ELSEVIER

FEBS Letters 339 (1994) 195–199

FEBS 13676

FEBS LETTERS

Expression of the human UDP-glucuronosyltransferase UGT1*6 in *Escherichia coli*

Influence of bacterial signal peptides on the production and localization of the recombinant protein

Mohamed Ouzzine[a], Sylvie Fournel-Gigleux[a,*], Thierry Pillot[a], Brian Burchell[b], Gérard Siest[a], Jacques Magdalou[a]

[a]*Centre du Médicament, URA CNRS no. 597, Faculté des Sciences Pharmaceutiques et Biologiques, 30 rue Lionnois, 54000 Nancy, France*
[b]*Department of Biochemical Medicine, Ninewells Hospital and Medical School, University of Dundee, DD1 9SY, Dundee, UK*

Received 6 December 1993; revised version received 3 January 1994

Abstract

The membrane-bound human liver UDP-glucuronosyltransferase UGT1*6 was expressed in *Escherichia coli*. Exchange of the natural signal peptide by the bacterial signal peptides of pelB or OmpT proteins considerably increased the level of expression and, as the natural signal peptide, targeted the protein to the membranes. The extent of maturation of ^{s}pelB-UGT1*6 precursor was about 30%. No processing of ^{s}OmpT-UGT1*6 occurred but the processing rate of this precursor could be significantly increased by mutagenesis of the first two amino acid residues of the mature sequence. These expression vectors allowed us to produce high levels of recombinant mature UGT1*6 required for further structural studies.

Key words: UDP-glucuronosyltransferase; Heterologous expression; Signal peptide; Membrane targeting; Human liver; *Escherichia coli*

1. Introduction

UDP-glucuronosyltransferase (UGT, EC 2.4.1.17), comprises a superfamily of isoenzymes, which detoxify a large variety of endogenous and foreign compounds [1,2]. Cloning and expression of the human liver UGT1*6 cDNA, one of the ten UGT cDNAs cloned to date in man, showed that this isoform presents a restricted specificity towards short and planar phenols and may therefore play a major role in the elimination of reactive molecular species [3–5].

UGTs are integral membrane proteins located in the endoplasmic reticulum (ER) of various cells. Comparison of UGT cDNAs sequences provided evidence for common topogenic elements proposed to mediate translocation and integration of UGT isoenzymes into the ER. Comparison of the N-terminal sequence of purified proteins to primary sequences deduced from cDNA cloning suggested that UGTs are biosynthesized as precursors with an amino-terminal signal peptide sequence. A stop-transfer sequence at the carboxy-terminal part of the protein stops the translocation and consequently integrates the polypeptide in the ER via a transmembrane anchor segment [6,7].

The difficulties encountered in the purification of these phospholipid-dependent enzymes from human tissues have delayed their molecular characterization. However, heterologous expression of UGT cDNAs in eukaryotic cells has allowed functional characterization of the substrate specificity of individual UGTs, but did not provide sufficient amounts of recombinant product for further structural analysis [4,7,8]. Bacterial expression systems have been developed for many proteins to overcome this problem but the possible advantages of this system have not been investigated in the case of UGT yet.

It has been shown that the export of heterologous proteins in *E. coli* may only require a correct protein

* Corresponding author. Fax: (33) 83 32 13 22.

Abbreviations: UGT, UDP-glucuronosyltransferase; PCR, polymerase chain reaction; IPTG, isopropyl-β-D-thiogalactopyranoside; PMSF, phenylmethylsulfonyl fluoride; RBS, ribosome binding site; EDTA, ethylenediaminetetraacetate; ER, endopasmic reticulum.

UGT1*6 [1] (HLUGP1, trivial name for human liver UDPGT phenol 1). The asterisk indicates that this isoform is derived from a primary transcript common to several isoforms encoded by the *UGT1* locus, by alternative splicing.

SSDI 0014-5793(94)00060-9

fusion between a prokaryotic signal peptide and the mature portion of the foreign protein, with the only restriction that the expressed protein itself should be translocated in its natural context [9]. In this report, we developed a bacterial expression system allowing the recovery of large amounts of a recombinant UGT1*6 protein and we examined the influence of the signal peptide on the level of production and on the processing of this protein expressed in *E. coli*. Furthermore, we provide evidence for the importance of the two first amino acid residues at the N-terminus of the mature protein on the rate of processing.

