Isolation and characterization of human *reg* protein produced in *Saccharomyces cerevisiae*

Takako Itoh[1], Hiroshige Tsuzuki[1], Takaaki Katoh[1], Hiroshi Teraoka[1], Kohichi Matsumoto[1], Nobuo Yoshida[1], Kimio Terazono[2], Takuo Watanabe[2], Hideto Yonekura[2], Hiroshi Yamamoto[2] and Hiroshi Okamoto[2]

[1] *Shionogi Research Laboratories, Shionogi & Co. Ltd., Fukushima-ku, Osaka 553, Osaka, Japan* and [2] *Department of Biochemistry, Tohoku University School of Medicine, Sendai 980, Miyagi, Japan*

Received 7 August 1990

reg was originally identified a a gene expressed during the regeneration of insulin-producing pancreatic β-cells of the rat. We built an expression vector containing human *reg* cDNA to drive *Saccharomyces cerevisiae* to synthesize the *reg* protein, and purified it from the culture medium. The 144-amino acid sequence of the recombinant protein was consistent with that deduced from the cDNA and genomic DNA sequence except that the signal sequence of 22 amino acids was eliminated, and the amino-terminal residue of the protein was pyroglutamic acid. The secondary structure of the *reg* protein was predicted by determination of the intramolecular cystine linkage and of α-helix and β-sheet contents.

reg protein; Secondary structure; Pancreatic stone protein; Yeast; Human pancreas

1. INTRODUCTION

In screening a cDNA library derived from rat regenerating islets [1–4], a novel gene, named *reg* (regenerating gene), was isolated [5–7]; the gene was expressed in mouse hyperplastic islets as well as in rat regenerating islets, suggesting that its expression is associated with replication, growth or maturation of pancreatic endocrine cells [5–8]. A human homologue to *reg* was also identified in a human pancreas-derived cDNA library; human and rat *reg* cDNAs encoded 166 and 165 amino acid proteins containing signal sequences [5–7]. Recently, two proteins named pancreatic stone protein (PSP) and pancreatic thread protein (PTP) were isolated from human pancreatic juice, and the amino acid sequences were determined [9,10]. Since the amino acid sequence of human *reg* protein deduced from the *reg* cDNA sequence [5–7] contains the sequences of PSP and of PTP and the human *reg* gene is a single copy gene [11], it is concluded that the *reg* protein, PSP, and PTP are just different names for a single protein existing in several molecular forms but deriving from the *reg* gene [11].

Correspondence address: H. Okamoto, Department of Biochemistry, Tohoku University School of Medicine, 2-1 Seiryomachi, Aoba-ku, Sendai 980, Miyagi, Japan

Abbreviations: PSP, pancreatic stone protein; PTP, pancreatic thread protein; TPCK, *N*-tosyl-1-phenylalanyl chloromethyl ketone; PTH, phenylthiohydantoin

In this study, we constructed an expression plasmid to direct the synthesis of human *reg* protein in *S. cerevisiae* and characterized the primary and secondary structures of the purified recombinant *reg* protein.

2. EXPERIMENTAL

2.1. *Materials*

A yeast-*Escherichia coli* shuttle vector pAM82 carrying the PHO5 promoter of the acid phosphatase gene [12] was a kind gift from Dr Kenichi Matsubara (Osaka University). Restriction endonucleases were purchased from Takara Shuzo Co. (Kyoto). S-Sepharose was from Pharmacia. Pyroglutamyl aminopeptidase from Boehringer Mannheim. Carboxypeptidases A and Y from Oriental Yeast Co. (Tokyo). Lysyl endopeptidase from Wako Pure Chemical Industries (Osaka). TPCK-treated trypsin from Sigma. Mouse anti-human *reg* protein monoclonal antibody was prepared in our laboratory using recombinant human *reg* protein [8,11]. Pancreas extract was prepared as described [11].

2.2. *Construction of expression plasmid for human* reg *protein*

Human *reg* cDNA in λgt10 [5] was digested with *Eco*RI and subcloned in pBS vector. The *Afl*II/*Eco*RI fragment containing the 5′-terminal half of the human *reg* cDNA and the *Eco*RI/*Xba*I fragment containing the 3′-terminal half were excised from the plasmids, and the *Afl*II site (24 bp upstream of translational initiation codon) and the *Xba*I site (24 bp downstream of the stop codon) were converted into *Xho*I and *Hin*cII sites by utilizing *Xho*I and *Sal*I linkers, respectively, thus forming a *Xho*I/*Hin*cII fragment, which was inserted into pAM82 previously digested with *Xho*I and *Pvu*II. The resultant recombinant plasmid (pAM82-human *reg*), shown in Fig. 1, thus provides all the appropriate transcriptional and translational signals for producing human *reg* protein in yeast.

