Volume 269, number 1, 244-246 FEBS 08778 August 1990

Influence of the second and third codon on the expression of recombinant hirudin in *E. coli*

Eric Degryse

Transgene S.A., 11 rue de Molsheim, 67000 Strasbourg, France

Received 7 June 1990; revised version received 3 July 1990

The effect of all possible codons corresponding to the second and third amino acid (isoleucine and threonine) on the expression level of hirudin in *E. coli* has been analysed. These levels could not be correlated with changes in primary and secondary mRNA structure. A decrease in the rate of synthesis and of product accumulation follows the introduction for ile of the ATA codon which is of very low usage, and for thr of the ACC codon, which results in homology of the mRNA with the 3'-end of 16S rRNA. The results are discussed according to current concepts of protein expression in *E. coli*

Recombinant hirudin; Expression; Codon usage; *E. coli*

1. INTRODUCTION

Gene expression in *E. coli* has been shown to be influenced by many factors. Of relevance to this study are: (i) Codon usage: for many analysed *E. coli* genes the expression level is correlated with the frequency of occurrence of codons [1] and with the cognate tRNA concentration [2]. (ii) The primary structure of the mRNA; the ribosome binding site [3] and the sequence following the initiation codon, which might bind the 5'-end of the 16S rRNA [4]. Both make up the translation initiation region. (iii) The secondary structure in which the ribosome binding site and the initiating ATG are presented to the ribosome [5].

To gain a better understanding of the factors controlling expression of heterologous genes in *E. coli*, the expression level of recombinant hirudin, driven by the strong PL promoter of bacteriophage lambda, was analysed for all possible codon combinations for ile and thr, the second and third amino acid.

This was surmised to be a sensitive model since: (i) The initial steps in the translation process will determine whether protein synthesis proceeds or aborts and sets the overall rate. (ii) Influences on the rate of protein synthesis will be more pronounced for a small gene, like hirudin (198 nucleotides).

Finally, at the end of the induction procedure, hirudin is one of the very few proteins still synthesised. Use of this model excludes therefore possible competition with the synthesis of other proteins.

Correspondence address: E. Degryse, Transgene S.A., 11 rue de Molsheim, 67000 Strasbourg, France

2. MATERIALS AND METHODS

When not explicitly stated, materials and methods were as previously described [5,6].

Synthetic oligonucleotides (TG 460 5' TATGATHACNT 3' and TG461 5' ATANGTDATCA 3', corresponding to all codon combinations for ile and thr, were phosphorylated, hybridised at 65°C, cooled at 4°C and ligated with pTG720 [6], cut with restriction enzymes NdeI and AccI and treated with alkaline phosphatase. Initial transformants were sequenced by T track analysis and segregated by transformation at low concentration. Candidates were confirmed by sequence analysis with all four nucleotides. Possible RNA secondary structures were analysed as described previously for α_1-antitrypsin [5].

E. coli strain TGE901 (F^- su ilv his (lambda delta Bam cI857 delta H1) transformed with the different plasmids was grown at 30°C and induced at 37°C as described [6], except that LB medium was used throughout.

De novo synthesis was analysed by pulse labelling a sample of the bacterial culture corresponding to 1 ml of 1 unit A600 with 1/100 volume of [^{35}S]cysteine (Amersham; specific activity >600 Ci/mmol) for 90 s at 37°C. Subsequent treatment of the samples was as described previously [5].

Hirudin activity was measured as described [7], after storage of the bacterial samples for at least one week at −20°C. Values are expressed in anti-thrombin units/A600/l [6].

3. RESULTS AND DISCUSSION

Hirudin accumulation was after 2 h, 3 h, 4 h and 5 h of induction. De novo synthesis was followed by densitometric scanning of the autoradiographs, using the N gene product (at molecular weight 14 kDa) as an internal standard. This analysis reveals that the relative intensity of the hirudin bands observed for the different plasmids (after 3 h and 5 h of induction, not shown) correlates well with the hirudin activity measured at these time points. Table I shows the quantity of recombinant hirudin activity after 4 h of induction together

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/90/$3.50 © 1990 Federation of European Biochemical Societies

Table I

Recombinant hirudin accumulation on induction

Plasmid number	Sequences	Induction level at 4h	Structure	ΔG
pTG7201	cat atg ATT ACG tac	2200	dimer	−15.1
pTG7202	cat atg ATT ACC tac	420	monomer	−12.0
pTG7203	cat atg ATT ACT tac	1900	dimer	−12.0
pTG7204	cat atg ATT ACA tac	1600	monomer	−12.0
pTG7205	cat atg ATC ACG tac	1100	monomer	−13.7
pTG7206	cat atg ATC ACC tac	800	monomer	−10.6
pTG7207	cat atg ATC ACT tac	1000	monomer	−10.6
pTG7208	cat atg ATC ACA tac	1100	monomer	−10.6
pTG7209	cat atg ATA ACG tac	500	monomer	−13.3
pTG7210	cat atg ATA ACC tac	−	monomer	−11.6
pTG7211	cat atg ATA ACT tac	900	dimer	−11.6
pTG7212	cat atg ATA ACA tac	−	dimer	−12.5

The free energy (ΔG) of the secondary structure formed by the CII ribosome binding site and the initial coding region of the hirudin gene has been calculated in a region comprising nucleotides −53 till +30, defined with respect to the initiating ATG, where A is at the null position.

with the sequence analysis of the transformants obtained. Differences in expression level are not due to altered copy numbers of the plasmids (not shown) nor to their multimerisation state.

