Acta Biochimica Polonica

Vol. 42 No. 1/1995

109–114

QUARTERLY

Synthesis, cloning and expression in *Escherichia coli* of the gene coding for the trypsin inhibitor from *Cucurbita pepo*

Bożenna Rempoła[a], Tadeusz Wilusz[b], Wojciech Markiewicz[c] and Magdalena Fikus[a]*

[a]*Institute of Biochemistry and Biophysics of the Polish Academy of Sciences, A. Pawińskiego 5a, 02–106 Warsaw, Poland*
[b]*Institute of Biochemistry, University of Wrocław, Wrocław, Poland*
[c]*Institute of Bioorganic Chemistry of the Polish Academy of Sciences, Poznań, POland*

Received: 5 January, 1995

Key words: proteinase inhibitor; gene cloning; *Escherichia coli*; expression and secretion

A chemically synthesized gene coding for the serine proteinase inhibitor CPTI II was cloned in *E. coli* and its expression was investigated in cytoplasmic and secretion systems. Under all conditions investigated the biologically active form of the inhibitor was found only in the latter system, although the yield was rather low.

It has been suggested that proteinase inhibitors play an important role in many vital processes. Various inhibitors were found in different cells and tissues of animal and plant origin, and their structure and function have frequently been studied [1].

Several short polypeptides, about 30 amino acids in length, with highly conserved primary structure, exhibiting trypsin inhibitory activity were identified and isolated from the seeds of *Cucurbitaceae* plants family [2–6]. Their compact structure, stabilized by three S-S bridges, makes these proteins extremely resistant to the action of various biological and physico-chemical agents. Their biological function in plants is unknown.

Polypeptides of this type are attractive models for studies on protein folding processes with formation of transient and final stable tertiary structures. They may be investigated in the form of protein:protein complexes with corresponding serine proteinases. In plants, identification of the respective genes can help in determining the role of these inhibitors at different stages of plant development. Substantial amounts of the inhibitor and parallel investigations of the plant genome are needed for these purposes.

Inhibitors from the squash family have been isolated from the seeds [2–4, 7] in a wide range of yields (1–100 mg/kg). Chemical synthesis of these polypeptides and of their derivatives was found possible [8], but the yield was rather low, and the overall procedure costly. As an alternative, recombinant DNA technology was used with various degrees of success [7, 9, 10].

In this work expression of the synthetic gene for CPTI II in *E. coli* was investigated in two different systems. In the first system cytoplasmic accumulation of the expression product was expected, whereas in the second one the recombinant protein was secreted to the surrounding medium.

---

*To whom correspondence should be adressed.

[1]Abbreviations used: BAPNA, Nα-benzoyl-DL-arginine-*p*-nitroanilid; CPTI, *Cucurbita pepo* trypsin inhibitor; LB, Luria broth; IPTG, isopropyl β-D-thiogalactopyranoside.

MATERIALS AND METHODS

Reagents and kits. All chemicals used were of the highest purity grade. Restriction enzymes, T4 DNA ligase, T4 polynucleotide kinase and the dideoxy sequencing kit were from Amersham. DNA Taq polymerase was from Cetus Perkin-Elmer.

Cloning procedures. Nucleotide sequences were determined by the dideoxy chain termination method. All standard cloning procedures were performed according to Sambrook *et al.* [11]. PCR was performed according to PCR Protocols [12].

E. coli strains. For cloning and expression of genes *E. coli* K12 strain DH5α, and *E. coli* B strain BL21(DE3)pLys S [13] were used as indicated.

Oligonucleotides. Oligonucleotides composing the gene sequence were synthesized by the phosphoramidite approach using 2-cyanoethyl phosphate protecting groups, according to [14].

Estimation of the yield of the recombinant inhibitor. Trypsin inhibitory activity was qualitatively determined by electrophoresis on polyacrylamide gels [15] with copolymerized edestin as a substrate for trypsin. The quantitative test was based on the determination of the residual trypsin activity against BAPNA as a substrate [16] after partial purification of CPTI II by Sepharose 4B anhydrotrypsin affinity chromatography [3].

Putative inclusion bodies were prepared and processed according to Staley and Kim [17].

RESULTS

Design and construction of the gene coding for CPTI II

The nucleotide sequence coding for CPTI II was programmed according to its amino-acid sequence [2] and codon preference established for *E. coli* [18], except for lysine^5 AAA changed to AAG in order to create a *Bgl*II restriction site. (Fig. 1).