2. Materials and methods

Isopropyl-β-D-thiogalactopyranoside (IPTG) was from Boehringer. Restriction enzymes, T4 DNA ligase were purchased from New England Biolabs and Taq polymerase was from Perkin-Elmer Cetus. Culture media for bacteria were from Difco.

2.1. Bacterial strain and expression vectors

E coli strain BL21 (F^- *ompT* r^-_B m^-_B)(DE3)pLysS was from Novagen. The pET-3d vector was kindly provided by Dr. F. W. Studier (Brookhaven National Laboratory, Brookhaven, USA) [15]. The pET-12a, pET-20b plasmids were from Novagen and the TA cloning kit from Invitrogen.

*2.2. Construction of the vector containing the full-length UGT1*6 cDNA (pUGT1*6)*

Cloning and sequencing of *UGT1*6* cDNA have been reported elsewhere [3]. The cDNA sequence of UGT1*6 was amplified by polymerase chain reaction (PCR) using the synthetic oligomer A, 5′-CGCCATGGCCTGCCTCCTTCGC-3′ as N-terminal primer. This primer was designed to create a *Nco*I site required for subcloning into the pET-3d vector and hybridized to a sequence encoding the first 6 amino acid residues of the precursor protein. The C-terminal primer B, 5′-GCGGATCCTCAATGGGTCTTGGATTT-3′ hybridized to a sequence coding for the 5 last amino acid residues of the protein and contains a *Bam*HI site for subcloning into the vector. Linearized pKCRH2-HLUGP1 vector [3] was used as a template. The amplification was performed as follows: denaturation at 94°C for 1 min, annealing at 55°C for 1 min, gradual temperature increase to 72°C over a 2-min period, and further primer extension at 72°C for 10 min. Amplification was performed for 20 cycles (Perkin-Elmer Cetus). The PCR product was ligated into pCR II vector (TA cloning kit). This recombinant vector was then digested with *Bsu*36I and *Bam*HI and the resulting fragment of 1490bp coding for 497 C-terminal amino acid residues was exchanged with the same fragment of the non-amplified sequence. Finally, the vector was digested by *Nco*I and *Bam*HI and the resulting modified cDNA, encoding UGT1*6 with its natural signal peptide was subcloned into *Nco*I and *Bam*HI sites of pET-3d to generate pUGT1*6.

*2.3 Construction of the vector designed for the expression of UGT1*6 fused to OmpT signal peptide (ptUGT1*6)*

The construction of ptUGT1*6 was performed according to the same strategy than that developed for pUGT1*6 described above, except that the oligomer C, 5′-GGTCGACGGACAAGCTGCTGGTGGTCCCTCAG-3′ was used as N-terminal primer. This primer contains a *Sal*I site required for subcloning into the pET-12a vector and hybridized to a sequence encoding the first 8 amino acid residues of the mature sequence of the protein. For the mutagenesis of the first two amino acid residues (Asp-1, Lys-2) to (Pro-1, Glu-2) of the mature sequence of UGT1*6 in the ^{s}OmpT-UGT1*6 fusion, the following oligonucleotide D, 5′-GGTCGACGCCTGAACTGCTGGTGGTCCCTCAG-3′ was used as N-terminal primer. In both cases the PCR product was digested with *Sal*I and *Bam*HI generating a fragment coding for the mature native or mutated UGT1*6 protein, which was isolated and subcloned into *Sal*I–*Bam*HI sites of pET-12a in frame with the sequence coding for the OmpT signal peptide.

*2.4. Construction of the vector designed for the expression of UGT1*6 fused to pelB signal peptide (pbUGT1*6)*

The construction of pbUGT1*6 was performed according to the same strategy than that developed for pUGT1*6 described above except that the synthetic oligomer E, 5′-CGCCATGGACAAGCTGCTGGTGGTC-3′ was used as N-terminal primer. This primer was designed to create a *Nco*I site and hybridized to a sequence encoding the first 6 amino acid residues of the mature protein. The PCR product was digested with *Nco*I and *Bam*HI generating a fragment coding for the mature UGT1*6 protein, which was isolated and subcloned into *Nco*I and *Bam*HI sites of pET-20b in frame with the sequence coding for the pelB signal peptide. The final recombinant plasmids were sequenced and used for overexpression.

