Enhanced expression of a furin-cleavable proinsulin

C W Hay and K Docherty

Department of Molecular and Cell Biology, University of Aberdeen, Institute for Medical Sciences, Foresterhill, Aberdeen AB25 2ZD, UK

(Requests for offprints should be addressed to K Docherty, Department of Molecular and Cell Biology, University of Aberdeen, Institute of Medical Sciences, Foresterhill, Aberdeen AB25 2ZD, UK; Email: k.docherty@abdn.ac.uk)

Abstract

Cell engineering or gene therapy may represent an alternative to current methods of treating diabetes mellitus. Cells could be engineered to secrete insulin *ex vivo* for transplantation or the insulin gene could be administered directly by injection into muscle. A problem has been that non-neuroendocrine cells lack the endoproteases (PC3/1 and PC2) that are responsible for the processing of proinsulin to insulin. This can be surmounted by engineering the paired basic amino acid processing sites within proinsulin to sites that would be recognized by the ubiquitously expressed protease, furin. However, in every study to date, the expression of the furin-cleavable construct was greatly reduced relative to that of the unmodified proinsulin construct. We investigated possible causes for this, including mRNA stability, the presence of additional CpG islands, and the amino acid substitutions within furin-cleavable proinsulin. Several furin-cleavable rat proinsulin I cDNAs were engineered and used to transfect human HEK293, rat L6 and mouse C₂C₁₂ cell lines. The stability of wild-type and furin-cleavable proinsulin mRNA in transfected C₂C₁₂ cells was measured by RT-PCR. Comparison of the decay rates in the presence of actinomycin D showed no significant difference between the two species of mRNA. A furin-cleavable proinsulin cDNA was created to contain the same distribution of CpG islands as wild-type proinsulin. Comparison of insulin-like immunoreactivity in all three cell lines transfected with either this construct or a widely used furin-cleavable proinsulin containing additional CpG islands showed that the presence of the extra CpG islands had no effect. Studies to examine amino acid substitutions used to create furin consensus sequences showed that the addition of basic residues at the C-peptide/A-chain junction was responsible for the reduced production of furin-cleavable proinsulin. Using this information, we engineered a cDNA for furin-cleavable rat proinsulin I that was efficiently processed to mature insulin and expressed at the same level as wild-type proinsulin.

Journal of Molecular Endocrinology (2003) **31**, 597–607

Introduction

The intracellular conversion of all known mammalian proinsulins to insulin involves cleavage at two paired basic sites present at either end of the C-peptide (Steiner 1998). In both rat and human proinsulins, these are Arg³¹-Arg³² at the B-chain/C-peptide junction and Lys⁶⁴-Arg⁶⁵ at the C-peptide/A-chain junction. The cleavage is carried out by the endopeptidases PC2 and PC3/1 (Seidah *et al.* 1990, Smeekens & Steiner 1990, Smeekens *et al.* 1991), which are restricted to pancreatic islets and several specialised endocrine cells. The two basic residues at the newly created COOH termini of the B-chain and C-peptide are then removed by carboxypeptidase E/H (Fricker

et al. 1982, Fricker 1988) to yield fully active insulin, in which the B- and A-chains are joined by two disulphide bridges.

Autoimmune destruction of the pancreas β -cells leads to insulin deficiency (Bach 1994) and type I (insulin-dependent) diabetes. Type II diabetes has several causes, but insulin therapy is required in up to 50% of cases (Berger 1996). Unfortunately, the tight glycaemic control required to prevent long-term microvascular complications in both types of diabetes is difficult to attain using subcutaneous infusion of insulin. Alternative approaches have been pursued; however, a lack of suitable donors and fears of transmission of agents of infection (Weiss 1998) have limited the application of pancreas transplantation and

xenografts respectively. Gene therapy for diabetes has, therefore, become the focus of much research (Docherty 1997, Halban *et al.* 2001).

Most cell engineering approaches target non-neuroendocrine cells that lack the specific endopeptidases (PC2 and PC3/1) required to process proinsulin into active mature insulin. To overcome this obstacle, various groups have used a variety of site-directed mutations to engineer proinsulin to be a substrate for furin (Yanagita *et al.* 1992, Groskreutz *et al.* 1994, Muzzin *et al.* 1997, Short *et al.* 1998, Yamasaki *et al.* 1999, Shaw *et al.* 2002). This enzyme (also known as PACE) is a Golgi-associated propeptide endoprotease that is present in the constitutive secretory pathway of virtually all cells (Van de Ven *et al.* 1990). The introduction of furin consensus sequences at the B-chain/C-peptide and C-peptide/A-chain junctions has been shown to increase significantly the processing of proinsulin to insulin in a wide variety of non-neuroendocrine cells, including fibroblasts, myoblasts, epithelial cells and lymphocytes. Despite the efficient conversion of mutated furin-cleavable proinsulin into insulin (Nishigori *et al.* 1996), the efficacy of all gene therapy strategies for diabetes has been severely hampered by the greatly reduced expression of all furin-cleavable proinsulins in comparison with the wild-type protein.

In the experiments reported here, we investigated various factors that could account for this reduction by creating site-directed mutations in rat proinsulin I and expressing the protein in human, rat and mouse cell lines. The results clearly show that production of furin-cleavable proinsulin is sensitive to the amino acid substitutions used. Consequently, we were able to create a furin-cleavable rat proinsulin that was expressed at the same levels as wild-type protein and processed with 77% efficiency in HEK293 cells.