2.3. *Production and purification of recombinant protein*

S. cerevisiae AH22 cells were transformed with the recombinant

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

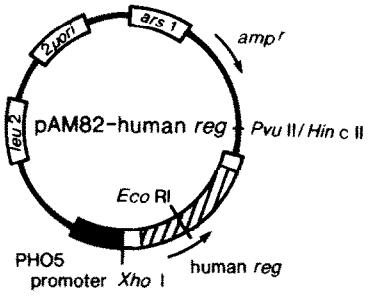

Fig. 1. Schematic representation of the recombinant plasmid used for the expression of human *reg* protein. The hatched box in the plasmid represents the entire protein coding region in human *reg* cDNA. Open boxes represent 5′ and 3′ untranslated regions. The closed box represents the promoter region of PHO5 in a shuttle vector (pAM82) with a replication origin (2 μori) and selection marker (leu2) that function in yeast. The direction of transcription of the human *reg* cDNA and the ampicillin resistant gene (*amp*^{r}) is shown by arrows.

plasmid by the lithium acetate method [13], and grown in Burkholder medium [14] containing histidine (20 mg/l) at 30°C with aeration. Production of recombinant proteins was induced by eliminating P_i. S-Sepharose equilibrated with 0.05 M acetate buffer (pH 3.4) was added to the culture medium that was adjusted to the same pH, followed by overnight agitation at 4°C, then packed into a column (4.4 × 50 cm), and washed, after which the proteins were eluted with 0.5 M NaCl in the buffer. The eluates were dialyzed, reapplied to another S-Sepharose column (1.5 × 43 cm), and eluted with a linear gradient of 0.2 M to 0.5 M NaCl in the buffer. The fractions containing proteins that displayed a single band at 16 kDa on SDS-PAGE were collected, concentrated and desalted by ultrafiltration with an Amicon YM-10 membrane.

2.4. *SDS-PAGE and Western blot analysis*

Samples were subjected to SDS-PAGE [15], transferred to a nitrocellulose filter [16] and immunostained with an anti-human *reg* monoclonal antibody as described [11].

2.5. *Determination of amino acid composition and sequence*

The purified recombinant protein and peptide fragments were hydrolyzed in 4 M methanesulfonic acid containing 0.2% 3-(2-aminoethyl)indol at 110°C for 24 h, and analyzed by a Hitachi model 835 amino acid analyzer. Amino acid sequence was determined by automated Edman degradation using a protein sequencer 477A and a PTH analyzer 120A (Applied Biosystems Inc.).

3. RESULTS AND DISCUSSION

Yeast cells carrying pAM82-human *reg* secreted predominantly a 16 kDa protein (Fig. 2A, lane 2, upper arrow), which reacted strongly with an antibody to human *reg* protein (Fig. 2B, lane 2). The 16 kDa protein migrated on SDS-PAGE to almost the same position as the smaller molecular weight forms of pancreatic *reg* protein (Fig. 2B, lane 4). The 16 kDa recombinant protein purified by two cycles of S-Sepharose chromatography showed a single band on SDS-PAGE (Fig. 2A and B, lane 3). About 1 mg of the purified *reg* protein was obtained per liter of the culture medium.

No PTH-amino acid was liberated at the first cycle of Edman degradation of the recombinant protein, but after the protein was digested with pyroglutamyl aminopeptidase the degradation reaction efficiently proceeded over 30 cycles, indicating that the amino-terminal residue is a pyroglutamic acid. As shown in Fig. 3, the amino terminal sequence was in agreement with positions 1–30 of the sequence deduced from human *reg* cDNA and gene [5,11], indicating that the 22-amino acid signal sequence had been eliminated in the 16 kDa protein. Carboxypeptidase A digestion of the *S*-carboxymethylated [17] protein released an asparagine residue, and subsequent carboxypeptidase Y digestion permitted confirmation of the 7 carboxyl-terminal residues, Phe(138)-Val-Cys-Lys-Phe-Lys-Asn(144)-COOH for the deduced human *reg* protein (Fig. 3). The amino acid composition of the 16 kDa protein was in good agreement with that deduced from the human *reg* cDNA and gene [5,11] (data not shown).

