In vivo processing of N-terminal methionine in *E. coli*

Henrik Dalbøge, Stephen Bayne and John Pedersen

Novo Nordisk A/S, DK-2880 Bagsvaerd, Denmark

Received 4 April 1990

The processing of amino-terminal methionine from cytosolic proteins in *E. coli* has been investigated in vivo, using amino-terminal-extended human growth hormone (hGH) as a model system. Twenty different hGH-genes with the sequence Met-Xxx-Glu-Glu-hGH where Xxx denotes each of the 20 different amino acids, were constructed and expressed in *E. coli*. Following purification of the products, the N-terminal amino acid sequences (10 cycles) were determined. The results demonstrate that the removal of methionine is dependent on the amino acid adjacent to methionine, and that the processing is strongly correlated to the radius of gyration of this amino acid. In addition, measurement of the hGH expression level from the 20 clones demonstrated that the small difference in the amino acid extension leads to a change in the specific hGH expression rate.

Processing, in vivo; Methionyl aminopeptidase; Amino-terminal-extended human growth hormone

1. INTRODUCTION

Protein synthesis in *E. coli* is under normal circumstances initiated at the translation initiation codon AUG which codes for the amino acid, methionine. As a consequence, the newly synthesized protein has as its N-terminal amino acid a methionyl residue.

E. coli possesses an enzyme activity in the cytoplasm with the capacity to remove N-terminal methionine. However, in vivo and in vitro experiments have demonstrated that the enzyme has some limitations and that processing is dependent on the amino acid adjacent to methionine [1,2].

When synthesising recombinant proteins of pharmaceutical interest, it is of the utmost importance that the final product has an N-terminal identical to the naturally occurring counterpart, as it otherwise may have antigenic properties. The correct removal of methionine is therefore of great concern.

Recently we reported a new method for the production of recombinant interleukin-1 and human growth hormone (hGH) with the correct N-termini [3,4]. The proteins are produced as precursors, N-terminally extended with a few amino acids and are subsequently converted to the correct molecules by an enzymatic process in vitro. However, the method requires that the extension has an even number of amino acids, typically 2 or 4, and we were therefore especially interested in whether the N-terminal methionine is removed in vivo or not.

Correspondence address: H. Dalbøge, Novo Nordisk A/S, DK-2880 Bagsvaerd, Denmark

Abbreviations: PVDF, polyvinylidene difluoride; PTH, phenylthiohydantoin

In the present work, an analysis of the influence of the amino acid adjacent to the methionine on the processing capacity is described.

2. MATERIALS AND METHODS

2.1. *Construction of hGH expression vectors*

The hGH expression vectors were constructed from the plasmid pHD86-3SP13 [4]. The plasmid was cut with *ClaI/Eco*RI, purified by gel electrophoresis and ligated to new synthetic linkers encoding the 20 different amino-terminal extensions (Table I). The linkers were designed to have *ClaI/Eco*RI overhangs:

```
( CGATG XXX GAA G
     TAC XXX CTT CTTAA)
```

The correct DNA sequence of the individual linkers was verified after cloning, directly on the plasmid by use of a synthetic primer complementary to a part of the promoter sequence. *E. coli* MC1061 [5] harbouring the expression plasmids was propagated in 50 ml TB medium containing 50 mg ampicillin/litre. The cultures were started from an overnight culture at an OD_{600} = ~0.2 and incubated in a shaker (250 rpm) at 30°C. Samples (1 ml) were collected in duplicate from each culture at time intervals of 2, 4, 6 and 18 h. OD_{600} and the contents of hGH were determined in each sample using a hGH-specific ELISA [6]. Following growth for 18 h (overnight samples), cells were harvested by centrifugation, resuspended in 5 ml of 10 mM Tris-HCl, pH 7.5, 50 mM EDTA, and sonicated on ice (4 × 30 s).