Materials and methods

Cell culture

Human embryonic kidney cells (HEK293) were cultured in low glucose (1 mg/ml) DMEM supplemented with 10% foetal calf serum (both from Gibco, Paisley, UK). The mouse muscle cell line C₂C₁₂ (obtained from the European Collection of Animal and Cell Cultures, Salisbury, UK) and the rat muscle cell line L6 were grown in DMEM

with the addition of 10% foetal calf serum and 4 mM glutamine. All cell lines were maintained at 37 °C under 5% CO₂ in humidified air without the addition of antibiotics.

Expression plasmids

Wild-type rat preproinsulin I cDNA and rat preproinsulin I cDNA modified at the B-chain/C-peptide and C-peptide/A-chain junctions for cleavage by furin (Yanagita *et al.* 1992, 1993) were provided by Professor T. Takeuchi (Institute for Endocrinology, Gunma University, Japan). The cDNAs had been cloned into the mammalian expression plasmid pCXN2 (Niwa *et al.* 1991), in which expression of preproinsulin is driven by the chicken β -actin promoter and CMVIE enhancer. The proinsulin coding region followed the chicken cytoplasmic β -actin first intron and the rabbit β -1 globin type 1 allele intron for enhanced mRNA stability. The region downstream of the cDNA inserts contained the rabbit β -1 globin 3' flanking sequence including the polyadenylation signal. The plasmid also contained genes for selection by ampicillin and neomycin resistance.

In situ mutagenesis

The native rat preproinsulin I cDNA was excised from the pCXN2 plasmid by digestion with EcoRI and subcloned into pUC18. Mutagenesis was carried out using the QuikChange Multi Site-directed Mutagenesis kit (Stratagene, Amsterdam, Netherlands) with 5' phosphorylated oligonucleotides according to the manufacturer's methods. The furin-cleavable consensus sequence at the B-chain/C-peptide junction in pCXN2rIns3 and pCXN2rIns6 was created using the oligonucleotide 5'-CCCAAGTCCCGTCGTAAAAGGCAGGACCCGCAAGTG-3'. Similarly, the furin-cleavable consensus sequence at the C-peptide/A-chain junction in pCXN2rIns3 and pCXN2rIns7 was made with the oligonucleotide 5'-CTGGAGGTTGCCCGGAGGAAGCGTCGCATTGTG-3'.

The native rat preproinsulin I cDNA subcloned into pUC18 was also used to create wild-type and furin-cleavable protein sequences (pCXN2rIns4 and pCXN2rIns5 respectively) that were completely lacking CpG islands, by performing several

rounds of *in situ* mutagenesis. The same methodology as above was used, with the product of each round of mutagenesis being used as the substrate for the subsequent round after analysis by restriction enzyme digestion or DNA sequencing to confirm that the desired mutations were present. Firstly, pCXN2rIns4 was created using the following oligonucleotides: first round 5'-GCCCTGTG GATGAGATTCCTGCCCCCTG-3', 5'-CTGGC CCTGCTTGCTCCTCTGGGAG3', 5'-AAGTGG AGGACCCTCAAGTGCCACAAC-3', 5'-GCCC GGCAGAAGAGAGGCATTGTGGATC-3' and 5'-GAGCTCGAATTCAAGCTCGCCCCG-3'; second round 5'-GGGTGGAGGCCCTGAGGC TGGGGATCTTCAG-3', 5'-ACTGGAGGTTG CCAGGCAGAAGAGAGGC-3', 5'-CTCGAAT TCAAGCTCACAGGGGATCCAACATGG-3' and 5'-TGCAACTGAGTCCAGGGATCCTCT AGAGATC-3'; third round 5'-TACACACCCA AGTCCAGGAGAGAAGTGGAGGACCCTC-3'; fourth round 5'-TCTAGAGATCTCTAGAGTT GAGGAATTCGTAATCATG-3'; and fifth round 5'-CCTGGTGTGTGGGGAAAGAGTTTCTT CTACACACC-3'. Finally, furin cleavage sites at both cleavage sites in pCXN2rIns5 were created using the two oligonucleotides 5'-CCCAAGTCC AGGAGAAAAAGGGAGGACCCTCAAGTG-3' and 5'-CTGGAGGTTGCCAGGAGGAAGAGA GGCATTGTG-3', which changed the amino acid sequences at the B:C and C:A splice sites respectively.

The mutated rat preproinsulin I cDNAs were subcloned back into the mammalian expression plasmid vector pCXN2. The mutations and integrity of final products were confirmed by DNA sequencing.

DNA transfection

Plasmids used for transfection studies were prepared using Qiagen (Crawley, West Sussex, UK) endotoxin-free plasmid maxiprep kits in accordance with the manufacturer's instructions. Plasmid purity and concentration were determined spectrophotometrically by measuring absorbance at 260 nm and 280 nm, and by quantitative agarose gel electrophoresis. Cells were allowed to grow to near confluence in six-well tissue culture plates. One microgram of plasmid DNA and 10 µl Lipofectamine (Invitrogen, Paisley, UK) were used per well. The plasmid and Lipofectamine were

incubated for 30 min before transfection in accordance with the manufacturer's procedure. The cells were washed with phosphate buffered saline and, upon transfection, were incubated in a final volume of 1 ml serum-free medium for 5 h. After this period an equal volume of medium containing 20% foetal calf serum was added and the cells were incubated for a further 18 h. The medium was then replaced and the cells were cultured for 24 h, after which time the culture medium was collected.