In a previous report, Tessier et al. [5] showed that stabilisation of the secondary structure of the mRNA containing the cII ribosome binding site and the region involving initiation of transcription of a lambda cII:α1-antitrypsin gene fusion correlates with increased expression levels. High-level expression of unfused α1-antitrypsin was obtained when this secondary structure was reconstructed, but only in the presence of certain silent point mutations in the α1-antitrypsin coding region. Using the same computer-based analysis for the hirudin mRNA structural features reminiscent of those reported were found for plasmid pTG7201. It explains its relatively high expression level (10^5 molecules/cell). The stability of this structure is equally reduced for plasmids pTG7205 and pTG7209 and correlates with a decreased expression. For the other plasmids, the computed structure no longer contains the redundant third nucleotide as part of a secondary structure, hence minor stability changes of the mRNA structure do not correlate with expression levels.

However, a single nucleotide change downstream of the initiating ATG can affect the product yields obtained (Table I). Noteworthy are the drops in hirudin expression level when ATA and ACC stand for ile and thr, respectively.

The ile codons ATT and ATC have a corresponding tRNA concentration 9–19-fold higher [2,8] than the ATA codon. Table I shows that the presence of ATA as second codon is indeed correlated with a decreased expression level (compare for example pTG7209 and pTG7205). Yet, pTG7211 yields an expression level exceeding that of certain plasmids containing ATT or ATC (e.g. pTG7102). It is to be noted that only pTG7211 generates a stretch (AACT) of 4 bases complementary to the 5′-end of 16S rRNA. As this region is suggested to be part of the translation initiation region [4], it could account for a higher rate of protein synthesis. Thus, apart from secondary mRNA structure and codon usage, the primary structure of the mRNA at the beginning of the hirudin gene clearly influences the expression levels.

Evidence for the influence of secondary structure on expression comes from the analysis of the data obtained with the thr codons. There is no codon of particular low usage for thr, yet ACC decreases the expression level. It is possible that, as a result of the introduction of an ACC codon, hybridisation of the hirudin mRNA with the 3′-end of 16S rRNA becomes feasible. This suggestion is compatible with the model of Stern and collaborators [9] since the incriminated nucleotides (1503–1506) are not involved in secondary RNA structures. The energy contribution [10] is calculated to be ΔG = −19 kcal. This secondary structure (see Fig. 1) could impede initiation of protein synthesis. This proposal is substantiated by the fact that pTG7202, which can form an additional A-T bond, expresses two-fold less than pTG7206. Furthermore, pTG7210 which has both ATA and ACC codons, synthesises very low amounts of hirudin. The structure shown in Fig. 1 explains the specific effect of ACC, since the other threonine codons yield sequences that lack this region of complementarity to the 3′-end of 16S rRNA.

Although the expression level of hirudin could not be increased over that already obtained with pTG720 [6],

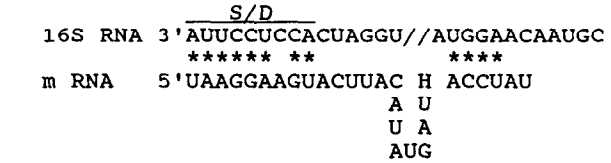

Fig. 1. Secondary structure of the 5′ extremity of the ACC containing hirudin mRNA.

various factors which determine hirudin expression in *E. coli* have been identified. Nonetheless predictions about their relative contributions remain difficult.

Acknowledgements: I am indebted to Drs J.P. Lecocq and M. Courtney for support throughout this study and to Dr M.P. Kieny and to Dr C. Roitsch for critically reading the manuscript. Many thanks go to D. Villeval for sequence analysis and to L. Stefani for technical support.

REFERENCES

[1] Grantham, R., Gautier, C., Gouy, M., Jacobzone, M. and Mercier, R. (1981) Nucleic Acids Res. 9, r43–74.
[2] Ikemura, T., (1985) Mol. Biol. Evol. 2, 13–34.
[3] Stormo, G.D., Schneider, T.D. and Gold, L. (1982) Nucleic Acids Res. 10, 2971–2996.
[4] Petersen, G.B., Stockwell, P.A. and Hill, D.F. (1988) EMBO J. 7, 3957–3962.
[5] Tessier, L.-H., Jallat, S., Sauvageot, M., Crystal, R.G. and Courtney, M. (1986) FEBS Lett. 208, 183–188.
[6] Harvey, R.P., Degryse, E., Stefani, L., Schamber, F., Cazenave, J.-P., Courtney, M., Tolstoshev, P. and Lecocq, J.-P. (1986) Proc. Natl. Acad. Sci. USA 83, 1084–1088.
[7] Riehl-Bellon, N., Carvallo, D., Acker, M., Van Dorsselaer, A., Marquet, M., Loison, G., Lemoine, Y., Brown, S.W., Courtney, M. and Roitsch, C. (1989) Biochemistry 28, 2941–2949.
[8] Varenne, S. and Lazdunski, C. (1986) J. Theor. Biol. 120, 99–110.
[9] Stern, S., Weiser, B. and Noller, H.F. (1988) J. Mol. Biol. 204, 447–481.
[10] Tinoco, I., Borer, P.N., Dengler, B., Levine, M.D., Uhlenbeck, O.L., Crothers, D.M. and Gralla, J. (1973) Nature New Biol. 246, 40–41.