Appl Microbiol Biotechnol (1995) 42:703–708

ORIGINAL PAPER

M. Yabuta · Y. Suzuki · K. Ohsuye

High expression of a recombinant human calcitonin precursor peptide in *Escherichia coli*

Received: 11 March 1994/Received revision: 2 May 1994/Accepted: 27 May 1994

Abstract Human calcitonin (hCT) is a C-terminus α-amidated peptide hormone consisting of 32 amino acids. The amidated structure is essential for its biological activities, and the C-terminal-glycine-extended precursor peptide, hCT[G], is converted to bioactive hCT by a C-terminus-α-amidating enzyme. An efficient production method is described for the hCT[G] peptide, as a part of the fusion protein consisting of a modified *E. coli* β-galactosidase, linker amino acids and hCT[G]. Stable inclusion bodies of the fusion protein in *E. coli* were expressed by focusing on the amino acid charge, and the fusion protein was modified by inserting a basic amino acid sequence into its linker region. This modification greatly affected the formation of inclusion bodies. *E. coli* strain W3110/pG97S4DhCT [G]R4 could produce a large amount of stable inclusion bodies, and the hCT[G] peptide was released quantitatively from the fusion protein by *S. aureus* V8 protease. This enabled a large-scale production method to be established for the hCT[G] precursor peptide in *E. coli* to produce mature hCT.

Introduction

Calcitonin (CT) is a C-terminus-α-amidated peptide hormone consisting of 32 amino acids; it is produced and secreted by the parafollicular cells of the thyroid gland in mammals and by the ultimobranchial gland of fish (Wallis et al. 1985). This C-terminus-α-amidated structure is essential for its biological activities (Guttmann 1980). It has been suggested that a C-terminus-α-amidating enzyme converts the C-terminal-glycine-extended precursor peptide to mature CT during the post-translational modification steps (Bradbury et al. 1982). CT is an endogenous regulator of calcium homeostasis, acting principally on bone (Austin and Heath 1981). Chemically synthesized CT of different origins, e.g., salmon, porcine, eel and human, have been used therapeutically for various bone metabolism diseases. However, all CT except the human type (hCT) are potentially antigenic in humans, and the long-term use of salmon CT is followed by a decrease in its activity (Azria 1989). Therefore, establishing an efficient expression system for precursor hCT[G] in microorganisms and converting this precursor hCT[G] to mature hCT by a C-terminus-α-amidating enzyme would be very advantageous for industrial production and therapeutic use.

Many short peptide hormones have been expressed as a part of a fusion protein in *Escherichia coli*, and target hormones have been released from fusion proteins by chemical or enzymatic cleavage (Uhlen and Moks 1990). In such cases, highly expressed fusion proteins tend to accumulate as inclusion bodies in *E. coli*. Since inclusion bodies form an insoluble aggregate, there are advantages in using them as a starting material. For example, the isolation steps for inclusion bodies are relatively simple, and concentrated material can be easily obtained. Although the mechanism for inclusion-body formation is still not clear, factors such as the growth temperature, average charge, turn-forming residue fraction, cysteine fraction, proline fraction, hydrophilicity and total number of residues have been suggested to affect inclusion body formation (Schein and Noteborn 1988; Wilkinson and Harrison 1991). Among these factors, we focused on the amino acid charge to increase the amount of fusion protein expressed as an inclusion body in *E. coli*.

In this paper, we describe the construction of a series of expression plasmids that encode fusion proteins containing the hCT[G] precursor under the control of the lactose promoter in *E. coli*. The fusion

M. Yabuta (✉) · Y. Suzuki · K. Ohsuye
Suntory Bio-Pharma Tech Center, 2716-1 Kurakake,
Akaiwa, Chiyoda-machi, Ohra-Gun,
Gunma 370-05, Japan

proteins consisted of (a) a modified *E. coli* β-galactosidase, (b) linker amino acid residues consisting of various numbers of basic amino acids and (c) the hCT[G] precursor peptide. We studied the relationship between the expression of a fusion protein and the total amino acid charge caused by changing the amino acid residues in the linker region. As a result, it was found that the insertion of basic amino acid residues into the linker region greatly affected the formation of inclusion bodies. The *E. coli* strain harbouring one of the constructed plasmids, pG97S4DhCT[G]R4, could produce a large amount of the fusion protein as an inclusion body, and the hCT[G] precursor peptide was released from it by *Staphylococcus aureus* V8 protease with high efficiency.