mRNA stability studies

C₂C₁₂ cells were transfected with either pCXN2rIns1 expressing wild-type proinsulin or pCXN2rIns2 expressing furin-cleavable proinsulin. Twenty-four hours after transfection, actinomycin D (Melford Laboratories, Ipswich, UK) was used at 5 µg/ml to block transcription and the cells were harvested at different time points (6, 12, 24 and 36 h). Total RNA was prepared from cell pellets using Trizol (Invitrogen, Paisley, UK) and the concentration determined by spectroscopy. One microgram aliquots of total RNA were digested with 1 unit DNaseI (Invitrogen Amplification Grade) for 15 min at room temperature and hybridised with 5 ng oligo-(dT)₁₂₋₁₈. First-strand synthesis was carried out using 200 units SuperScript II reverse transcriptase (Invitrogen) in the presence of 1 µl RNasin (Promega, Southampton, UK) and 0.5 mM deoxynucleotide triphosphates in a volume of 20 µl. PCR amplification was performed using 1 µl of the reverse transcriptase reaction in a total volume of 25 µl containing 0.2 µM each specific primers, 3 mM MgCl₂, 0.2 mM deoxynucleotide triphosphates (equimolar amounts) and 1.25 units Taq polymerase (Promega). Thermal cycling was carried out using an initial denaturing step at 94 °C for 2 min and thereafter 30 s each at 94 °C, 66.5 °C and 72 °C (27 cycles) with a final step of 4 min at 72 °C for rat preproinsulin I. PCR amplification for mouse GADH was identical, except that an annealing temperature of 68.0 °C was used, with 21 cycles. In order to avoid possible contamination from genomic or plasmid DNA, the following primers were used. Rat preproinsulin I mRNA was amplified with a forward primer that spanned the intron splice site in the chicken β-actin/rabbit β-1 globin type 1 allele sequence upstream of the

cDNA insert (5'-CGTTACTCCCACAGCTCCT-3'). The reverse primer was specific for the 3' end of the coding region of rat insulin (5'-TTGCAGTAGTTCTCCAGTTGGT-3'). Similarly, the mouse GADH mRNA was amplified with a forward primer, which spanned an intron splice site (5'-AAGGGCTCATGACCACAGTC-3'). The reverse primer was complementary to the 3' region of coding sequence, with the amplified region covering two introns in the genomic sequence (5'-CTTACTCCTTGGAGGCCATGT-3').

The products of RT-PCR were analysed by 2.0% agarose Tris acetate EDTA gel electrophoresis, followed by staining with ethidium bromide. Gels were analysed using the GelDoc system and DNA bands quantified against standard amounts of DNA using MultiAnalyst software (Bio-Rad). Care was taken to avoid pixel saturation by using appropriate camera aperture and image attenuation.

Gel exclusion chromatography

Ten grams of Sephadex G-50 Superfine beads were allowed to swell overnight in column buffer (50 mM phosphate buffer pH 6.8, 0.5% (w/v) BSA and 0.02% sodium azide). After degassing, the beads were packed into a 1 cm × 120 cm Bio-Rad chromatography column. Compounds greater than 30 kDa are excluded from this column, with resolution of proteins with molecular weights of between 1500 and 30 000. Samples (0.4–0.5 ml) were prepared from the medium of transfected cells by centrifugation at 1000 *g* for 3 min. The supernatant was either applied directly to the column or acetic acid was added to a final concentration of 20 mM and stored at –70 °C. The supernatants of media containing low amounts of (pro)insulin were concentrated using Millipore Centricon YM-3 centrifugation filters (Millipore, Watford, UK). The filters have a molecular weight cutoff of 3000 and were used in accordance with the manufacturer's instructions, using a Sorvall SS34 rotor. Column fractions of 1 ml were collected and insulin-like immunoreactivity (ILI) determined by radioimmunoassay.

Radioimmunoassays

Total ILI in culture medium and gel exclusion chromatography column fractions was detected with an in-house radioimmunoassay using guinea

pig polyclonal antibodies raised against human (pro)insulin (1011 from Linco Research, Biogenesis, Poole, UK). Radioimmunoassays were calibrated using rat insulin (Linco Research) and performed in duplicate. ILI results were calculated as ng/ml per 24 h.

Statistical analysis

Values were calculated as means ± S.E.M. Two-way ANOVA was used to compare the time-courses of the mRNA half-lives.

Results

Gel filtration analysis of insulin production by cultured cell lines

In order to investigate expression and processing of furin-cleavable proinsulin, use was made of three cell lines from different species: human embryonic kidney cell line HEK293, mouse myoblast cell line C₂C₁₂ and rat myoblast cell line L6. Each of these cell lines was transfected with either pCXN2rIns1 expression plasmid encoding wild-type rat proinsulin I or pCXN2rIns2 expressing rat proinsulin I that had been mutated to permit specific cleavage by furin. These are described in Table 1, in which changes from the wild-type sequence are underlined. Culture medium was collected after 24 h and subjected to Sephadex G-50 gel filtration. Culture medium from HEK293 cells expressing wild-type proinsulin was resolved into two immunoreactive peaks, with 85% eluting in the position of proinsulin and only 15% in the position of mature insulin (Fig. 1A). In contrast, culture medium from furin-cleavable proinsulin expressing HEK293 cells showed that 94% of the ILI product was mature insulin (Fig. 1B). Similar results were observed using C₂C₁₂ and L6 cell lines, which processed furin-cleavable proinsulin by 95% and 100% respectively, as opposed to 29% and 7% respectively for wild-type proinsulin (Fig. 1C). Transfection of all three cell lines with empty pCXN2 vector produced no detectable ILI.

Comparison of furin-cleavable and wild-type ILI production

It can also be seen from Fig. 1A and B that, in comparison with the expression of wild-type proinsulin, the total amount of ILI was significantly lower when furin-cleavable cDNA was expressed.