2.2. *Purification of amino-terminal-extended hGH*

The cell extracts were centrifuged and the resulting supernatants were dialysed for 18 h against 200 vols of 25 mM Tris-HCl, pH 7.5. The dialysed extracts were loaded on a DEAE-Sepharose column (Pharmacia) (0.8 cm^2 × 5 cm), and equilibrated with 25 mM Tris-Cl, pH 7.5. The column was washed with 10 ml of the above buffer, and the hGH eluted with 100 ml buffer containing NaCl in a gradient from 0 to 200 mM. The fractions containing hGH were pooled. 50–200 pmol of each sample were separated by SDS-PAGE [7], and electroblotted onto Immobilon PVDF membranes (Millipore) [8]. After transfer, proteins were stained with 0.1% Coomassie blue R250 in 10% acetic acid. The protein bands corresponding to amino-terminal-extended hGH were excised and stored at −20°C in an argon atmosphere until sequence analysis.

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/90/$3.50

Table I

DNA and amino acid sequence of the amino terminal-extension of the 20 different clones

Clone	Encoded N-terminus	Linker DNA sequence
pHD204	Met-Gly-Glu-Glu	ATG GGT GAA GAA
pHD107	Met-Ala-Glu-Glu	ATG GCT GAA GAA
pHD202	Met-Ser-Glu-Glu	ATG TCT GAA GAA
pHD212	Met-Tyr-Glu-Glu	ATG TAC GAA GAA
pHD211	Met-Cys-Glu-Glu	ATG TGC GAA GAA
pHD118	Met-Thr-Glu-Glu	ATG ACC GAA GAA
pHD203	Met-Pro-Glu-Glu	ATG CCG GAA GAA
pHD205	Met-Val-Glu-Glu	ATG GTT GAA GAA
pHD210	Met-Asp-Glu-Glu	ATG GAC GAA GAA
pHD207	Met-Asn-Glu-Glu	ATG AAC GAA GAA
pHD201	Met-Leu-Glu-Glu	ATG CTG GAA GAA
pHD213	Met-Ile-Glu-Glu	ATG ATC GAA GAA
pHD208	Met-Gln-Glu-Glu	ATG CAG GAA GAA
pHD116	Met-Glu-Glu-Glu	ATG GAA GAA GAA
pHD209	Met-His-Glu-Glu	ATG CAC GAA GAA
pHD214	Met-Met-Glu-Glu	ATG ATG GAA GAA
pHD115	Met-Phe-Glu-Glu	ATG TTC GAA GAA
pHD206	Met-Lys-Glu-Glu	ATG AAA GAA GAA
pHD216	Met-Trp-Glu-Glu	ATG TGG GAA GAA
pHD200	Met-Arg-Glu-Glu	ATG CGT GAA GAA

2.3. *Sequence analysis*

The PVDF membranes containing the hGH samples were subjected to automatic Edman degradation using an ABI 477A protein/peptide sequencer with on-line analysis of the released PTH-amino acids. The TFA cleavage step was modified so as to incorporate gas-phase cleavage instead of liquid phase. This enabled a more positive identification of the amino acids. All samples contained a background level of amino acids, particularly in the first two cycles of Edman degradation; but this level constituted less than 5% of the relevant N-termini.

3. RESULTS AND DISCUSSION

Twenty genes expressing amino-terminal-extended hGH were constructed. The genes were identical except for the amino acid adjacent to the Met (Table I). Following expression of the individual hGH genes, cells were harvested by centrifugation, opened by sonication and recentrifuged. The cleared cell extracts were analyzed on native and SDS-PAGE (data not shown) and purified by ion-exchange chromatography as described in section 2. The hGH contents were determined by a hGH-specific ELISA [6] in samples collected from the individual cell clones during exponential (2 and 4 h samples), early stationary phase (6 h samples) and late stationary growth phase. The expression was found to be approximately 5–10 μg/ml (OD_{600} = 1) in the 18 h samples, except for the clone pHD205 and pHD212 which both expressed approximately 1 μg/ml (OD_{600} = 1).