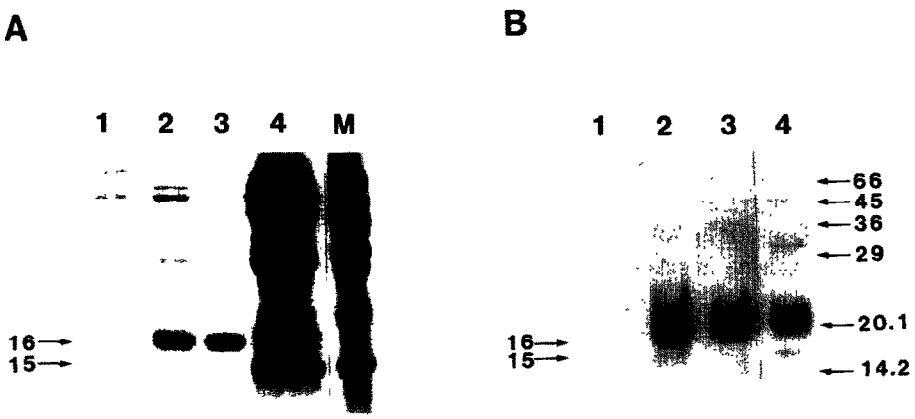

Fig. 2. SDS-PAGE and immunoblotting analyses of recombinant and pancreatic *reg* proteins. Lane 1, condensed culture medium of untransformed AH22 cells; lane 2, condensed culture medium of pAM82-human *reg*-transformed AH22 cells; lane 3, purified recombinant human *reg* protein (1 μg); lane 4, human pancreas extract (30 μg as protein); lane M, protein molecular mass markers (66, 45, 36, 29, 20.1, 14.2 kDa). Culture medium was condensed about 100-fold, after which an aliquot (5 μl) was loaded on each lane. Protein bands were visualized with Coomassie blue staining (A) or blotted onto nitrocellulose filters (B). Arrows indicate *reg* proteins. Molecular sizes (in kDa) as estimated from marker proteins are shown on the left.

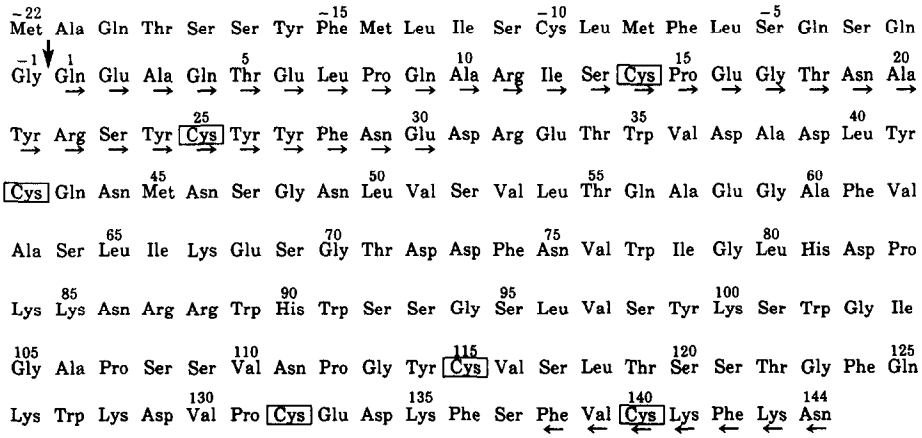

Fig. 3. Amino acid sequence of human *reg* protein. The sequence is 166 amino acids long [5] and numbered every 5 amino acids. The signal peptide is indicated by minus numbers. The vertical arrow shows the position at which the mature protein begins. The horizontal arrows indicate the residues determined by amino acid sequencing. The 6 cysteine residues involved in disulfide bond formation are boxed.

To determine the disulfide linkage location between the 6 cysteine residues in the recombinant *reg* protein, the protein was digested with lysyl endopeptidase, and a part of the digest was *S*-carboxymethylated. As shown in Fig. 4A, when the HPLC elution profiles of non-*S*-carboxymethylated and *S*-carboxymethylated digests were compared, three peaks, 2, 4 and 5, in the non-*S*-carboxymethylated sample were seen to have disappeared as a result of *S*-carboxymethylation. By amino acid composition and sequence analyses, peak 4 was revealed to contain undigested *reg* protein: peak 2 contained two peptides, Asp-129–Lys-135 and Ser-101–Lys-126, indicating the interconnection of Cys-115 and Cys-132 by a disulfide bond (see Fig. 3); peak 5 was composed of two peptides containing four half-cystine residues. The peak 5 peptides were then subjected to tryptic digestion, followed by separation on HPLC, which gave three peaks, 5TA, 5TB, and 5TC (Fig. 4B). 5TA consisted of the single polypeptide chain Glp-1–Arg-11 with no cysteine residue, 5TB consisted of two peptides, Ile-12–Arg-22 and Ser-23–Arg-32, and 5TC consisted of two peptides, Glu-33–Lys-67 and Phe-136–Lys-141, indicating that the position of disulfide bridges in the peak 5 peptides