Table 1 The furin-cleavable rat proinsulin I consensus sequences that were engineered and used to transfect human HEK293, rat L6 and mouse C₂C₁₂ cell lines, showing cleavage sites with changes from the wild-type sequence underlined, and CpG island distributions

Construct	B:C cleavage site	C:A cleavage site	CpG island distribution
pCXN2rIns1	lys ser arg arg : glu val glu	arg gln lys arg : gly ile val	Wild-type
pCXN2rIns2	lys ser arg arg : <u>lys</u> <u>arg</u> glu	arg <u>arg</u> lys arg : gly ile val	4 additional
pCXN2rIns3	lys ser arg arg : <u>lys</u> <u>arg</u> glu	arg <u>arg</u> lys arg : gly ile val	Wild-type
pCXN2rIns4	lys ser arg arg : glu val glu	arg gln lys arg : gly ile val	No CpG islands
pCXN2rIns5	lys ser arg arg : <u>lys</u> <u>arg</u> glu	arg <u>arg</u> lys arg : gly ile val	No CpG islands
pCXN2rIns6	lys ser arg arg : <u>lys</u> <u>arg</u> glu	arg gln lys arg : gly ile val	Wild-type
pCXN2rIns7	lys ser arg arg : glu val glu	arg <u>arg</u> lys arg : gly ile val	Wild-type

Experiments using equal numbers of cells with identical transfection and culturing conditions showed that the ratios of furin-cleavable ILI to wild-type ILI in HEK293, C₂C₁₂ and L6 cells were 46%, 25% and 27% respectively. Many other groups using human and rat proinsulin sequences containing a variety of mutations to facilitate furin cleavage in different cell types have also reported reductions in furin-cleavable proinsulin expression of up to 90% (Yanagita *et al.* 1992, Hunt *et al.* 1996, Simonson *et al.* 1996, Short *et al.* 1998, Shaw *et al.* 2002, Scougall *et al.* 2003).

Preproinsulin mRNA stability

The relative half-lives of wild-type and furin-cleavable rat preproinsulin I mRNAs were studied using C₂C₁₂ cells in which expression of furin-cleavable proinsulin is only 25% that of wild-type protein. Cell cultures were transiently transfected with either pCXN2rIns1 (wild-type) or pCXN2rIns2 (furin-cleavable), exactly as for the ILI expression studies. Twenty-four hours after transfection, the cells were treated with actinomycin D (5 µg/ml) and the relative amounts of the two forms of preproinsulin mRNA were determined by RT-PCR. The two proinsulin constructs were expressed equally (data not shown) and expression of the housekeeping gene, *GADH*, was used to normalise the RNA content, as this gene has been shown to be a suitable internal control (Lee *et al.* 2002). Control experiments showed that, despite a DNase I digestion step, RNA preparations from transfected cells contained expression vector that was subject to amplification (data not shown). Appropriate primers were designed to avoid

possible contamination from genomic or plasmid DNA (described in Materials and methods). The RT-PCR procedures were also optimised to ensure that amplification occurred in the linear phase. The decay curves for the wild-type and furin-cleavable preproinsulin mRNA were similar (Fig. 2). Statistical analysis of the data by two-way ANOVA revealed no significant difference in the stability of wild-type and furin-cleavable preproinsulin mRNAs ($P=0.78$).

Effect of CpG islands on proinsulin expression

In the light of reports that the CpG dinucleotide content of the sequence can have a significant effect on transcription in mouse cells (Scrabble & Stambrook 1997), the distribution of CpG dinucleotides within the rat proinsulin I was examined. Comparison of the CpG dinucleotide arrangement in wild-type rat proinsulin cDNA (Fig. 3A) with the furin-cleavable form (Fig. 3B) showed that four additional CpG dinucleotides were present as a result of the mutations used to create the furin consensus sequences. Furthermore, these additional CpG dinucleotides formed two clusters. To determine whether these additional CpG dinucleotides were responsible for the reduced expression of the furin-cleavable proinsulin, wild-type rat proinsulin I cDNA was mutated, using alternative codon usage, to create a furin-cleavable proinsulin cDNA that coded for the same amino acid sequence as the previously used furin-cleavable proinsulin but with an unaltered wild-type CpG dinucleotide distribution. This cDNA was inserted into the pCXN2 expression vector to create pCXN2rIns3 (Table 1) and used to transfect both

Figure 1 Proinsulin processing. Medium from cell cultures that had been transiently transfected with either pCXN2rIns1 or pCXN2rIns2 was loaded on a Sephadex G-50 Super Fine column. Column fractions (1.0 ml) were collected and assayed for insulin-like immunoreactivity (ILI) by radioimmunoassay. (A) Gel filtration profile of medium from HEK293 cells transfected with pCXN2rIns1 encoding wild-type proinsulin. (B) Gel filtration profile of medium from HEK293 cells transfected with pCXN2rIns2 encoding furin-cleavable proinsulin. (C) Degree of proinsulin processing to insulin determined by Sephadex G-50 chromatography for all three cell types expressing either wild-type or furin-cleavable proinsulin.

Figure 2 Comparison of the stability of wild-type or furin-cleavable preproinsulin mRNAs. C₂C₁₂ cells were transfected with either pCXN2rIns1 or pCXN2rIns2 expressing wild-type (◆) or furin-cleavable preproinsulin (■) respectively, and subsequently treated with 5 µg/ml actinomycin D after 24 h. RNA was extracted at various times (6, 12, 24 and 36 h) after treatment and equal amounts were used for semiquantitative RT-PCR as described in Materials and methods. Three separate RNA preparations were used and RT-PCR assays were performed in duplicate. Data (means ± S.E.M.) were normalised for the expression of endogenous *GADH* gene and expressed as percent of the value at time 0.

rodent cell lines, C₂C₁₂ and L6, and the human HEK293 cells. There was no difference in ILI between furin-cleavable proinsulin sequence containing the four additional CpG dinucleotides (pCXN2rIns2) and the wild-type CpG dinucleotide distribution (pCXN2rIns3) in all three cell lines (Fig. 4).