Although the contents of hGH per cell (except for the two clones mentioned above) were in the same order of magnitude, there was a significant difference in the specific production of hGH per cell between the individual clones. The reason(s) for this difference which was observed both in the exponential growth phase and in the early and late stationary growth phase, is not obvious, but factors such as codon usage, secondary structures or stability of either the mRNA or the product may account for the observed differences. It could also be expected that processing of the methionine could have some influence on the hGH production level. However, as no correlation between processing of the methionine and the final hGH contents is observed, this is probably not the case.

The ratio of hGH per cell between the exponential growth phase, and the late stationary phase (18 h samples), was, for most of the clones, between 2 and 4. Plasmids such as pBR322 or derivatives as used in this study, are known to increase their copy number per cell when the growth rate declines; thus the above-mentioned difference in hGH per cell between the exponentially and stationary cells may be explained by an increase in plasmid copy number.

In a recently published paper [12], 3 of the clones (pHD115, pHD116, pHD118) were compared under various controlled growth conditions in fermentors. This resulted in essentially the same observations of differences in expression levels between the clones reported here, and it is concluded that the differences are related to the small changes in the N-terminal regions.

The purified hGH samples were subjected to N-terminal analysis by automatic Edman degradation. The results are presented in Table II. The processing of the methionine residues is in accordance with the interpretation that this amino acid is only cleaved if the adjacent amino acid has a relatively small side chain. Others have shown that the radius of gyration [11] has to be 1.43 Å or less, preferably smaller than 1.29 Å in order to allow for processing of the Met in yeast [1,2,9]. However, in our experiments, processing was also found to take place if the radius of gyration was larger than 1.43 Å, as seen for Gln 1.75 Å, Ile 1.56 Å and Met 1.80 Å.

If the radius of gyration is the sole factor in determining the processing capacity of the enzyme, it would be expected that, for example, processing at Ala (0.77 Å) would be more efficient than at the Ser (1.08 Å) or Cys (1.22 Å). This does not seem to be the case as 100% processing is observed at the Met-Ser and Met-Cys, thus implying that the enzyme, although dependent on the radius of gyration of the adjacent amino acid, may have some additional preferences. The various hGH molecules differ from each other by only one amino acid, and this change may alter the solubilization properties or the charge of the protein, thus affecting the processing by making the protein less accessible to the enzyme. We have shown earlier that the processing capacity can be 'overloaded' if the expression level of hGH is high [4]. Since the expression levels of e.g. Met-Ala-Glu-Glu-hGH, Met-Thr-Glu-

Table II

Comparison of the 20 different clones

Clone	Encoded N-terminus	Gyration radius of amino acid adjacent to Met (Å) [11]	Expression level mg/ml $OD_{600}=1$ Exp./overnight	Observed N-terminus (% molecules with N-terminal Met) This report/ref. [10]
pHD204	Met-Gly-Glu-Glu	0.00	0.7/3.7	0/3
pHD107	Met-Ala-Glu-Glu	0.77	1.1/4.6	5/4
pHD202	Met-Ser-Glu-Glu	1.08	2.4/8.7	0/6
pHD211	Met-Cys-Glu-Glu	1.22	2.6/8.0	0/19
pHD118	Met-Thr-Glu-Glu	1.24	3.3/5.4	10/10
pHD203	Met-Pro-Glu-Glu	1.25	2.1/8.7	5/12
pHD205	Met-Val-Glu-Glu	1.29	0.3/1.2	10/12
pHD210	Met-Asp-Glu-Glu	1.43	1.2/5.7	100/84
pHD207	Met-Asn-Glu-Glu	1.45	1.8/6.1	100/84
pHD201	Met-Leu-Glu-Glu	1.54	2.5/6.6	100/84
pHD213	Met-Ile-Glu-Glu	1.56	3.6/4.9	90/82
pHD208	Met-Gln-Glu-Glu	1.75	1.4/4.8	40/100
pHD116	Met-Glu-Glu-Glu	1.77	0.8/2.6	100/100
pHD209	Met-His-Glu-Glu	1.78	0.5/5.4	100/100
pHD214	Met-Met-Glu-Glu	1.80	2.0/5.8	60/100
pHD115	Met-Phe-Glu-Glu	1.90	2.9/3.2	100/100
pHD206	Met-Lys-Glu-Glu	2.08	3.3/7.1	100/100
pHD212	Met-Tyr-Glu-Glu	2.13	0.2/0.7	100/100
pHD216	Met-Trp-Glu-Glu	2.21	0.7/2.2	100/100
pHD200	Met-Arg-Glu-Glu	2.38	3.0/9.1	100/100