All oligonucleotides, except the terminal ones, were 5′ phosphorylated with polynucleotide kinase, annealed and ligated in a single step. The ligation product of the expected size was isolated from a polyacrylamide gel, ligated to plasmid pUC18 and sequenced.

Construction of the cytoplasmic system for expression of CPTI II.

The cytoplasmic system consisted of the *E. coli* B strain (see Materials and Methods) and pMW172 vector which make possible specific transcription of the gene cloned under the T7 RNA polymerase promoter.

The *Eco* RI/*Bgl*II fragment of the gene coding for CPTI II cloned in pUC 19, was exchanged for the synthetic oligonucleotide *Bam*HI/*Bgl*II; this was confirmed by sequencing. The *Bam*HI/*Hin*dIII fragment containing the gene coding for CPTI II was inserted downstream from the inducible promoter of the expression vector pMW 172 kindly offered by Dr. M. Way to Dr. K. Bolewska of the Institute of Biochemistry and Biophysics. The physical map of the recombined vector was confirmed by restriction analysis. The short *Nde*I fragment was removed and *E. coli* BL21(DE3)/pLys S cells were transformed with pMW172--CPTI II plasmid (Fig. 2).

Cells were grown at 37°C in LB medium supplemented with ampicillin (100 μg/ml), chloramphenicol (30 μg/ml), and when the culture reached A_{600} = 1.0, expression of the cloned gene was induced with IPTG (0.4 mM). They were harvested and frozen 3 h later.

After thawing, the cells were sonicated and centrifuged. Antitrypsin activity was qualitatively determined in samples of the concentrated crude extract, corresponding to 4 or 8 ml of the initial culture.

Direct expression of the gene to yield an active form of the inhibitor in the cytoplasm was not qualitatively detected in cellular lysates and in putative inclusion bodies.

Construction of the secretion — expression system for CPTI II synthesis

In this system, vector sequences coding for the signal peptide and transcription termination signal were from the *E. coli* gene coding for the outer membrane Omp A protein. Expression was regulated by the efficient *lpp* gene constitutive promoter and the inducible *lac*UV5 promoter/operator fragment [19]. These sequences were excised from pA2T1 plasmid kindly offered by Dr. Ulrich Hahn.

The *SphI/Hin*dIII fragment of the above plasmid, containing the entire vector sequences, was ligated to a 104 bp sequence coding for CPTI II, giving rise to pBM9 plasmid (Fig. 3). Site directed mutagenesis based on the PCR method was used to create a proper reading frame downstream from the coding region of the OmpA signal peptide. Subsequently, the *XbaI/Hin*dIII fragment containing the cloned gene and the upstream sequences coding for the leader peptide were used as a template for PCR. The PCR product was cloned and sequenced in pUC 19.

Restriction fragment *XbaI/Hin*dIII from the above construct was ligated to the vector *XbaI/Hin*dIII fragment from plasmid pA2T1, and *E. coli* DH5-α strain was transformed with the resulting plasmid pIN-III-ompA2-CPTI II (Fig. 3). Transformants were selected after restriction analysis of plasmids from single colonies. Preliminary experiments showed antitrypsin activity secreted by transformed clones to the culture medium, following IPTG induction.

The growth and induction conditions of transformed *E. coli* cells giving the maximum yield of the recombinant peptide were established. *E. coli* cells containing the plasmid were grown in shaker flasks at 37°C in LB medium with ampicillin (100 μg/ml). IPTG was added and induction continued for the next 20 h. Cells were harvested and trypsin inhibitory activity was qualitatively estimated in the periplasm [20] and in culture medium, after its concentration by freeze-drying.

All conclusions were based on visual estimations of active fractions on the gel, compared to appropriate controls run under identical conditions (Fig. 4). All preparations migrated as single active bands.

The search for maximal yields of the expression product is summarized in Table 1. Optimal growth conditions were established as: LB medium, 37°C, induction at the mid-log phase with 2 mM IPTG for 24 h.

Antitrypsin activity was not found in 10-fold concentrated periplasmic solution and in putative inclusion bodies. The rate of growth of the cells containing the recombinant plasmid under all conditions tested was identical with that for the cells containing the vector.