It has also been reported that the total CpG dinucleotide content of a sequence can have a direct bearing upon its level of transcription in mouse (Cronin *et al.* 2001). A drastic reduction in the number of CpG dinucleotides leads to a significantly increased expression, particularly when the sequence is flanked by CpG-poor regions. As the pCXN2 expression vector used in these studies also contains CpG-poor regions both upstream and downstream of the cDNA insert, studies were carried out to investigate the influence of total CpG dinucleotide content on expression of rat proinsulin. *In situ* mutagenesis was used to create both wild-type and furin-cleavable rat proinsulin I cDNAs containing no CpG dinucleotides. This was accomplished by means of alternative codon usage and resulted in the elimination of all 10 CpG dinucleotides within the coding sequence and six in the multiple cloning

Figure 3 Distribution of CpG islands in wild-type and furin-cleavable preproinsulin cDNAs. Individual CpG islands within the rat preproinsulin I cDNA are displayed as vertical lines. (A) Wild-type sequence. (B) Furin-cleavable sequence.

sequences immediately flanking the inserts. The integrity of the mutated proinsulin constructs was confirmed by DNA sequence analysis and shown to encode for the same proteins as the unmutated forms. All three cell lines (C₂C₁₂, L6 and HEK293) were transfected with wild-type rat proinsulin I with normal CpG distribution (pCXN2rIns1), wild-type rat proinsulin I devoid of all CpG dinucleotides (pCXN2rIns4), furin-cleavable rat proinsulin I containing four additional CpG dinucleotides (pCXN2rIns2) or furin-cleavable rat proinsulin I devoid of all CpG dinucleotides (pCXN2rIns5). No major differences were detected in (pro)insulin expression as determined by ILI between constructs containing CpG dinucleotides and those devoid of CpG dinucleotides for both wild-type and furin-cleavable proinsulins (Fig. 4).

Effect of amino acid sequence on proinsulin expression

In order to investigate the effect on proinsulin production of the amino acid substitutions used to create the commonly used furin consensus sequence present in pCXN2rIns2, two further rat proinsulin I cDNAs were made. The expression vector pCXN2rIns6 contained two amino acid substitutions at the B:C cleavage site to create a

furin consensus site, whereas the A:C cleavage site was unaltered. In pCXN2rIns7, the B:C cleavage site was unaltered and the C:A cleavage site was mutated (Table 1).

These constructs were used to transfect all three cell lines and the ILI was measured as before (Fig. 4). Interestingly, the level of expression of proinsulin using the pCXN2rIns6 construct was markedly increased compared with the other mutated cDNAs and equalled the levels of wild-type expression of proinsulin in HEK293 and C₂C₁₂ cell lines. ILI for pCXN2rIns7 was intermediate between those of pCXN2rIns6 and constructs encoding both mutated cleavage sites (pCXN2rIns2, -3 and -5).

The unmutated C:A cleavage site in rat proinsulin I encoded by pCXN2rIns6 contained a naturally occurring furin consensus site. The degree of processing into mature insulin was determined using HEK293 cells and Sephadex G-50 chromatography and found to be 77% (data not shown). These results suggest that efficient expression and processing of human proinsulin in non-endocrine cells could be achieved by using a construct in which the B:C cleavage site is mutated to contain the furin consensus sequence used in these studies and the C:A cleavage site is mutated at a single residue to contain the naturally occurring furin consensus present in rat insulin.

Figure 4 Expression of insulin-like immunoreactivity (ILI) in different cell types. Cells were transfected using pCXN2rlns constructs expressing various rat proinsulin I cDNAs described in Table 1. Media were collected after 24 h and the ILI determined. (A) HEK293 cell line. (B) C₂C₁₂ cell line. (C) L6 cell line. Mean and S.E.M. values ($n=6$) were calculated from duplicate ILI assays and the results were representative on two or three occasions.

Discussion

In order to overcome the absence of the proinsulin-processing endopeptidases PC2 and PC3 in non-endocrine cells, proinsulin is mutated at the cleavage sites to contain consensus sequences for furin, which is Arg⁻⁴-X⁻³-Lys/Arg⁻²-Arg⁻¹ ↓ X⁺¹ (Hosaka *et al.* 1991). One of the most commonly used sets of mutations (Yanagita *et al.* 1992) that creates two tetrabasic furin recognition sites (Table 1) has been used in these studies, as the resulting insulin has wild-type primary sequence. Furthermore, rat proinsulin I and human proinsulin containing these furin-cleavable consensus sequences have been shown to elicit a hypoglycaemic response in diabetic mice (Kon *et al.* 1999) and rats (Shaw *et al.* 2002) respectively. Another popular furin consensus sequence used in human proinsulin (with amino acid substitutions underlined) is B:C Arg-Thr-Lys-Arg-:Glu-Ala, C:A Arg-Gln-Lys-Arg-:Gly-Ile (Groskreutz *et al.* 1994). Although this proinsulin is processed, the insulin formed contains a mutation in the B chain, Lys^{B29} to Arg. Ideally, a natural product is preferable for therapeutic purposes.