Materials and methods

Materials

E. coli strain W3110 was used as a bacterial host (Bachmann 1972). Synthetic oligonucleotides were prepared with a model 380B DNA synthesizer (Applied Biosystems). DNA restriction enzymes and T4 DNA ligase were purchased from Takara-shuzo (Kyoto, Japan). L-[^{35}S]Methionine was purchased from Amersham, and *S. aureus* V8 protease (endoprotease Glu-C) was obtained from Boehringer Mannheim-Yamanouchi. Restriction enzymes were used by following the manufacturer's instructions.

Media

The SB medium consisted of glycerol (0.5%), yeast extract (2.4%), tryptone (1.2%) and 100 mM potassium phosphate buffer (pH 7.5) in deionized water. The production medium was composed of yeast extract (0.4%), K_2HPO_4 (0.4%), KH_2PO_4 (0.4%), Na_2HPO_4 (0.27%), $(NH_4)_2SO_4$ (0.12%), NH_4Cl (0.02%), L-methionine (2 mg/ml), $MgSO_4 \cdot 7H_2O$ (0.2%) and the mineral solution reported by Mori et al. (1979). Glucose or glycerol was added to the production medium as a carbon source; tetracycline (10 μg/ml) was added to maintain the plasmid.

p*I* value calculation

The isoelectric point (p*I*) of each fusion protein was calculated with a computer program (Skoog et al. 1986), the DNASIS-Mac software being obtained from Hitachi Software Engineering Co.

Pulse-chase experiment

Specific pulse-labelling of the newly synthesized fusion protein was performed by using L-[^{35}S]methionine (Cantrell et al. 1991). Seed cultures were grown overnight at 37°C in the production medium supplemented with 0.5% glycerol and 50 μg/ml each amino acid. The cells were washed by centrifugation in the same medium without methionine, cysteine and yeast extract. The washed cells were cultured again with the same medium at 37°C, and when the A_{660} value reached 0.2, L-[^{35}S]methionine was added to a final concentration of 0.74 MBq/ml. The cells were labelled for 2 min, before unlabelled L-methionine was added at 1 mg/ml. Aliquots of the culture were taken 0, 2, 5, 10 and 20 min after the unlabelled L-methionine was added and then analysed by 16% sodium dodecyl sulphate/polyacrylamide gel electrophoresis (SDS-PAGE). After the gels had been dried on paper, an autoradiograph was taken.

Fermentation experiment

Fed-batch cultivation was carried out with 30-l fermentors, 20-l production medium supplemented with 2% glucose being used as a start-up medium. During the batch fermentation period, the temperature and pH were maintained at 32°C and 7.0, respectively, and the dissolved oxygen level was controlled to 30% of the air saturation value by controlling the agitation speed. After the glucose had been consumed, the temperature was increased to 37°C, and an 80% glycerol solution was fed to promote transcription from the lactose promoter. The fed-batch cultivation was continued until cell growth had ceased.

Reversed-phase high-performance liquid chromatography (HPLC) analysis

Reversed-phase HPLC analysis was performed in a C18 column (YMC-packed A-302 type; 4.6 × 150 mm, YMC, Kyoto, Japan). Elution was carried out with a linear gradient of 10% to 38% acetonitrile in a 0.1% trifluoroacetic acid solution at a flow rate of 1.0 ml/min, the effluent being monitored by its absorbance at 214 nm.