2.5. Expression of the recombinant proteins

The bacteria were grown in Luria-Bertani (LB, Difco) medium containing ampicillin (50 μg/ml) and chloramphenicol (25 μg/ml), at 37°C until an OD_{600} of 0.5–0.6 was reached. IPTG (0.4 mM) was added and the incubation was continuated for 2.5 h. The cells were harvested by centrifugation at 5,000 × *g* for 15 min at 4°C.

2.6. Subcellular fractionation

The cells were resuspended in 50 mM Tris-HCl (pH 8.0) buffer containing 1 mM EDTA and 1 mM PMSF (lysis buffer) and were lysed by three 10-s pulses from a 0.5 cm diameter probe of a sonicator (Bioblock Vibra-cell) at 60% of maximal power. The sonicated lysates were then centrifuged at 5,000 × *g* for 10 min to remove cell debris. The supernatants were further centrifuged at 12,000 × *g* for 20 min to precipitate inclusion bodies [11]. The supernatant of the 12,000 × *g* centrifugation was subjected to a further 100,000 × *g* centrifugation for 60 min to separate cell membranes from cytosol, as previously described [12]. Sodium carbonate washes were carried out by homogenization of membrane pellets in 100 mM Na_2CO_3 (pH 11.0), as previously described [13]. Following incubation with the sodium carbonate solution on ice for 10 min, membranes were pelleted again by centrifugation.

2.7. Protein analysis

Protein concentration was measured according to Bradford [14] using Bio-Rad reagent with bovine serum albumin as standard. SDS-PAGE was performed according to Laemmli [15]. After electrophoresis, the proteins were transferred onto Immobilon-P^R membrane (Millipore). Detection of the recombinant protein was carried out using polyclonal anti-rat liver anti-UGT antibodies and anti-sheep IgG $(Fab')_2$-alkaline phosphatase conjugate (Jackson Immunoreagents) as secondary antibody as previously described [16].

3. Results

The coding sequence of mature UGT1*6 protein was fused to the natural UGT1*6 signal peptide or to two different bacterial signal peptides, ^{s}OmpT and ^{s}pelB. The amino acid sequences of the signal peptides are listed in Fig. 1 and differ in the length, in the degree of hydrophobicity of the core region and in the turn region preceeding the signal peptidase cleavage site.

The expression cassettes used were under the control of the T7 gene 10 promoter and ribosome binding site [10]. The production of UGT1*6 protein from these vectors was compared in the same bacterial strain BL21(DE3)pLysS and in identical culture conditions. The results of Western blot analysis of the cell lysates obtained from IPTG-induced cells using anti-UGT antibodies are illustrated in Fig. 2. Derepression by IPTG resulted in the expression of the three recombinant pro-

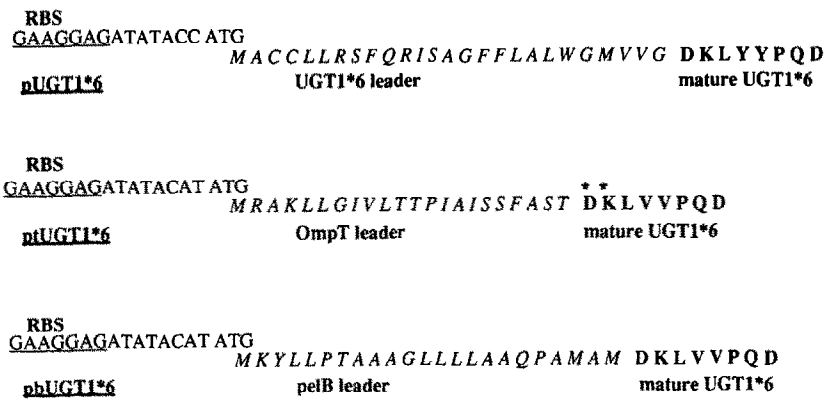

Fig. 1. Schematic representation of the expression cassette contained in pUGT1*6, ptUGT1*6 and pbUGT1"*6 constructs designed to express UGT1*6 with natural and bacterial signal peptides in bacteria. The ribosome binding site (RBS) sequence is underlined. Mutated amino acid residues of the mature UGT1*6 sequence are indicated by an asterisk.

