RNA–protein interactions in the 5′ untranslated region of preproinsulin mRNA

S. W. Knight and K. Docherty

Department of Medicine, University of Birmingham, Queen Elizabeth Hospital, Birmingham B15 2TH, U.K.

Requests for offprints should be addressed to K. Docherty

ABSTRACT

A comparison between species of the 5′ untranslated region of preproinsulin mRNA revealed conserved sequences associated with a potential stem–loop structure. The present study was undertaken to determine whether specific protein interactions exist with mRNA sequences involved in the formation or stabilization of this structure in the 5′ untranslated region. 32P-labelled RNA probes corresponding to sequences from this region were synthesized by an in-vitro transcription reaction and used in electrophoretic mobility shift and u.v.-crosslinking studies with cytoplasmic protein extracts from a number of cell lines. Specific protein–RNA interactions were mapped to a sequence located between nucleotides −21 and −50 upstream of the AUG start codon. A number of proteins of molecular mass 25 kDa, 40 kDa, 46 kDa, 58 kDa, 69 kDa, 97 kDa, 110 kDa and 160 kDa were specifically crosslinked to this sequence. The observed specific protein–RNA interactions in the 5′ untranslated region may affect the activity of preproinsulin mRNA.

Journal of Molecular Endocrinology (1992) 8, 225–234

INTRODUCTION

Glucose is the major physiological regulator of plasma insulin levels (Ashcroft, Bunce, Lowry et al. 1978). Glucose stimulates the secretion of insulin from stored vesicles in the β cells of the islets of Langerhans and the concomitant enhancement of the synthesis of insulin (Schuit, Kickens & Pipeleers, 1991) and other secretory vesicle proteins (Guest, Bailyes, Rutherford & Hutton, 1991). Acute stimulation of insulin synthesis by glucose occurs translationally (Permutt & Kipnis, 1972; Permutt, 1974; Itoh & Okamoto, 1980; Itoh, Ohshima, Nose & Okamoto, 1982; Welsh, Scherberg, Gilmore & Steiner, 1986), while prolonged exposure to elevated glucose levels increases transcription (Brunstedt & Chan, 1982; Nielsen, Welsh, Casadaban & Steiner, 1985) and mRNA stability (Welsh, Nielsen, MacKrell & Steiner, 1985; Philippe & Missotten, 1990).

Sequence and structural motifs in the untranslated regions (UTRs) of eukaryotic mRNA have both been associated with the regulation of translation and mRNA stability (Hershey, 1991; Kozak, 1991a). It is possible that cis-acting elements within the preproinsulin mRNA or nascent polypeptide may account for the observed effect of glucose on translation. Preproinsulin mRNA consists of a coding region flanked by both 5′ and 3′ UTRs, the sequence and intronic organization of which are highly conserved in evolution. The functional significance of the conserved sequences (Bell & Pesca-dor, 1984) and secondary structure (Lomedico, Rosenthal, Efstratiadis et al. 1979) of the preproinsulin mRNA 5′ UTR remains undetermined.

The aim of the present study was to investigate specific protein–RNA binding within the 5′ UTR using electrophoretic mobility shift assays and in-vitro u.v. crosslinking, on the basis that such interactions may be involved in the regulation of activity of preproinsulin mRNA.

MATERIALS AND METHODS

Cell lines

The cell lines used in this study were HIT T15, a hamster β-cell line (Santere, Cook, Crisel et al. 1981), RINm5F, a β-cell line derived from a rat insulinoma (Gazdar, Chick, Oie et al. 1980), monkey COS-7 kidney cells (Gluzman, 1981), human HEP G2 liver cells (Aden, Fogel, Plotkin et al. 1979) and mouse NIH/3T3 fibroblast-like cells (Jainchill, Aaronson & Todaro, 1969).
Preparation of cytoplasmic extracts

Cells were grown to confluence, washed twice with 5 ml ice-cold phosphate-buffered saline and harvested by scraping with a rubber spatula. The cells were centrifuged at 400 g for 10 min and the pellet washed once with phosphate-buffered saline. The cell pellet was resuspended in five packed cell volumes of ice-cold 10 mm Hepes-KOH, pH 7.9, 10 mm KCl, 1.5 mm MgCl2, 0.5 mm dithiothreitol, 0.5 mm phenylmethylsulphonyl fluoride, 2.5 mg aprotinin/ml, 0.5 mm leupeptin, 0.5 mm pepstatin A and 0.5 mm antipain, and homogenized with 100 strokes of a Dounce homogenizer. All buffers used in the preparation of the cytoplasmic extracts contained the above cocktail of proteinase inhibitors. The cell lysate was then centrifuged at 400 g for 20 min at 4°C and the post-nuclear supernatant collected and centrifuged at 100000 g for 20 min at 4°C. The supernatant was collected and solid ammonium sulphate slowly added to give a final concentration of 330 mg/ml. Following centrifugation at 25000 g for 20 min at 4°C, the precipitated proteins were resuspended in 20 mm Hepes, pH 7.9, containing 25% (v/v) glycerol, 50 mm KCl, 1.5 mm MgCl2, 0.2 mm EDTA, and 0.5 mm dithiothreitol, and dialysed overnight against 100 volumes of the same buffer. The protein concentration was determined using the Bio-Rad modification of the Bradford assay (Bradford, 1976) and the extracts were then stored in aliquots at -70°C.