Results

Construction of the expression plasmids

Plasmid pG97S4DhCT[G] has previously been constructed to express precursor hCT[G] as part of a fusion protein in *E. coli* by Magota K, Tanaka S (in preparation). As shown in Fig. 1, the chemically synthesized hCT[G] gene was placed after a modified *E. coli* β-galactosidase gene (encoding 97 amino acids from the N terminus) via a linker-peptide (Glu-Phe-Leu-Glu) coding gene. The glutamic acid residue preceding the first amino acid (cysteine) of hCT[G] was designed to release the hCT[G] peptide with *S. aureus* V8 protease (Houmard and Drapean 1972). When W3110/pG97S4DhCT[G] was cultured and analysed for expression of the encoding fusion protein (βGH[G]), the expression level was low (Magota K, Tanaka S, personal communication). The βGH[G] protein is an acidic protein; in other words, the calculated p*I* (pI_{cal}) value of the protein is 4.43. Since the intracellular pH value for *E. coli* is 7.5–7.9 (Slonczewski et al. 1981), it was suggested that the solubility of βGH[G] might be high in *E. coli* cells. One possible explanation for the low expression is that the high solubility of βGH[G] might have prevented the formation of insoluble inclusion bodies, and proteolytic degradation occurred. In order to improve the expression level of hCT[G] containing the fusion protein, we introduced basic amino acid residues into the linker region to increase the pI_{cal} value of the fusion protein from the acidic to neutral

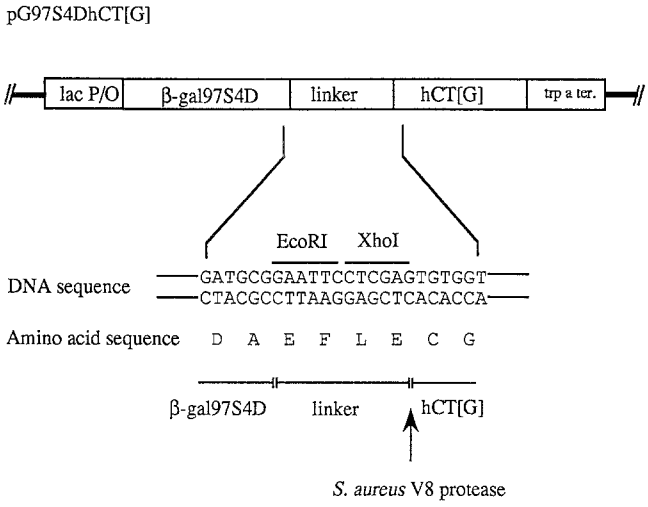

Fig. 1 Schematic drawing of the pG97S4DhCT[G] expression plasmid. Plasmid pG97S4DhCT[G] is a pBR322-derived high-copy plasmid, the βGH[G] fusion protein gene being transcribed under the control of the *E. coli* lactose promoter. The chemically synthesized hCT[G] gene encodes the C-terminal-glycine-extended precursor peptide of hCT. A chemically synthesized *trp a* terminator region (*trp a ter.*) was introduced to increase the expression efficiency. *lac P/O*, *E. coli* lactose promoter operator region; *β-gal97S4D*, a modified *E. coli* β-galactosidase (consisting of the N-terminal 97 amino acids with the glutamic acids located at 40, 41, 71 and 75 substituted with aspartic acids). The cleavage site by *S. aureus* V8 protease is indicated by an *arrow*