Exp. = expression level during exponential growth phase (2 and 4 h samples)

Glu-hGH, Met-Ser-Glu-Glu-hGH and Met-Gly-Glu-Glu-hGH were of the same order of magnitude, but heterogeneous N-terminus was only observed for Met-Ala-Glu-Glu-hGH and Met-Thr-Glu-Glu-hGH, the expression level cannot explain the observed heterogeneity.

Our results on the processing of the methionine are, to a great extent, in agreement with the results published by Hirel et al. while this manuscript was in preparation [10]. In their paper, they show that the extent of methionine cleavage decreases proportionally to the increase of minimal side chain length of the amino acid adjacent to the methionine. The two sets of results are compared in Table II.

The greatest difference is seen for Gln (60% processed in our case vs 0%) and Met (40% vs 0%). The reason for this difference in the observed processing is not obvious, but the use of different model proteins, which may result in different secondary structure of the N-terminal regions and the use of different expression systems, may influence the catalytic events.

In conclusion, our results demonstrate that although dependent on the radius of gyration of the amino acid adjacent to the methionine, other factors, such as the expression level, the solubilization and charge properties of the protein and probably other as yet unknown factors, may affect the processing of the N-terminal methionine in vivo in *E. coli*.

Therefore, before choosing a final process based on in vivo processing of the methionine, it has to be considered that although processing should occur correctly, minor amounts of the product may contain N-terminal methionine. This 'contaminated' product can be difficult to separate from the correct protein in the subsequent purification steps and may also have undesirable biological side-effects.

Acknowledgements: We thank Heide Heinsøe and Jannie Rosendahl Christensen for excellent technical assistance, and Thorkild Christensen for the synthesis of DNA linkers.

REFERENCES

[1] Tsunasawa, S., Stewart, J.W. and Sherman, F. (1985) J. Biol. Chem. 260, 5382–5391.
[2] Ben-Bassat, A., Bauer, K., Chang, S.-Y., Myambo, K., Boosman, A. and Chang, S. (1987) J. Bacteriol. 169, 751–757.
[3] Dalbøge, H., Bayne, S., Christensen, T. and Hejnæs, K.R. (1989) FEBS Lett. 246, 89–93.
[4] Dalbøge, H., Dahl, H.-H.M., Pedersen, J., Hansen, J.W. and Christensen, T. (1987) Bio/Technology 5, 161–164.
[5] Casadaban, M.J. and Cohen, S.N. (1980) J. Mol. Biol. 138, 170–207.
[6] Dinesen, B. and Dalbøge, H. (1984) Anal. Chim. Acta 163, 119–125.
[7] Laemmli, U.K. (1970) Nature 277, 680–685.
[8] Matsudaira, P. (1987) J. Biol. Chem. 262, 10035–10038.
[9] Sherman, F., Stewart, J.W. and Tsunasawa, S. (1985) BioEssays 3, 27–31.
[10] Hirel, P.H.-H., Schmitter, J.-M., Dessen, P., Fayat, G. and Blanquet, S. (1989) Proc. Natl. Acad. Sci. USA 86, 8247–8251.
[11] Levitt, M. (1976) J. Mol. Biol. 104, 59–107.
[12] Jensen, E.B. and Carlsen, S. (1990) Biotechnol. Bioeng. 35, in press.