Synthesis of RNA probes

RNA probes were synthesized in an in-vitro transcription reaction, according to the method of Milligan, Withereill & Uhlenbeck (1987). Synthetic DNA oligonucleotides complementary to sequences within the human preproinsulin mRNA 5' UTR and the core iron response element (IRE) from the human ferritin 5' UTR (Rouault, Hentze, Haile et al. 1989) were purchased from Alta Bioscience (University of Birmingham). All oligodeoxynucleotides contained a 17 nucleotide 3' extension (5'-TAGTGAGTCGTATTA-3') corresponding to the promoter for T7 RNA polymerase. These oligodeoxynucleotides were annealed to a short primer oligodeoxynucleotide (5'-TAATACGACTCACTATAG-3'), such that a double-stranded T7 RNA polymerase primer was formed at the 3' end of the template oligodeoxynucleotide. For the annealing reaction, 3 μg primer and 3 μg complementary oligodeoxynucleotide were incubated in 10 mm Tris-HCl, pH 7.9, 2 mm MgCl2, 50 mm NaCl and 1 mm EDTA at 90°C for 5 min. The mixture was then allowed to cool slowly to room temperature.

For synthesis of the RNA probes, the annealed oligodeoxynucleotides (500 ng) were incubated for 60 min at 37°C in a reaction containing 40 mm Tris-HCl, pH 8.0, 15 mm MgCl2, 10 mm dithiothreitol, 1 mm ATP, 1 mm CTP, 1 mm UTP, 10 μCi [α-32P]GTP (Amersham International plc, Aylesbury, Bucks, U.K.), 70 units T7 RNA polymerase (Pharmacia, Milton Keynes, Bucks, U.K.) and 15 units RNAs (Pharmacia, Milton Keynes, Bucks, U.K.) in a final volume of 10 μl. Five units of DNase 1 (Pharmacia) were then added, and the mixture incubated for a further 15 min at 37°C, followed by 5 min at 70°C to inactivate the enzymes.

The products of the transcription reaction were separated on 7 m urea-20% polyacrylamide gels for the small (approximately 20 nucleotide) RNA probes and 7 m urea-8% polyacrylamide gels for the larger (approximately 60 nucleotide) RNA probes. Following electrophoresis, the gels were wrapped in Saran Wrap (Appleton Woods, Birmingham, U.K.) and exposed to X-ray film for a few minutes. The appropriately sized bands were identified by comparison with 32P-labelled oligodeoxynucleotide size markers, excised with a scalpel and then eluted from the gel with two changes over 16 h of 100 μl elution buffer containing 0.5 m ammonium acetate, 10 mm magnesium acetate, 1 mm EDTA and 0.1% (w/v) sodium dodecyl sulphate (SDS). The RNA probes were then ethanol-precipitated and resuspended in 20 μl water.

Electrophoretic mobility shift assays

32P-Labelled RNA probes were incubated for 30 min at room temperature in 10 mm Tris-HCl, pH 8.0, containing 20 mm KCl, 10-20 μg protein extract and 1 unit RNAs in a total volume of 10 μl. When unlabelled competitions were performed, the unlabelled RNA probe was added at the same time as the 32P-labelled RNA probe. Gel loading buffer (50% glycerol, 0.25% (w/v) xylene cyanol, 0.35% (w/v) bromophenol blue) was added, and the RNA-protein complexes separated from the free RNA probe on 6% non-denaturing polyacrylamide gels run in 45 mm Tris-HCl, pH 8.0, 45 mm sodium borate and 1 mm EDTA. The gel was dried and exposed to Fuji X-ray film for autoradiography.

Ultraviolet crosslinking

32P-Labelled RNA probes were incubated for 30 min at room temperature in 10 mm Tris-HCl, pH 8.0, containing 20 mm KCl, 20 μg cytoplasmic protein extract and 3 units RNAs in a total volume of 10 μl. The samples were then placed on ice and irradiated with u.v. light at 254 nm for 15 min using a Spectrolinker XL1000 (AMS Biotechnology, Burford, Oxon, U.K.). Following incubation with 1500 units T1 RNAs (GIBCO-BRL, Paisley, Strath-
clyde, U.K.) for 15 min at room temperature, the samples were subjected to SDS-polyacrylamide gel electrophoresis (Laemmli, 1970) and autoradiography.

