to show any effect on enhancing mRNA accumulation or protein production.

In the present study, we also showed that toxicity associated with the expression of the target gene correlates with the degradation of either its mRNA or the protein. One explanation is that the folding and the insertion into the *E. coli* membrane of the a subunit are the limiting steps for a successful over-expression and assembly of F_o. Mis-insertion of the protein into the membrane might trigger local disruption of the membrane, leading to the loss of impermeability towards ions and growth arrest. As suggested previously [29,30], ribosomes dissociate from mRNA, which then becomes more sensitive to nucleases [11,31,32]. At the membrane site, mis-inserted protein would also be degraded (Fig. 5). In agreement with this hypothesis, GFP fusions with fragments of the a subunit trigger the protein to form inclusion bodies, which preserve the membrane integrity of the cell, allowing a high level of protein production.

References

[1] Walker, J.E. (1998) Angew. Chem. Int. Ed. Engl. 37, 2309–2319.
[2] Stock, D., Gibbons, C., Arechaga, I., Leslie, A.G.W. and Walker, J.E. (2000) Curr. Opin. Struct. Biol. 10, 672–679.
[3] Walker, J.E., Gay, N.J., Saraste, M. and Eberle, A.N. (1984) Biochem. J. 224, 799–815.
[4] Gay, N.J. and Walker, J.E. (1981) Nucleic Acids Res. 9, 3919–3926.
[5] Brusilow, W.A., Porter, A.C.G. and Simoni, R.D. (1983) J. Bacteriol. 155, 1265–1270.
[6] Saraste, M., Gay, N.J., Eberle, A., Runswick, M.J. and Walker, J.E. (1981) Nucleic Acids Res. 9, 5287–5296.
[7] Walker, J.E., Gay, N.J. and Saraste, M. (1984) Biochim. Biophys. Acta 768, 164–200.
[8] Foster, D.L. and Fillingame, R.H. (1982) J. Biol. Chem. 257, 2009–2015.
[9] Jiang, W., Hermolin, J. and Fillingame, R.H. (2001) Proc. Natl. Acad. Sci. USA 98, 4966–4971.
[10] Jones, H.M., Brajkovich, C.M. and Gunsalus, R.P. (1983) J. Bacteriol. 155, 1279–1287.
[11] McCarthy, J.E.G. (1990) Mol. Microbiol. 4, 1233–1240.
[12] McCarthy, J.E.G. and Bokelmann, C. (1988) Mol. Microbiol. 2, 455–465.
[13] McCarthy, J.E.G., Gerstel, B., Surin, B., Wiedemann, U. and Ziemke, P. (1991) Mol. Microbiol. 5, 2447–2458.
[14] Patel, A.M. and Dunn, S.D. (1995) J. Bacteriol. 177, 3917–3922.
[15] Kanazawa, H., Kiyasu, T., Noumi, T. and Futai, M. (1984) J. Bacteriol. 158, 300–306.
[16] von Meyenburg, K., Jørgensen, B.B., Michelsen, O., Sørensen, L. and McCarthy, J.E.G. (1985) EMBO J. 4, 2357–2363.
[17] Eya, S., Maeda, M., Tomochika, K., Kanemasa, Y. and Futai, M. (1989) J. Bacteriol. 171, 6853–6858.
[18] Miroux, B. and Walker, J.E. (1996) J. Mol. Biol. 260, 289–298.
[19] Arechaga, I., Miroux, B., Karrasch, S., Huijbregts, R., de Kruijff, B., Runswick, M.J. and Walker, J.E. (2000) FEBS Lett. 482, 215–219.
[20] Arechaga, I., Butler, P.J.G. and Walker, J.E. (2002) FEBS Lett. 515, 189–193.
[21] Way, M., Pope, B., Hawkins, M. and Weeds, A.G. (1990) EMBO J. 9, 4103–4109.
[22] Canistrato, V.J., Subbarao, M.N. and Kennell, D. (1986) J. Mol. Biol. 192, 257–274.
[23] Baga, M., Giransson, M., Normack, S. and Uhlin, B.E. (1988) Cell 52, 197–206.
[24] Newbury, S.F., Smith, N.H., Robinson, E.C., Hiles, I.D. and Higgins, C.F. (1987) Cell 48, 297–310.
[25] Klug, G., Adams, C.W., Belasco, J., Doerge, B. and Cohen, S.N. (1987) EMBO J. 6, 2515–3520.
[26] Matten, S.R., Schneider, T.D., Ringquist, S. and Brusilow, W.S.A. (1998) J. Bacteriol. 180, 3940–3945.
[27] Angov, E. and Brusilow, W.S.A. (1988) J. Bacteriol. 170, 459–462.
[28] Hsu, D.K. and Brusilow, W.S.A. (1995) FEBS Lett. 371, 127–131.
[29] Dong, H., Nilsson, L. and Kurland, C.G. (1995) J. Bacteriol. 177, 1497–1504.
[30] Kurland, C.G. and Dong, H. (1996) Mol. Microbiol. 21, 1–4.
[31] Kennell, D.E. (1986) in: Maximising Gene Expression (Reznikoff, W. and Gold, L., Eds.), pp. 101–142, Butterworths, Boston.
[32] Belasco, J.G. and Higgins, C.F. (1988) Gene 72, 15–23.