CPTI II was partly purified from the culture medium by anhydrotrypsin affinity chromatography. Fractions showing elevated absorption at 260 μm were quantitatively assayed for antitrypsin activity. Under optimal conditions the

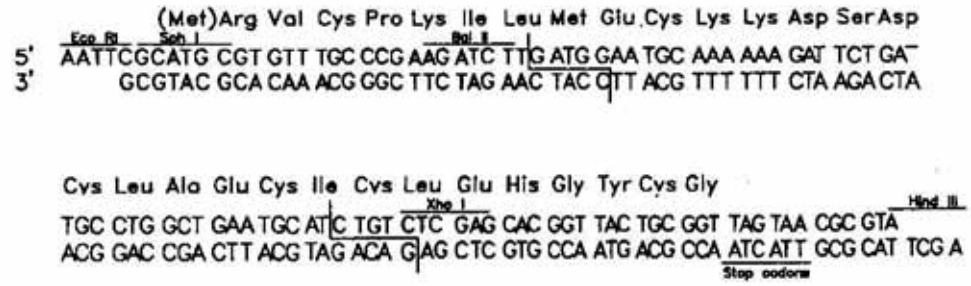

Fig. 1. Nucleotide sequence of the chemically synthesized gene coding for CPTI II, stop codons included, with amino-acid sequence of the corresponding inhibitor. Restriction sites and the size of appropriate oligonucleotides ligated to yield the gene sequence are indicated.

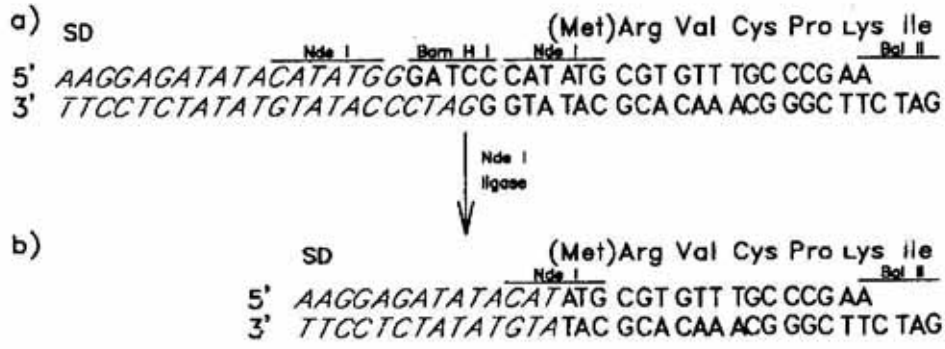

Fig. 2. (a) Nucleotide sequence connecting the pMW172 vector (fine italic lettering) with the synthetic gene coding for CPTI II. This sequence was modified by restriction with NdeI and subsequent ligation, giving rise to (b) the final nucleotide sequence connecting vector regulatory sequences to the synthetic gene. SD, Shine-Daelgarno sequence.