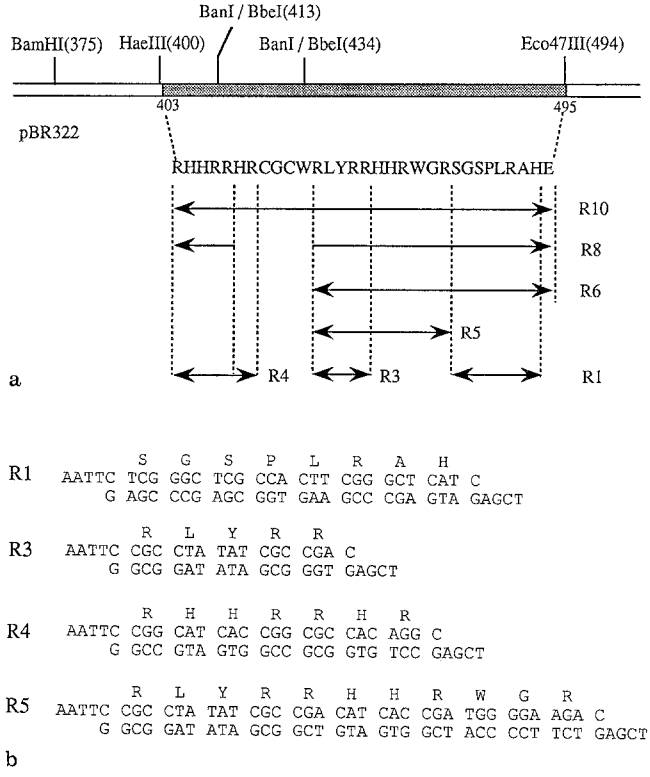

Fig. 2a, b Restriction map and linker amino acid sequences. **a** Location of the restriction-enzyme cleavage sites in the tetracycline-resistance gene of pBR322. Nucleotide 1 is the first thymidine residue in the *Eco*RI site of the pBR322. The nucleotide positions from 403 to 495 encode 31 amino acid residues. ↔ The truncated sequences (R1, R3, R4, R5, R6, R8 and R10) inserted into expression plasmid pG97S4DhCT[G]; *Rx* the number (*x*) of arginine residues in the linker amino acid sequence. **b** Sequences of the oligomers introduced into the *Eco*RI-*Xho*I site of pG97S4DhCT[G]. All the oligomers are flanked by *Eco*RI and *Xho*I sites

level. pG97S4DhCT[G] was used as the starting plasmid to construct a series of seven expression plasmids. As shown in Fig. 2a, the region of nucleotides at position 403–495 in pBR322 (the *Hae*III-*Eco*47III fragment) encodes 31 amino acid residues containing 10 arginine residues, this region being used to establish the basic amino acid sequences. A series of truncated sequences (R1–R10) from this region were introduced into the *Eco*RI-*Xho*I sites of the pG97S4DhCT[G] plasmid (Fig. 1). The methods for plasmid construction are described below.

Plasmids pG97S4DhCT[G]R1, pG97S4DhCT[G]R3, pG97S4DhCT[G]R4 and pG97S4DhCT[G]R5 were all prepared in the same manner. Chemically synthesized oligonucleotides R1, R3, R4 and R5 were ligated to the *Eco*RI- and *Xho*I-digested pG97S4DhCT[G] plasmid. The DNA sequences of these oligonucleotides are shown in Fig. 2b. The resulting plasmids were named pG97S4DhCT[G]R1, pG97S4DhCT[G]R3, pG97S4DhCT[G]R4 and pG97S4DhCT[G]R5, respectively.

In the case of pG97S4DhCT[G]R6 and pG97S4DhCT[G]R10, the 121-base-pair(bp) *Bam*HI-*Eco*47III fragment of pBR322 was isolated from a gel, further digested with *Ban*I, and blunt-ended with T4 DNA polymerase. The resulting 62-bp *Ban*I-*Eco*47III fragment was isolated and ligated with the *Eco*RI-digested, blunt-ended pG97S4DhCT[G] to obtain pG97S4DhCT[G]R6. pG97S4DhCT[G]R10 was constructed in a similar manner, except that the 95-bp *Hae*III-*Eco*47III fragment was used. Plasmid pG97S4DhCT[G]R8 was constructed by deleting the *Bbe*I(413–434) region from the pG97S4DhCT[G]R10 plasmid.

Table 1 shows the calculated p*I* value for each fusion protein and the newly constructed linker peptide sequences just described.

Expression of the fusion proteins

E. coli W3110 cells harbouring the constructed plasmids were grown at 37°C for 16 h in 100 ml SB medium contained in 1-l flasks. Since *E. coli* W3110 has only a single copy of the wild-type lactose-repressor gene, constitutive expression of the fusion proteins was observed. After harvesting, the total cellular proteins were analysed by SDS-PAGE. The molecular mass of each fusion protein was in the range from 18 kDa to 22 kDa, corresponding to the length of the linker peptide (Fig. 3). Fusion proteins ($pI_{cal} > 4.90$) containing

Table 1 Linker amino acid sequences and calculated p*I* values (pI_{cal}) for the fusion proteins. Amino acid residues in bold type were derived from *Eco*RI-*Xho*I sites. pI_{cal} values for the fusion proteins were determined by the method of Skoong and Wichman (1986) (see Materials and methods)