teins. A single polypeptide of an apparent molecular mass of 55,000 Da was produced from ^{s}OmpT-UGT1*6 sequence (Fig. 2, lane a), whereas expression of ^{s}pelB-UGT1*6 gave rise to two polypeptides of 52,000 and 55,000 Da (Fig. 2, lane b). In the case of the full-length UGT1*6 cDNA, a single polypeptide of an apparent molecular mass of 52,000 Da was expressed at a lower level than in the previous cases (Fig. 2, lane c). Expression of UGT1*6, as a mature form, from a truncated cDNA lacking the signal peptide coding sequence resulted in the appearance of a polypeptide band exhibiting the same apparent molecular mass of 52,000 Da (data not shown). These results suggested that the natural signal peptide was cleaved indicating that the precursor of UGT1*6 was correctly processed in bacteria (Fig. 2, lane c). Approximatively 30% of ^{s}pelB-UGT1*6 precursor was processed (Fig. 2, lane b) but no maturation of ^{s}OmpT-UGT1*6 precursor was observed (Fig. 2, lane a).

Interestingly, we have observed a significant maturation of ^{s}OmpT-UGT1*6 when changing the first two amino acid residues (Asp-1, Lys-2) of the mature protein into (Pro-1, Glu-2) of ^{s}OmpT-UGT1*6, as evidenced by the appearance of a second polypeptide of 52,000 Da recognized by immunoblot analysis (Fig. 3, lane c). This result suggested that the first two amino acid residues of the mature sequence were important for the conformation of the cleavage site and that the ^{s}OmpT-UGT1*6 protein was a poor substrate for the signal peptidase.

The subcellular localization of the expressed UGT1*6 proteins from the various constructs was determined by fractionation of IPTG-induced cell homogenates. Interestingly, the recombinant proteins were exclusively associated with the membrane fractions (Fig. 4). Treatment of the membrane fraction by sodium carbonate, a conventional treatment used to remove non-integral membrane proteins [13], did not release any of the three recombinant proteins from the membrane fractions. It is noteworthy that immunoblot analysis of membranes containing the recombinant proteins expressed with bacterial signal peptide showed that the processing of these polypeptides reached about 50% (Fig. 4, lane a and b). This was explained by the observation that part of the precursors ^{s}OmpT-UGT1*6 and ^{s}pelB-UGT1*6 precipitated as inclusion bodies (Fig. 4, lanes e and f).

4. Discussion

For the first time, the expression of a UGT isoform in a prokaryotic expression system was successfully achieved and the influence of the signal peptide, which may affect the rate of synthesis and the fate of the heterologous membrane-bound protein UGT1*6 in *E. coli*, was investigated.

For this purpose, expression plasmids carrying the coding sequence of mature human UGT1*6 fused to the

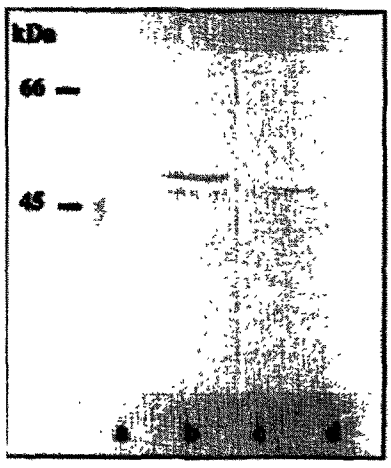

Fig. 2. Expression of native and modified UGT1*6 in *E. coli*. Western blot analysis was performed on total bacterial extracts harvested before and after IPTG-induction. Lane a, recombinant UGT1*6 expressed from the ptUGT1*6 vector (25 μg protein); lane b, recombinant UGT1*6 expressed from pbUGT1*6 vector (25 μg protein); lane c, recombinant UGT1*6 expressed from pUGT1*6 vector (50 μg protein); lane d, non-induced cells (50 μg protein).

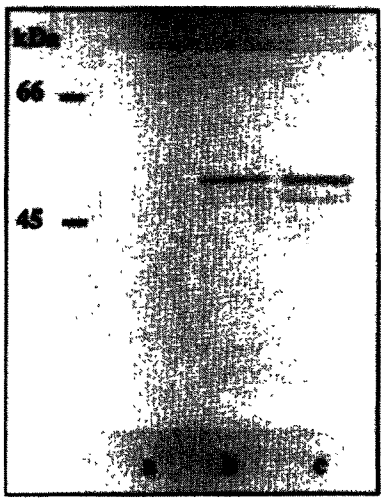

Fig. 3. Processing of non-mutated and mutated ^{s}OmpT-UGT1*6 recombinant proteins. Total bacterial extracts (25 μg protein per lane) harvested before and after IPTG-induction were analyzed by SDS-PAGE followed by immunoblotting. Lane a, non-induced cells; lane b, recombinant UGT1*6 expressed from ptUGT1*6; lane c, recombinant UGT1*6 expressed from mutated ptUGT1*6.