Other less common mutated proinsulins include: B:C Lys-Thr-Arg-Arg-:Lys-Ala, C:A Arg-Lys-Lys-Arg-:Gly-Ile (Yamasaki *et al.* 1999); B:C Lys-Thr-Arg-Arg-:Arg-Arg C:A Arg-Arg-Arg-Arg-:Gly-Ile (Short *et al.* 1998); B:C Arg-Thr-Arg-Arg-:Glu-Ala, C:A Arg-Gln-Lys-Arg-:Gly-Ile (Falqui *et al.* 1999) and Lys-Thr-Arg-Arg-:Glu-Ala, C:A Arg-Gln-Lys-Arg-:Gly-Ile (Hunt *et al.* 1996, O'Driscoll *et al.* 2002). The great variety of mutations used, some of which do not even create furin consensus sequences, is indicative of the central importance of proinsulin processing in the realisation of an effective long-term treatment for diabetes mellitus.

An attempt to circumvent the problem of enzymatic conversion of proinsulin has been made by replacing 35 amino acids of the C peptide with a short turn-forming heptapeptide linker (Lee *et al.* 2000). Although the single-chain insulin analogue possessed some biological activity, binding to the insulin receptor was only 27% of that of insulin. This fact, plus the desirability of not presenting the immune system with a novel form of insulin, suggest that it remains best to use furin-processed proinsulin.

Regardless of the mutation strategy, whenever production of insulin using furin-cleavable

proinsulin has been compared with natural proinsulin, expression of furin-cleavable proinsulin has always been drastically reduced. In the present study, transfection of cell lines from three different species, namely human HEK293, mouse C₂C₁₂ and rat L6, followed by Sephadex G-50 column chromatography of the culture media clearly demonstrated that the mutations used in our experiments greatly enhanced the processing of proinsulin to insulin. We observed that processing of proinsulin was 94% in HEK293 cells, 95% in C₂C₁₂ cells and 100% in L6. The efficient processing was accompanied by greatly reduced synthesis, with furin-cleavable (pro)insulin being expressed at 46%, 25% and 27% of wild-type (pro)insulin in HEK293, C₂C₁₂ and L6 cells respectively.

One suggestion to account for this reduction has been that sequences in the 3' untranslated region of the mRNA have a role in maintaining stability and expression. However, Shaw *et al.* 2002 made use of a full-length human proinsulin cDNA and observed a 90% reduction in expression upon creation of furin consensus sequences at the cleavage sites. As the furin-cleavable proinsulin cDNA remained identical to the wild-type form in every other respect, the role of untranslated regions in reducing expression can be excluded.

In view of the direct relationship between mRNA half-life and production of the encoded protein, we examined the relative stability of wild-type and furin-cleavable preproinsulin mRNA in C₂C₁₂ cells. Transcription was blocked with 5 µg/ml actinomycin D and the relative amounts of wild-type and furin-cleavable preproinsulin mRNA assayed by RT-PCR. No significant difference was detected in the half-lives of the two mRNAs (Fig. 2), so dissimilar mRNA stability cannot account for reduced expression of furin-cleavable proinsulin.

The possible effect of the introduction of four additional CpG dinucleotides in the furin-cleavable proinsulin cDNA as a consequence of the mutations used to generate the furin consensus sequences was investigated by means of alternative codon use to create a furin-cleavable proinsulin cDNA with unaltered wild-type distribution of CpG dinucleotides (pCXN2rIns3). No correlation between the presence of extra CpG islands and furin-cleavable proinsulin expression was observed. Similarly, the presence or total absence of CpG dinucleotides in the coding and flanking sequences

had no bearing on the expression of wild-type or furin-cleavable proinsulin in all three cell lines. Therefore, CpG dinucleotide distribution cannot account for decreased expression of furin-cleavable proinsulin.

The effect of basic amino acid substitutions on proinsulin expression was investigated by preparing a rat proinsulin I in which only the B:C cleavage site was mutated, to contain the tetrabasic furin recognition sequence by substituting glutamate⁺¹ and valine⁺² at the N-terminal of the C peptide with lysine and arginine residues. The C:A cleavage site was left unaltered (pCXN2rIns6). Transfection of this cDNA in all three cell types led to a large increase in proinsulin production, matching levels of unmutated protein in HEK293 and C₂C₁₂ (Fig. 4). Also, as this construct contained a natural furin consensus sequence at the C:A cleavage site, 77% processing was achieved in HEK293 cells. Mutation of the C:A cleavage site involved substitution of glutamine⁻³ at the C-terminal of the C-peptide with a basic arginine residue, and when this modified cleavage site alone was used (pCXN2rIns7), a decrease in expression of proinsulin was observed (Fig. 4). Therefore, the crucial factor in reducing the expression of furin-cleavable proinsulin is substitution with a basic amino acid at the C:A cleavage site, and this effect is compounded when further basic charges are introduced at the B:C cleavage site.

Once regarded as a random coil of minor importance, the C-peptide within proinsulin is now considered to be an autonomous folding unit (Brems *et al.* 1990). More recent studies have indicated that the highly conserved acidic N-terminal region of the C-peptide (Glu-Ala-Glu-Asp in human and Glu-Val-Glu-Asp in rat) is important in insulin precursor folding, and that the C-peptide may have an intramolecular chaperone-like function (Chen *et al.* 2002). The conservation of C-peptide residues near the C:A cleavage site may reflect functional constraints of the region, especially as similar structures are absent from the B:C cleavage site (Gross *et al.* 1989). Aromatic [¹H]-NMR and photo-CIDNP studies of human proinsulin have revealed that there is a local stable structure at the C:A cleavage site (referred to as the 'C:A knuckle') that, unlike the B:C cleavage site, alters the packing of the hydrophobic core of the insulin moiety (Weiss *et al.* 1990). Furthermore, these perturbations are released by cleavage of the

C:A cleavage site, but not of the B:C cleavage site. The substitution of two neutral amino acids to two basic residues in the centre of the C-peptide (gly¹⁵ and ala²⁰ to lys) in human proinsulin has been reported to have no effect on *in vitro* protein refolding (Chen *et al.* 2002). This would strongly suggest that local conformational perturbation, rather than overall charge conservation, has a significant role in affecting the stability of proinsulin.