Plasmid	Fusion protein	Linker amino acid sequence	pI_{cal}
pG97S4DhCT[G]	βGH[G]	**EFLE**	4.43
pG97S4DhCT[G]R1	βGH[G]R1	**EFSGSPLRAHLE**	4.70
pG97S4DhCT[G]R3	βGH[G]R3	**EFRLYRRLE**	4.90
pG97S4DhCT[G]R4	βGH[G]R4	**EFRHHRRHRLE**	5.80
pG97S4DhCT[G]R5	βGH[G]R5	**EFRLYRRHHREGRLE**	5.91
pG97S4DhCT[G]R6	βGH[G]R6	**ELRLYRRHHREG-RSGSPLRAHEQFLE**	6.01
pG97S4DhCT[G]R8	βGH[G]R8	**EFRHHRRLYRRHHREGRSG-SPLRAHEQFLE**	6.83
pG97S4DhCT[G]R10	βGH[G]R10	**EFRHHRRHRCGCWRLYRR-HHREGRSGSPLRHEQFLE**	7.85

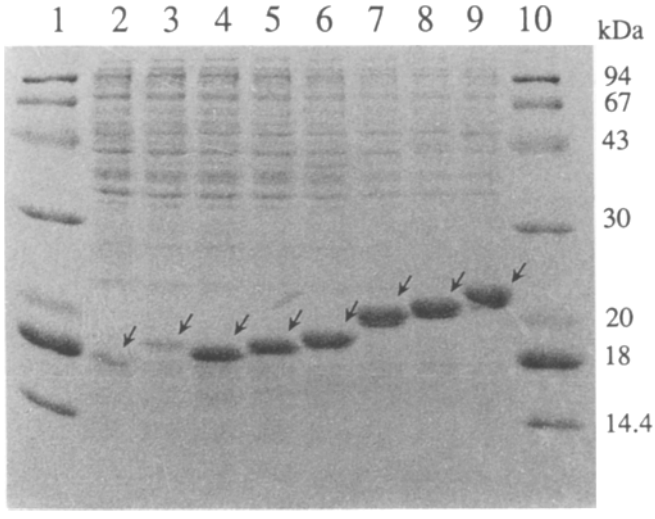

Fig. 3 SDS-PAGE analysis. After cultivation of each strain in the SB medium, the absorbance at 660 nm was measured. An equal number of cells were loaded onto 16% SDS-PAGE gels as described by Laemmli (1970). Lanes: *2* W3110/pG97S4DhCT[G]; *3* W3110/pG97S4DhCT[G]R1; *4* W3110/pG97S4DhCT[G]R3; *5* W3110/pG97S4DhCT[G]R4; *6* W3110/pG97S4DhCT[G]R5; *7* W3110/pG97S4DhCT[G]R6; *8* W3110/pG97S4DhCT[G]R8; *9* W3110/pG97S4DhCT[G]R10; *1, 10* molecular mass standards (kDa) The *arrows* indicate the position of the expressed fusion proteins

more than 3 arginine residues in the linker region were expressed to comprise over 20% of the total cellular protein (Fig. 3, lanes 4–9). These highly expressed fusion proteins formed inclusion bodies, since fusion proteins were collected by centrifugation after disruption of the cells (data not shown). On the other hand, in the case of W3110/pG97S4DhCT[G] (βGH[G], pI_{cal} = 4.43) and W3110/pG97S4DhCT[G]R1 (βGH[G]R1, pI_{cal} = 4.70), the amount of expressed fusion protein was very low (5% of total protein; Fig. 3, lanes 2 and 3). Formation of an inclusion body led to the high expression, since an inclusion body is considered to be insensitive to intracellular proteases.

Stability of the fusion proteins

The influence of post-translational protein turnover on production of the fusion protein was examined by a pulse-chase analysis with L-[^{35}S]methionine. By this analysis, the half-lives of each fusion protein were determined (Fig. 4). The half-lives were much longer (more than 20 min) for the high producers (βGH[G]R4; pI_{cal} = 5.80 and βGH[G]R6; pI_{cal} = 6.01) than that (10 min) for the low producer (βGH[G]; pI_{cal} = 4.43). In this experiment, labelled βGH[G]4 and βGH[G]R6 were predominantly observed in the insoluble fraction, even just after the chase (data not shown). To test whether these linker sequences contributed to the stabilization of transcribed mRNA, we analysed the stability of these transcribed mRNAs. The half-lives of the mRNAs were all about 2 min with the same *E. coli* strains (data not shown). Therefore, these data indicate that the stability of the fusion protein was increased by introducing the basic amino acid sequence into the fusion proteins. Rapid inclusion-body formation seems to have been a major factor for this stabilization.