natural signal peptide of UGT1*6 and to two different efficient secretion signal sequences [17] have been constructed. The production of the two recombinant proteins ^{s}OmpT-UGT1*6 and ^{s}pelB-UGT1*6 was higher than that of the native protein derived from full-length UGT1*6 cDNA, demonstrating that the use of bacterial signal peptides is favorable for the recovery of large amounts of recombinant UGT protein. The relatively low yield of production of UGT1*6 with its own signal peptide may result from the instability of this precursor in bacteria and/or to poor translation initiation.

In addition, the human UGT1*6 containing its natural signal peptide was apparently fully processed in bacteria. This result is in agreement with the hypothesis that UGTs are expressed as precursors with a cleaved signal sequence. As previously shown by Talmadge et al. [18] and by Watts et al. [19], *E. coli* leader peptidase correctly processed a pre-proinsulin and a mouse IgG κ-Chain fragment precursor, respectively, suggesting that the recognition between pre-proteins and leader peptidase was specific and conserved during evolution.

When the mature sequence of UGT1*6 was fused with pelB signal peptide, significant amount of processed protein was observed. The incomplete processing of this precursor could be due to the high rate production of this polypeptide, suggesting that the leader peptidase activity may be rate limiting. In the case of ^{s}OmpT-UGT1*6, no processing of the precursor occurred and we found that physiological parameters, such as temperature, duration of induction period and composition of the culture medium did not cause major changes in the extent of the maturation process.

Since signal peptides do not have a strict consensus sequence, one might expect these sequences to be fully interchangeable, but numerous studies, including our report, indicate that this is not the case. We provide further evidence that the primary sequence of the signal peptide must also be compatible with the mature polypeptide to which it is linked [20,21]. This observation may be explained by thermodynamic considerations of the conformation of the precursors, a matter which really concerns the whole protein and not the signal peptide only [20].

Mutation studies provided evidence that a region around the cleavage site that includes the last four amino-acid residues of the signal sequence and up to the first two amino-acid residues of the mature sequence is known to be important in the processing of several precursors [23]. In the case of the ^{s}OmpT-UGT1*6 presursor, the mutation of the two first amino acid residues of UGT1*6 mature sequence improved the maturation process providing further evidence that the sequence around the processing site is important for recognition by signal peptidase possibly due to the preference for some amino acid residues or merely for steric consideration.

UGT1*6 synthesized from the natural precursor was associated to the bacterial membrane. In addition, both the precursors ^{s}OmpT-UGT1*6 and ^{s}pelB-UGT1*6 were membrane-bound suggesting that cleavage of the signal peptide was not necessary for translocation, and that the leader peptidase did not catalyse the translocation of pre-proteins accross the membrane [22]. Binding of the recombinant proteins to the membrane is probably due to the carboxy-terminal stop transfer sequence of the UGT1*6 protein [7], which abolishes further export and

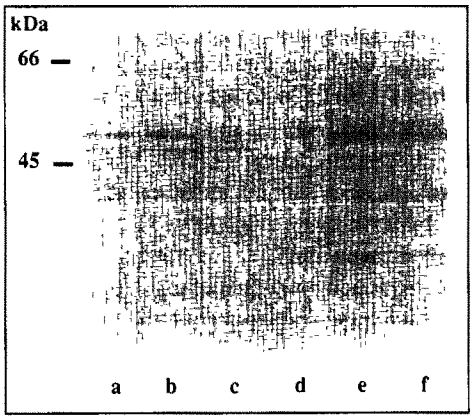

Fig. 4. Membrane localization of recombinant UGT1*6 expressed with natural and bacterial signal peptides. Subcellular fractions were prepared from bacteria harvested before and after IPTG-induction and analyzed by immunoblotting. Lane a, membrane fraction from bacteria transformed with mutated ptUGT1*6 (25 μg protein); lane b, membrane fraction from bacteria transformed with pbUGT1*6 (25 μg protein); lane c, membrane fraction from bacteria transformed with pUGT1*6 (50 μg protein); lane d, membrane fraction from non-transformed cells (50 mg protein); lane e, insoluble fraction from bacteria transformed with mutated ptUGT1*6 (25 μg protein); lane f, insoluble fraction from bacteria transformed with pbUGT1*6 (25 μg protein).

consequently integrates the protein into the membrane via the transmembrane anchor segment.