Taken together, these results suggest that the major cause of decreased expression of furin-cleavable proinsulin is the addition of basic charges at the C:A cleavage site. Hunt *et al.* (1996) made a different basic residue substitution at the C:A cleavage site in human proinsulin by replacing leucine⁻⁴ at the C-terminal of the C-peptide with an arginine residue. An 18-fold reduction in levels of immunoreactive insulin normalised against mRNA in CHO cells was reported, with no apparent reduction in the amount of mRNA.

The increased reduction in expression of proinsulin that we observed when both cleavage sites were mutated suggests that the effect of the introduction of basic charges at the C:A cleavage site is worsened by basic amino acid substitution at the B:C cleavage site. There could be several factors contributing to this. The N-terminal of the A-chain α -helix packs against the C-terminal of the B-chain in the crystal state, permitting transmission of perturbations (Weiss *et al.* 1990). In addition, the mutations at the B:C cleavage site resulted in glutamate⁺¹ and valine⁺² in the conserved acidic N-terminal of the C-peptide being substituted with basic lysine and arginine. A possible combination of altered folding of the insulin precursor and perturbations resulting from mutation of the C:A knuckle could lead to reduced protein stability. There are several possible mechanisms by which this could occur. These include impaired folding in the endoplasmic reticulum during synthesis, leading to degradation, or enhanced sorting to endosomal or lysosomal pathways. Further experiments are being undertaken to determine the fate of furin-cleavable proinsulin containing the mutations described here.

The rationale for strategies to convert dibasic proinsulin cleavage sites into substrates for tribasic-requiring furin has generally been to make the cleavage sites contain as many basic residues as possible. This has been based on the observation that many proproteins contain tetra- or pentabasic

cleavage sites (Duguay *et al.* 1997, Rosenblatt & Dickerson 1997). Therefore, increasing the number of basic residues in the furin consensus sequence would facilitate processing and greater yields of insulin. We propose that the opposite approach should be taken, and that researchers should consider designs using minimal substitutions with basic amino acids, especially at the C:A cleavage site. In addition, the relative merits of using either lysine or arginine at position -2 of the furin consensus sequence should be examined.

Acknowledgement

Supported by a grant from the Wellcome Trust.

References

- Bach JF 1994 Insulin-dependent diabetes mellitus as an autoimmune disease. *Endocrine Reviews* **15** 516–542.
- Berger M 1996 To bridge science and patient care in diabetes. *Diabetologia* **39** 749–757.
- Brems DN, Brown PL, Heckenlaible LA & Frank BH 1990 Equilibrium denaturation of insulin and proinsulin. *Biochemistry* **29** 9289–9293.
- Chen LM, Yang XW & Tang JG 2002 Acidic residues on the N-terminus of proinsulin C-peptide are important for the folding of insulin precursor. *Journal of Biochemistry* **131** 855–859.
- Cronin CA, Gluba W & Scrable H 2001 The lac operator-repressor system is functional in the mouse. *Genes and Development* **15** 1506–1517.
- Docherty K 1997 Gene therapy for diabetes mellitus. *Clinical Science* **92** 321–330.
- Duguay SJ, Milewski WM, Young BD, Nakayama K & Steiner DF 1997 Processing of wild-type and mutant proinsulin-like growth factor-1A by subtilisin-related proprotein convertases. *Journal of Biological Chemistry* **272** 6663–6670.
- Falqui L, Martinenghi S, Severini GM, Corbella P, Taglietti MV, Arcelloni C, Sarugeri E, Monti LD, Paroni R, Dozio N, Pozza G & Bordinon C 1999 Reversal of diabetes in mice by implantation of human fibroblasts genetically engineered to release mature human insulin. *Human Gene Therapy* **10** 1753–1762.
- Fricker L 1988 Carboxypeptidase E. *Annual Review of Physiology* **50** 309–321.
- Fricker LD, Supattapone S & Snyder SH 1982 Enkephalin convertase: a specific enkephalin synthesizing carboxypeptidase in adrenal chromaffin granules, brain, and pituitary gland. *Life Science* **31** 1841–1844.
- Groskreutz DJ, Sliwkowski MX & Gorman CM 1994 Genetically engineered proinsulin constitutively processed and secreted as mature, active insulin. *Journal of Biological Chemistry* **269** 6241–6245.
- Gross DJ, Villa-Komaroff L, Kahn CR, Weir GC & Halban PA 1989 Deletion of a highly conserved tetrapeptide sequence of the proinsulin connecting peptide (C-peptide) inhibits proinsulin to insulin conversion by transfected pituitary corticotroph (AtT20) cells. *Journal of Biological Chemistry* **264** 21486–21490.
- Halban PA, Kahn SE, Lernmark A & Rhodes CJ 2001 Gene and cell-replacement therapy in the treatment of type 1 diabetes: how high must the standards be set? *Diabetes* **50** 2181–2191.