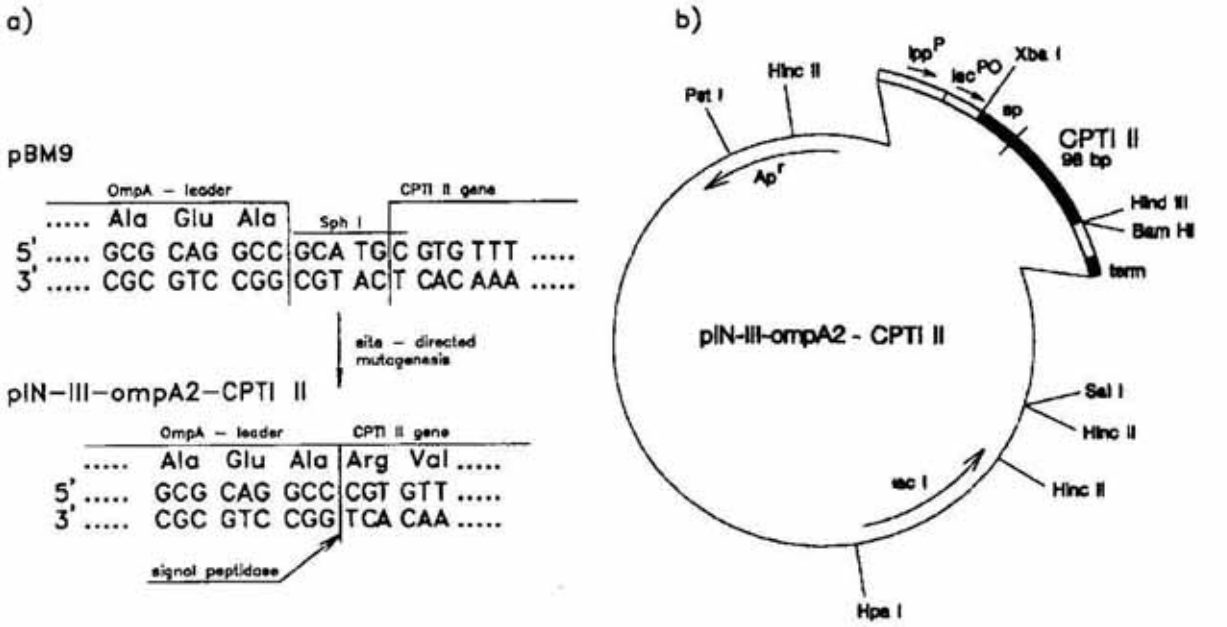

Fig. 3. Construction of the secretion — expression recombinant plasmid pIN-III-ompA2-CPTI II.
(a) Nucleotide sequence at the junction: ompA-leader — CPTI II gene in pBM9 plasmid. This region was next subjected to the site-directed mutagenesis *via* PCR; the resulting "in frame" construction is shown, together with the amino-acid sequence cleaved by the signal peptidase. (b) Physical map of the recombinant plasmid pIN-III-ompA2-CPTI II. sp, signal peptide; term, transcription terminator; lpp^P-lac^{PO}, constitutive and inducible promoters and *lac* operator region; appropriate restriction sites and selection marker are indicated.

yield of the active inhibitor was estimated as 60 μg per 1 litre of culture.

DISCUSSION

In the reducing conditions of the bacterial cytoplasm most of the cysteine-containing proteins do not easily fold into the properly oxidized form. However, several proteins and at least two different enzymes which catalyze oxidation of SH groups have been found in the periplasmic space of *E. coli* [21–23], and their assistance in protein folding was postulated. This suggested the use of secretion vectors for obtaining biologically active, oxidized CPTI II [19, 24].

When the *E. coli* secretion system was used for expression of recombinant CPTI II the yield was rather low. We believe this could be attributed to such properties of the system studied as:

– low molecular mass of the inhibitor. To our knowledge CPTI II is one of the smallest peptides synthesized in the secretion system of *E. coli* [25–27],

Fig. 4. Qualitative electrophoretic assay of trypsin inhibition.
White, nontransparent "bubbles" consist of undigested edestin, i.e. indicate localization of the band of the active inhibitor. If not otherwise indicated, samples are equivalents of 1.5 ml of culture. Lanes a, b, c: induction at time = 0 by 2 mM IPTG at 28°C, samples were taken 5, 12 and 24 h later, respectively; lanes d, e: inducer (2 mM IPTG) was added when culture reached at 28°C A_{600} = 0.3, samples were taken 8 and 20 h later, respectively; lane f: periplasmic sample equivalent to 15 ml of culture, standard conditions; lanes g, h, i: induction at time = 0 by 2 mM IPTG at 37°C, samples were taken 5, 12 and 24 h later, respectively; lanes j, k, l: inducer (2 mM IPTG) was added when culture at 37°C reached A_{600} = 0.4, samples were taken 2.5, 11.5 and 22.5 h later, respectively.

Table 1
Optimization of recombinant CPTI II secretion

Medium	IPTG conc. (mM)	Temperature (°C)	Induction time (h) (Inoculation of culture = 0)		Samples were taken at the following times after induction (h)			
LB	0.2	28	4	(A_{600} 0.3)	0	5	12	24
		37	2.5	(A_{600} 0.4)	0	5	12	24
	2.0	28	0	(A_{600} 0.065)	0	5	12	24
			4	(A_{600} 0.3)	0	5	12	24
		37	0	(A_{600} 0.090)	0	5	12	24
			2.5	(A_{600} 0.4)	0	5	13	24
M9 CAS	2.0	37	4	(A_{600} 0.1)	0		12	24

– positively charged N-terminal Arg^1 residue of the mature inhibitor amino-acid sequence. This was often implied as a factor lowering or even precluding secretion [28–30].

The recombinant CPTI II was not found in the periplasmic space and was secreted to the medium. This was probably due to its low molecular mass [24].