Cultivation of each strain in a fermentor

The *E. coli* strains were each cultured in a 30-l fermentor as described in Materials and methods. As shown in Table 2, all the *E. coli* strains, except for W3110/pG97S4DhCT[G]R10, could be grown to high final cell density (A_{660} = 94–120). Although W3110/pG97S4DhCT[G]R10 showed the highest expression efficiency per cell with flask culture (Fig. 3, lane 9), the cell growth was approximately half that of the other *E. coli* strains. Since the culture conditions of these strains were the same, we speculate for the case of W3110/pG97S4DhCT[G]R10 that the rapid formation of inclusion bodies might have inhibited cell growth or that the βGH[G]R10 fusion protein might have been toxic to the host cell.

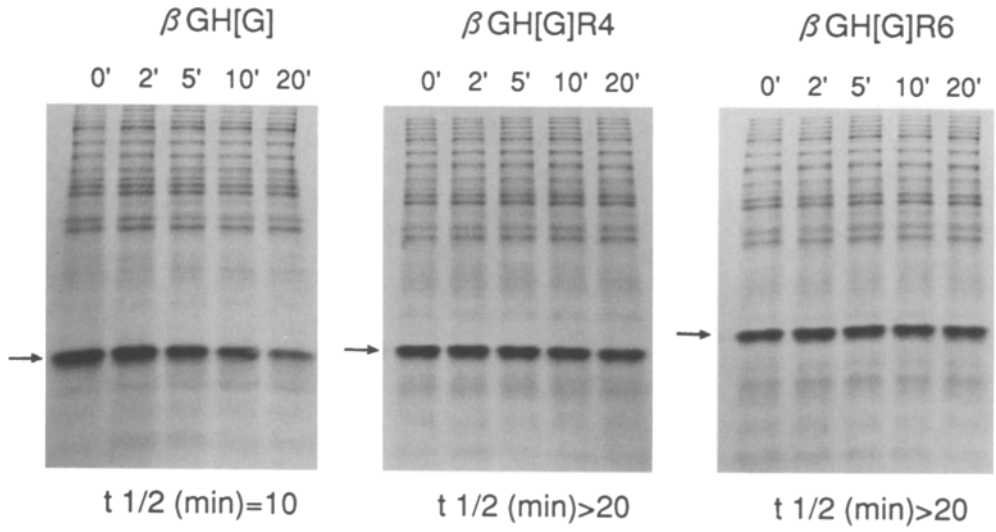

Fig. 4 Post-translational stability. *E. coli* strains W3110/pG97S4DhCT[G], W3110/pG97S4DhCT[G]R4 and W3110/pG97S4DhCT[G]R6 were labelled with L-[^{35}S]methionine, and an electrophoretic analysis of the fusion proteins was performed as described in Materials and methods. The amount of each labelled protein was determined by a densitometric scanning of the autoradiograms with a Molecular Dynamics 300A computing densitometer. These values were used to calculate the half-lives ($t_{1/2}$) of each fusion protein. Values above the lanes show the time after chasing and the *arrows* indicate the position of the labelled fusion proteins

Table 2 Effects of linker sequence on cell growth and hCT[G] release. The percentage of hCT[G] released from each fusion protein was determined by a reversed-phase HPLC analysis

Strain/plasmid	Final cell density (A_{660})	Release of hCT[G] (%)
W3110/pG97S4DhCT[G]R3	106	64
W3110/pG97S4DhCT[G]R4	94	97
W3110/pG97S4DhCT[G]R5	105	5
W3110/pG97S4DhCT[G]R6	110	34
W3110/pG97S4DhCT[G]R8	120	16
W3110/pG97S4DhCT[G]R10	53	90