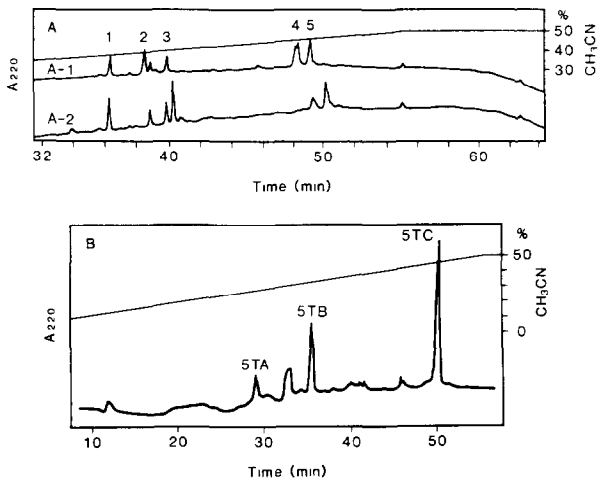

Fig. 4. Separation of disulfide bond- or *S*-carboxymethylcysteine-containing peptides of the protease-digested recombinant human *reg* protein by reverse phase HPLC. The peptides were separated on a Vydac $5C_4$ column (4.6 × 150 mm) with 50-min linear gradient from 0% to 50% acetonitrile in 0.1% trifluoroacetic acid at a flow rate of 1.0 ml/min. (A) Lysyl endopeptidase digested peptides: (A-1) non-carboxymethylated peptides; (A-2) carboxymethylated peptides. (B) Non-carboxymethylated tryptic peptides derived from the peptide in peak 5 (Fig. 4A).

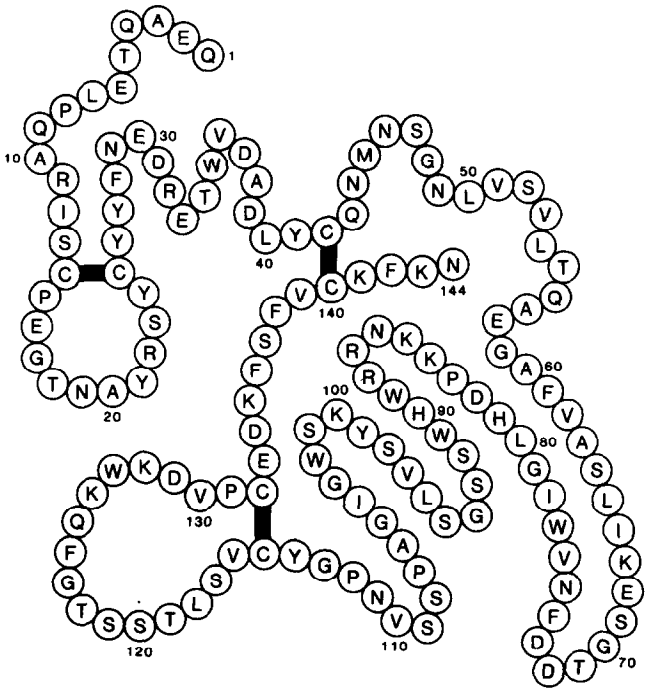

Fig. 5. Schematic representation of the secondary structure of recombinant human *reg* protein. The disulfide bonds are shown as the vertical or horizontal bars joining cysteine residues. The single-letter code for amino acids is used.

are Cys-14–Cys-25 and Cys-42–Cys-140. The α-helix and β-sheet contents of the recombinant human *reg* protein estimated by circular dichroism and by the Chou and Fasman algorithm [18] were 12–17% and 22–26%, respectively (data not shown). Thus, the secondary structure of the *reg* protein is predicted as shown in Fig. 5.

The recombinant *reg* protein exhibited the following similarities to the naturally occurring human *reg* protein (PSP/PTP). (i) The amino acid sequence of the recombinant protein was consistent with those of PSP [9,19] and PTP [10]. (ii) The recombinant protein had the same cystine linkages (Fig. 5) as PSP [20]. (iii) The amino-terminal glutamine residue of the recombinant protein was converted to pyroglutamic acid residue as in the case of PSP [9]. (iv) The recombinant protein showed almost the same electrophoretic mobility as the smaller molecular weight form of the pancreatic protein (Fig. 2). (v) The recombinant protein showed immunoreactivity to an anti-*reg* protein antibody as did the pancreatic protein.