The failure of experiments aiming at cytoplasmic expression of the gene coding for CPTI II could be explained as due mainly to the metabolic instability of expression products (mRNA, polypeptide) [31]. Moreover, the low yield of CPTI II production in the secretion system indicates that a more efficient system of expression may be required.

Such phenomena as the instability of mRNA, intracellular proteolysis prior to translocation, incorrect folding, aggregation, incomplete processing, defects in translocation, were not investigated because the available assay system for proteinase inhibitors of this family allows only the estimation of the biologically active peptide.

We are greatly indebted to Dr. Ulrich Hahn from Institut für Kristallographie, Freie Universität Berlin for offering plasmid pA2T1. The excellent technical assistance of Ms. Teresa Rak is appreciated.

REFERENCES

1. *Innovations in Proteases and their Inhibitors.* (1993), (Aviles, F. X., ed.) pp. 530, Walter de Gruyter & Co., Berlin, New York.
2. Wieczorek, M., Otlewski, J., Cook, J., Parks, K., Leluk, J., Wilimowska-Pelc, A., Polanowski, A., Wilusz, T. & Laskowski, M., Jr. (1985) The squash family of serine proteinase inhibitors. Amino acid sequences and association equilibrium constants of inhibitors from squash, summer squash, zucchini, and cucumber seeds. *Biochem. Biophys. Res. Commun.* **126**, 646–652.
3. Otlewski, J., Whatley, H., Polanowski, A. & Wilusz, T. (1987) Amino-acid sequences of trypsin inhibitors from watermelon (*Citrullus vulgaris*) and red bryony (*Bryonia dioica*) seeds. *Biol. Chem. Hoppe-Seyler* **368**, 1505–1507.
4. Otlewski, J. (1993) The squash inhibitors of serine proteinases; in *Innovations in Proteases and their Inhibitors* (Aviles, F.X., ed.) pp. 369–388, Walter de Gruyter & Co., Berlin, New York.
5. Matsuo, M., Hamato, N., Takano, R., Kamei-Hayashi, K., Yasuda-Kamatani, Y., Nomoto, K. & Hara, S. (1992) Trypsin inhibitors from bottle gourd (*Lagenaria leucantha Rusby var. Depressa Makino*) seeds. Purification and amino acid sequences. *Biochim. Biophys. Acta* **1120**, 187–192.
6. Favel, A., Mattras, H., Coletti-Previero, M.A., Zwilling, R., Robinson, E.A. & Castro, B. (1989) Proteinase inhibitors from *Ecballum elaterium* seeds. *Int. J. Peptide Protein Res.* **33**, 202–208.
7. Ling, M.-H., Qi, H.-y. & Chi, C.-w. (1993) Protein, cDNA and genomic DNA sequences of the towel gourd trypsin inhibitor. *J. Biol. Chem.* **268**, 810–814.
8. Rolka, K., Kupryszewski, G., Ragnarsson, U., Otlewski, J., Krokoszyńska, I. & Wilusz, T. (1991) Chemical synthesis of new trypsin, chymotrypsin and elastase inhibitors by amino-acid substitutions in a trypsin inhibitor from squash seeds (CMTI III). *Biol. Chem. Hoppe-Seyler* **372**, 63–68.

9. Chen, X.-M., Qian, Y.-W., Chi, C.-W., Gan, K.-D.E., Zhang, M.-F. & Chen, C.-Q. (1991) Chemical synthesis, molecular cloning and expression of the gene coding for the *Trichosantes* trypsin inhibitor. *J. Biochem.* **112**, 45–51.

10. Rempoła, B. (1993) Trypsin inhibitor gene cloned and expressed in *Saccharomyces cerevisiae*. *Acta Biochim. Polon.* **40**, 8–11.

11. Sambrook, J., Fritch, E.F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, vol. 1, 2nd edn., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

12. *PCR Protocols. A Guide to Methods and Applications* (1990) (Innis, M.A., Gelfand, D.H., Sninsky, J.J. & White, T.J., eds.) Academic Press, Inc., Harcourt Brace Jovanovich, Publ. San Diego, New York, Berkeley, Boston.

13. Pagratis, N.C. & Revel, H.R. (1993) Broad host range, regulated expression system utilizing bacteriophage T7 RNA polymerase and promoter. *Biotechnology and Bioengineering* **41**, 837–845.