This report describes for the first time the expression of a human UGT isoform in *E. coli*. Targeting the protein to the membranes, like in its natural environnement, was achieved by using natural and bacterial signal peptides. This approach will be applicable to the other members of this large family of isoenzymes and will provide appropriate tools for further structural and physicochemical analysis.

Acknowledgements: This work was supported by grants from the CNRS and INSERM (CRE 920308), from the Institut des Biotechnologies (Nancy, France) and from the Wellcome Trust (UK). T.P. is the recipient of a research grant from Laboratoires SERVIER (Orléans, France). Dr. H. Souhaili El Amri is gratefully acknowledged for constant support to this work.

References

[1] Burchell, B., Nebert, D.W., Nelson, D.R., Bock, K.W., Iyanagi, T., Jansen, P.L.M., Lancet, D., Mulder, G.J., Chowdhury, J.R., Siest, G., Tephly, T.R. and Mackenzie, P.I. (1991) DNA Cell. Biol. 10, 487–494.
[2] Mulder, G.J. (1992) Annu. Rev. Pharmacol. Toxicol. 32, 25–49.
[3] Harding, D., Fournel-Gigleux, S., Jackson, M.R. and Burchell, B. (1988) Proc. Natl. Acad. Sci. USA 85, 8381–8385.
[4] Fournel-Gigleux, S., Sutherland, L., Sabolovic, N., Burchell, B. and Siest, G. (1991) Mol. Pharmacol. 39, 177–183.
[5] Bock, K.W., Forster, A., Gschaidmeier, H., Brück, M., Münzel, P., Schareck, W., Fournel-Gigleux, S. and Burchell, B. (1993) Biochem. Pharmacol. 45, 1809–1814.
[6] Nilsson, T., Jackson, M. and Peterson, P.A. (1989) Cell 58, 707–718.
[7] Mackenzie, P. (1986) J. Biol. Chem. 261, 6119–6125.
[8] Fournel-Gigleux, S., Jackson, M.R., Wooster, R. and Burchell, B. (1989) FEBS Lett. 243, 119–122.
[9] Duffaud, G.D., March, P.E. and Inouye, M. (1987) Methods Enzymol. 153, 492–506.
[10] Studier, F.W., Rosenberg, A.H., Dunn, J.J. and Dubendorff, J.W. (1990) Methods Enzymol. 185, 60–89.
[11] Marston, F.A.O. (1987) in: DNA Cloning. Volume III. Practical Approach Series, pp. 59–88 (D. Rickwood and D.D. Hames, Eds.).
[12] Clark, B.J. and Waterman, M.R. (1991) J. Biol. Chem. 266, 5898–5904.
[13] Fujiki, Y., Hubbard, A.L., Fowler, S. and Lazarow, P.B. (1982) J. Cell Biol. 93, 97–102.
[14] Bradford, M.M. (1976) Anal. Biochem. 72, 248–254.
[15] Laemmli, U.K. (1970) Nature 227, 680–685.
[16] Coughtrie, M., Burchell, B., Leakey, J.E.A. and Hume, R. (1988) Mol. Pharmacol. 34, 729–735.
[17] Stader, J.A. and Silhavy, T.J. (1990) Methods Enzymol. 185, 166–187.
[18] Talmadge, K., Kaufman, J. and Gilbert, W. (1980) Proc. Natl. Acad. Sci. USA 77, 3988–3992.
[19] Watts, C., Wickner, W. and Zimmermann, R. (1984) Proc. Natl. Acad. Sci. USA 80, 2809–2813.
[20] Denèfle, P., Kovarik, S., Ciora, T., Gosselet, N., Bénichou, J.C., Latta, M., Guinet, F., Ryter, A. and Mayaux, J.F. (1989) Gene 85, 499–510.
[21] Pugsley, A.P. (1989) Protein Targeting, Academic Press, San Diego.
[22] Dalbey, R.E. and Wickner, W. (1985) J. Biol. Chem. 260, 15925–15931.
[23] Koshland, D., Sauer, R.T. and Botstein, D. (1982) Cell 30, 903–914.