PSP has been shown to have *O*-linked oligosaccharides [9] at Thr5, which accounts for the presence of larger molecular forms (Fig. 2B, lane 4). However, the recombinant human protein showed a single band at 16 kDa on SDS-PAGE (Fig. 2B, lane 3), which corresponded to the smaller molecular form of the pancreatic protein. This suggests that the recombinant human *reg* protein synthesized in yeast is either not or only a little *O*-glycosylated. In addition, minor 15 kDa proteins detected by SDS-PAGE (Fig. 2A, lane 2) and Western blotting (Fig. 2B, lane 2) began with Ile12 (Itoh, T., unpublished results), indicating that the 15 kDa protein was generated by proteolytic cleavage between Arg-11 and Ile-12 (see Fig. 3). The 15 kDa protein of 133 amino acids is thought to correspond to PSP S1, which has been isolated from human pancreatic juice [19]. We have also obtained recombinant rat *reg* protein in a similar way (unpublished results). These recombinant proteins should be useful for studying the biological functions of the *reg* protein or PSP/PTP both in vivo and in vitro.

REFERENCES

[1] Yonemura, Y., Takashima, T., Miwa, K., Miyazaki, I., Yamamoto, H. and Okamoto, H. (1984) Diabetes 33, 401–404.
[2] Okamoto, H. (1985) BioEssays 2, 15–21.
[3] Okamoto, H., Yamamoto, H. and Yonemura, Y. (1985) in: ADP-Ribosylation of Proteins (Althaus, F.R., Hilz, H. and Shall, S. eds) pp. 410–416, Springer-Verlag, Berlin.
[4] Yonemura, Y., Takashima, T., Matsuda, Y., Miwa, K., Sugiyama, K., Miyazaki, I., Yamamoto, H. and Okamoto, H. (1988) Int. J. Pancreatol. 3, 73–82.
[5] Terazono, K., Yamamoto, H., Takasawa, S., Shiga, K., Yonemura, Y., Tochino, Y. and Okamoto, H. (1988) J. Biol. Chem. 263, 2111–2114.
[6] Okamoto, H., Yamamoto, H., Takasawa, S., Inoue, C., Terazono, K., Shiga, K. and Kitagawa, M. (1988) in: Frontiers in Diabetes Research. Lessons from Animal Diabetes II (Shafrir, E. and Renold, A.E. eds) pp. 149–157, John Libbey and Company Ltd., London, Paris.
[7] Terazono, K., Watanabe, T. and Yonemura, Y. (1990) in: Molecular Biology of the Islets of Langerhans (Okamoto, H. ed.) pp. 301–313, Cambridge University Press, Cambridge.
[8] Terazono, K., Uchiyama, Y., Ide, M., Watanabe, T., Yonekura, H., Yamamoto, H. and Okamoto, H. (1990) Diabetologia 33, 250–252.
[9] De Caro, A.M., Adrich, Z., Fournet, B., Capon, C., Bonicel, J.J., De Caro, J.D. and Rovery, M. (1989) Biochim. Biophys. Acta 994, 281–284.
[10] Gross, J., Carlson, R.I., Brauer, A.W., Margolies, M.N., Warshaw, A.L. and Wands, J.R. (1985) J. Clin. Invest. 76, 2115–2126.
[11] Watanabe, T., Yonekura, H., Terazono, K., Yamamoto, H. and Okamoto, H. (1990) J. Biol. Chem. 265, 7432–7439.
[12] Miyanohara, A., Toh-e, A., Nozaki, C., Hamada, F., Ohtomo, N. and Matsubara, K. (1983) Proc. Natl. Acad. Sci. USA 80, 1–5.
[13] Itoh, H., Fukuda, Y., Murata, K. and Kimura, A. (1983) J. Bacteriol. 153, 163–168.
[14] Bostian, K.A., Lemire, J.M., Cannon, L.E. and Halvorson, H.O. (1980) Proc. Natl. Acad. Sci. USA 77, 4504–4508.
[15] Laemmli, U.K. (1970) Nature 227, 680–685.
[16] Burnette, W.N. (1981) Anal. Biochem. 112, 195–203.
[17] Crestfield, A.M., Moore, S. and Stein, W.H. (1963) J. Biol. Chem. 238, 622–627.
[18] Chou, P.Y. and Fasman, G.D. (1978) Adv. Enzymol. 47, 45–148.
[19] De Caro, A.M., Bonicel, J.J., Rouimi, P., De Caro, J.D., Sarles, H. and Rovery, M. (1987) Eur. J. Biochem. 168, 201–207.
[20] Rouimi, P., De Caro, J., Bonicel, J., Rovery, M. and De Caro, A. (1988) FEBS Lett. 229, 171–174.