Release of hCT[G] from the fusion proteins by *S. aureus* V8 protease

E. coli cells that had been cultured as just described were collected and mechanically disrupted with the Manton-Gaullin homogenizer (model 15M8TBA) at a pressure of 500 kg/cm^2. The inclusion bodies were collected by centrifugation and washed twice with deionized water. The washed inclusion bodies were then suspended in an 8 M urea solution (20 mM TRIS-HCl at pH 8.0 and 2 mM dithiothreitol) and diluted to a 2 M urea solution with deionized water. After preincubating at 30°C for 10 min, *S. aureus* V8 protease was added to the solution at an enzyme:substrate ratio of 1:5000 (w/w), which was then incubated for 60 min. The efficiency of release of hCT[G] from the fusion protein was analysed by reversed-phase HPLC. As shown in Table 2, there were notable differences in releasing efficiency among the fusion proteins. The βGH[G]R4 fusion protein was cleaved efficiently (97%), but the βGH[G]R5 fusion protein was only cleaved by 5%. Figure 5 depicts the HPLC pattern for the *S. aureus*-V8-protease-treated βGH[G]R4 fusion protein. The released hCT[G] was eluted at the same retention time (6.0 min) as that of authentic reduced hCT[G]. The hCT[G] peak was also detected in the case of the other digested fusion proteins.

Discussion

In this report we have described an attempt to form stable inclusion bodies in *E. coli* cells by changing the amino acid charge in the linker region of the hCT[G] fusion protein. The starting βGH[G] fusion protein comprised only 5% of the total protein (Fig. 3, lane 2), and the [^{35}S]methionine-labelled fusion protein was mainly detected in the soluble fraction (data not shown). On the other hand, the insertion of a sequence with more than 3 basic amino acids into the βGH[G] fusion protein promoted the formation of inclusion bodies (Fig. 3, lanes 4–9) and increased the stability of the fusion proteins (Fig. 4). It has been reported that the intracellular pH value for *E. coli* is 7.5–7.9 (Slonczewski et al. 1981). Hence, it is suggested that the change in p*I* from acidic to neutral might have decreased the intracellular solubility of the fusion protein and might have stimulated the formation of inclusion bodies. However, further studies are needed to elucidate the mechanism of inclusion-body formation caused by the insertion of basic amino acids.

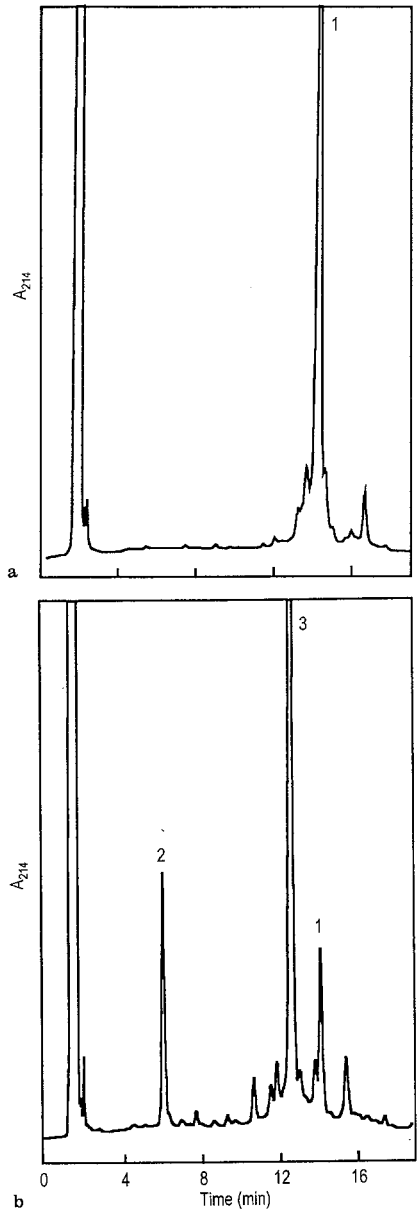

Fig. 5a, b HPLC analysis of the βGH[G]R4 fusion protein, **a** No treatment with *S. aureus* V8 protease; **b** 60 min treatment with *S. aureus* V8 protease. Peaks: *1* βGH[G]R4 fusion protein; *2* hCT[G] (reduced form); *3* partner peptide (modified β-galactosidase plus R4 linker)

The βGH[G]R4 fusion protein was cleaved efficiently (97%), whereas the βGH[G]R5 fusion protein was only cleaved by 5% (Table 2). Despite βGH[G]R6, βGH[G]R8 and βGH[G]R10 fusion proteins having the same 24-amino-acid sequence in the linker region, the releasing efficiency was quite different. The conformation of the fusion protein in the 2 M urea solution might have affected the substrate recognition of *S. aureus* V8 protease.