- Hosaka M, Nagahama M, Kim WS, Watanabe T, Hatsuzawa K, Ikemizu J, Murakami K & Nakayama K 1991 Arg-X-Lys/Arg-Arg motif as a signal for precursor cleavage catalyzed by furin within the constitutive secretory pathway. *Journal of Biological Chemistry* **266** 12127–12130.
- Hunt SMN, Tait AS, Gray PP & Sleight MJ 1996 Processing of mutated human proinsulin to mature insulin in the non-endocrine cell line, CHO. *Cytotechnology* **21** 279–288.
- Kon OL, Sivakumar S, Teoh KL, Lok SH & Long YC 1999 Naked plasmid-mediated gene transfer to skeletal muscle ameliorates diabetes mellitus. *Journal of Gene Medicine* **1** 186–194.
- Lee HC, Kim SJ, Kim KS, Shin HC & Yoon JW 2000 Remission in models of type 1 diabetes by gene therapy using a single-chain insulin analogue. *Nature* **408** 483–488.
- Lee PD, Sladek R, Greenwood CM & Hudson TJ 2002 Control genes and variability: absence of ubiquitous reference transcripts in diverse mammalian expression studies. *Genome Research* **12** 292–297.
- Muzzin P, Eisensmith RC, Copeland KC & Woo SL 1997 Hepatic insulin gene expression as treatment for type 1 diabetes mellitus in rats. *Molecular Endocrinology* **11** 833–837.
- Nishigori T, Yanagita M & Takeuchi T 1996 Proinsulin cleaved by furin is processed to chromatographically mature insulin by carboxypeptidases in nonneuroendocrine cells. *Peptides* **17** 789–796.
- Niwa H, Yamamura K & Miyazaki J 1991 Efficient selection for high-expression transfectants with a novel eukaryotic vector. *Gene* **108** 193–200.
- O'Driscoll L, Gammell P & Clynes M 2002 Engineering Vero cells to secrete human insulin. *In Vitro Cellular and Developmental Biology. Animal* **38** 146–153.
- Rosenblatt MI & Dickerson IM 1997 Endoproteolysis at tetrabasic amino acid sites in procalcitonin gene-related peptide by pituitary cell lines. *Peptides* **18** 567–576.
- Scougall KT, Maltin CA & Shaw JAM 2003 Tetracycline-regulated secretion of human insulin in a transfected non-endocrine cell line. *Journal of Molecular Endocrinology* **30** 331–346.
- Scrabble H & Stambrook PJ 1997 Activation of the lac repressor in the transgenic mouse. *Genetics* **147** 297–304.
- Seidah NG, Gaspar L, Mion P, Marcinkiewicz M, Mbikay M & Chrétien M 1990 cDNA sequence of two distinct pituitary proteins homologous to Kex2 and furin gene products: tissue-specific mRNAs encoding candidates for pro-hormone processing proteinases. *DNA and Cell Biology* **9** 415–424.
- Shaw JAM, Delday MI, Hart AW, Docherty HM, Maltin CA & Docherty K 2002 Secretion of bioactive human insulin following plasmid-mediated gene transfer to non-neuroendocrine cell lines, primary cultures and rat skeletal muscle *in vivo*. *Journal of Endocrinology* **172** 653–672.
- Short DK, Okada S, Yamauchi K & Pessin JE 1998 Adenovirus-mediated transfer of a modified human proinsulin gene reverses hyperglycemia in diabetic mice. *American Journal of Physiology* **275** E748–E756.
- Simonson GD, Groskreutz DJ, Gorman CM & MacDonald MJ 1996 Synthesis and processing of genetically modified human proinsulin by rat myoblast primary cultures. *Human Gene Therapy* **7** 71–78.
- Smeeckens SP & Steiner DF 1990 Identification of a human insulinoma cDNA encoding a second putative protein structurally related to the yeast dibasic processing protease Kex2. *Journal of Biological Chemistry* **265** 2997–3000.
- Smeeckens SP, Avruch AS, LaMendola J, Chan SJ & Steiner DF 1991 Identification of a cDNA encoding a second putative prohormone convertase related to PC2 in AtT20 cells and islets of Langerhans. *Proceedings of the National Academy of Sciences of the United States of America* **88** 340–344.
- Steiner DF 1998 The proprotein convertases. *Current Opinion in Chemical Biology* **2** 31–39.
- Van de Ven WJ, Voorberg J, Fontijn R, Pannekoek H, van den Ouweland AM, van Duijnhoven HL, Roebroek AJ & Siezen RJ 1990 Furin is a subtilisin-like proprotein processing enzyme in higher eukaryotes. *Molecular Biology Reports* **14** 265–275.
- Weiss MA, Frank BH, Khait I, Pekar A, Heiney R, Shoelson SE & Neuringer LJ 1990 NMR and photo-CIDNP studies of human proinsulin and prohormone processing intermediates with application to endopeptidase recognition. *Biochemistry* **29** 8389–8401.
- Weiss RA 1998 Transgenic pigs and virus adaptation. *Nature* **391** 327–328.
- Yamasaki K, Sasaki T, Nemoto M, Eto Y & Tajima N 1999 Differentiation-induced insulin secretion from nonendocrine cells with engineered human proinsulin cDNA. *Biochemical and Biophysical Research Communications* **265** 361–365.
- Yanagita M, Nakayama K & Takeuchi T 1992 Processing of mutated proinsulin with tetrabasic cleavage sites to bioactive insulin in non-endocrine cells, COS-7. *FEBS Letters* **311** 55–59.
- Yanagita M, Hoshino H, Nakayama K & Takeuchi T 1993 Processing of mutated proinsulin with tetrabasic cleavage sites to mature insulin reflects the expression of furin in nonendocrine cell lines. *Endocrinology* **133** 639–644.

Received 18 June 2003

Accepted 26 August 2003