14. *Oligonucleotide Synthesis: A Practical Approach* (1984) (Gait, M.J., ed.), IRL Press, Oxford.

15. Pham, T.-C., Konopska-Waliszkiewicz, J. & Leluk, J. (1986) Trypsin inhibitors in the aleurone grains of *Cucurbita pepo* var. *patissonia* (white bush) cotyledons. *Biochem. Physiol. Pflanzen* **181**, 565–569.

16. Erlanger, B.F., Kokowsky, N. & Cohen, W. (1961) The preparation and properties of two new chromogenic substrates of trypsin. *Arch. Biochem. Biophys.* **95**, 271–278.

17. Staley, J.P. & Kim, P.S. (1994) Formation of a native-like subdomain in a partially-folded intermediate of bovine pancreatic trypsin inhibitor. *Protein Science* **10**, 1822–1832.

18. Sharp, P.M., Cowe, E., Higgins, D.G., Shields, D.C., Wolfe, K.H. & Wright, F. (1988) Codon usage pattern in *Escherichia coli, Bacillus subtilis, Saccharomyces cerevisiae, Schizosaccharomyces pombe, Drosophila melanogaster,* and *Homo sapiens*; a review of the considerable within-species diversity. *Nucleic Acids Res.* **16**, 8207–8211.

19. Duffaud, G.D., March, P.E. & Inouye, M. (1987) Expression and secretion of foreign proteins in *Escherichia coli*. *Methods Enzymol.* **153**, 492–507.

20. Quaas, R., McKeown, Y., Stanssens, P., Frank, R., Blocker, H. & Hahn, U. (1988) Expression of the chemically synthesized gene for ribonuclease T1 in *Escherichia coli* using a secretion cloning vector. *Eur. J. Biochem.* **173**, 617–622.

21. Joly, J.C. & Swartz, J.R. (1994) Protein folding activities of *Escherichia coli* protein disulfide isomerase. *Biochemistry* **33**, 4231–4236.

22. Missiakas, D., Georgopoulos, C. & Raina, S. (1994) The *Escherichia coli* Dsbc (Xpra) gene encodes a periplasmic protein involved in disulfide bond formation. *EMBO J.* **13**, 2013–2020.

23. Wulfing, C. & Pluckthun, A. (1994) Protein folding in the periplasm of *Escherichia coli*. *Molec. Microbiol.* **12**, 685–692.

24. Molina, M.A., Aviles, F.X. & Querol, E. (1992) Expression of a synthetic gene encoding potato carboxypeptidase inhibitor using a bacterial secretion vector. *Gene* **116**, 129–138.

25. Maywald, F., Boldicke, T., Gross, G., Frank, R., Blocker, H., Meyerhans, A., Schwellnus, K., Ebbers, J., Bruns, W., Reinhardt, G., Schnabel, E., Schroder, W., Fritz, H. & Collins, J. (1988) Human pancreatic secretory trypsin inhibitor (PSTI) produced in active form and secreted from *Escherichia coli*. *Gene* **68**, 357–369.

26. MacIntyre, S. & Henning, U. (1990) The role of the mature part of secretory proteins in translocation across the plasma membrane and in regulation of their synthesis in *Escherichia coli*. *Biochimie* **72**, 157–167.

27. Pang, S.-Z., Oberhaus, S.M., Rasmussen, J.L., Knipple, D.C., Bloomquist, J.R., Dean, D.H., Bowman, K.D. & Sanford, J.C. (1992) Expression of a gene encoding a scorpion insectotoxin peptide in yeast, bacteria and plants. *Gene* **116**, 165–172.

28. Hussain, K., Bowler, C., Roberts, L.M. & Lord, J.M. (1989) Expression of ricin B chain in *Escherichia coli*. *FEBS Lett.* **244**, 383–387.

29. Summers, R.G., Harris, C.R. & Knowles, J.R. (1989) A conservative aminoacids substitution, arginine for lysine, abolishes export of a hybrid protein in *Escherichia coli*. *J. Biol. Chem.* **264**, 20082–20088.

30. Boyd, D. & Beckwith, J. (1990) The role of charged amino acids in the localization of secreted and membrane proteins. *Cell* **62**, 1031–1033.

31. Enfors, S.-O. (1992) Control of *in vivo* proteolysis in the production of recombinant proteins. *Trends Biotechnol.* **10**, 310–315.