It was observed that modification of the linker sequence affected inclusion-body formation, cell growth and releasing efficiency by *S. aureus* V8 protease. In respect of hCT[G] productivity, we conclude that pG97S4DhCT[G]R4 was the best expression plasmid of those constructed, the yield from a 20-l culture of the strain harbouring this plasmid being as high as 10 g hCT[G].

Although not shown, the purified hCT[G] precursor was converted to C-terminus-α-amidated mature hCT by a *X. laevis* C-terminus-α-amidating enzyme, which was derived from a recombinant Chinese hamster ovary cell line (Furukawa et al. 1993). hCT thus obtained showed full biological activity (data not shown). In conclusion, we developed a high level expression system for the hCT[G] precursor peptide by changing the total amino acid charge of the fusion protein. This expression system is capable of producing mature hCT on an industrial scale.

Acknowledgements We are grateful to Prof. Dr. Toshiomi Yoshida and Dr. Masakazu Kataoka for giving us helpful advice.

References

Austin LA, Heath H (1981) Calcitonin: physiology and pathophysiology. N Engl J Med 304:269–278

Azria M (1989) Calcitonin in therapeutic use. In: The calcitonins. Karger, Switzerland, p 133

Bachmann BJ (1972) Pedigrees of some mutant strains of *Escherichia coli* K-12. Bacteriol Rev 36:525–557

Bradbury AF, Finnie MDA, Smyth DG (1982) Mechanism of C-terminal amide formation by pituitary enzymes. Nature 298:686–688

Cantrell AS, Burgett SG, Cook JA, Smith MC, Hsiung HM (1991) Effects of second-codon mutation on expression of the insulin-like growth factor-II-encoding gene in *Escherichia coli*. Gene 98:217–223

Furukawa K, Okuno K, Onai S, Sugimura K, Yoko-o Y, Ishibasi Y, Ohshima T, Tsuruoka N, Magota K, Tanaka S, Ohsuye K (1993) Production of an α-amidating enzyme (α-AE) in recombinant CHO cells. In: Animal cell technology: basic and applied aspects, vol 5. Kluwer, Netherlands, pp 493–499

Guttmann S (1980) Chemistry and structural-activity relationship of natural and synthetic calcitonins. In: Pecile A (ed) Calcitonin. Int Congr Ser 540. Excerpta Medica, Amsterdam p 11

Houmard J, Drapeau GR (1972) Staphylococcal protease: a proteolytic enzyme specific for glutamoyl bonds. Proc Natl Acad Sci USA 69:3506–3509

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227:680–685

Mori H, Yano T, Kobayashi T (1979) High density cultivation of biomass in fed-batch system with DO-stat. J Chem Eng Jpn 12:313–319

Schein CH, Noteborn MHM (1988) Production of soluble recombinant proteins in bacteria. Biotechnology 6:291–294

Skoog B, Wichman A (1986) Calculation of the isoelectric point of polypeptides from the amino acid composition. Trends Anal Chem 5:82–83

Slonczewski JL, Rosen BP, Alger JR, Macnab RM (1981) pH homeostasis in *Escherichia coli*: measurement by ^{31}P nuclear magnetic resonance of methylphosphonate and phosphonate. Proc Natl Acad Sci USA 78:6271–6275

Uhlen M, Moks T (1990) Gene fusion for purpose of expression: an introduction. Methods Enzymol 185:129–143

Wallis M, Howell SL, Taylor KW (1985) Parathyroid hormone and calcitonin. In: The biochemistry of the polypeptide hormones, chapter 13. Wiley, New York

Wilkinson DL, Harrison RG (1991) Predicting the solubility of recombinant protein in *Escherichia coli*. Biotechnology 9:443–448