RESULTS

An examination of the 5′ untranslated sequences of preproinsulin mRNAs from a number of different species (Fig. 1) demonstrated that the overall length and specific sequences are highly conserved. Each of the 5′ UTRs were between 57 and 63 nucleotides in length, except for the sequence from dog which had a deletion of the first 11 nucleotides of the second exon. The 5′ UTRs were interrupted by an intron in the pre-mRNA and, as noted previously (Bell & Seino, 1990), the position of the splice junction was conserved, with the exception of the murine preproinsulin 5′ UTRs which had the site translocated three bases further upstream.

In addition to the sequence similarities, the use of algorithms for determining RNA secondary structure (Zuker & Stiegler, 1981) predicted the presence of a stem–loop structure in the 5′ UTR of preproinsulin mRNA for all the species listed in Fig. 1. The structures adopted by the rat (I) and human preproinsulin 5′ UTRs are shown in Fig. 2. The most stable mRNA secondary structures for each species were not directly superimposable. However, regions of homologous sequence were located in almost identical positions in the stem–loop structure.

Given that the interaction of proteins with stem–loop structures within the 5′ UTR of a number of mRNAs has been shown to modulate the translational activity of the mRNA, we were interested in determining whether specific protein interactions occurred with sequences in this region of preproinsulin mRNA. A set of overlapping 32P-labelled RNA probes were synthesized (Fig. 3): probe A (−1 to −20), probe B (−11 to −30), probe C (−21 to −40), probe D (−31 to −50) and probe E (−41 to −59), together with the full-length 5′ UTR of human preproinsulin mRNA (−1 to −59) and an RNA probe corresponding to the IRE (Rouault et al. 1989). The RNA probes were synthesized in an in-vitro transcription reaction using complementary oligodeoxynucleotide templates (Milligan et al. 1987). The initial reaction product was heterogeneous (data not shown). However, it was possible to excise the appropriately sized RNA oligonucleotide from denaturing polyacrylamide gels. The highly purified RNA probes were then used in electrophoretic mobility shift and u.v.-crosslinking studies.

Tissue distribution of proteins binding to sequences in the 5′ UTR of preproinsulin mRNA

Binding of proteins to the RNA probe corresponding to the complete preproinsulin 5′ UTR sequence was observed with cytosolic extracts from all the cell lines tested, i.e. HIT T15, RINm5F, HEP G2, COS-7 and NIH/3T3 (Fig. 4). The mobility of the retarded bands differed slightly among cell extracts. These differences might indicate the formation of different protein complexes in the various extracts or that, despite the presence of proteinase inhibitors, the extracts had undergone some proteolysis.

To characterize the binding of proteins to the preproinsulin 5′ UTR further, mobility shift assays were performed using RNA probes corresponding to discrete regions within the 5′ UTR (probes A–E) (Fig. 5). A major retarded complex of approximately the same mobility was observed for all the probes using cytoplasmic extracts from the COS-7, HIT T15, NIH/3T3 and HEP G2 cell lines. The major complex formed with the RINm5F extract migrated ahead of the major complex formed with the other extracts.

![Figure 1](https://example.com/figure1.png)

**Figure 1.** Comparison of the 5′ untranslated regions (UTRs) of mammalian preproinsulin mRNAs.

The sequences of the 5′ UTRs were aligned and colons introduced to maximize the homology. Asterisks indicate the 5′-cap sites and the underlined nucleotides indicate the splice junction. Sequences are quoted from Bell & Pescador (1984); Wentworth, Schaefer, Villa-Komaroff & Chirgwin (1986) and Chan, Episkopou, Zeitlin et al. (1984).

*Journal of Molecular Endocrinology* (1992)
Investigation of sequence-specific interactions within the 5' UTR

Sequence-specific interactions within the 5' UTR were investigated using a cytoplasmic extract from COS-7 cells. The interaction of proteins with probes A, B and E was not sequence-specific since, in addition to self-competition, competition was also observed with all of the other unlabelled RNA probes (Fig. 7). However, protein interactions with RNA probes C and D were specific, since self-competition was observed, but competition with the other RNA probes was not observed. Since cross-competition was observed with both C and D, and since these RNA sequences overlap, it is possible that the sequence-specific interaction observed within these probes represents a single RNA-protein binding site.

In-vitro u.v. crosslinking of proteins to RNA probes

In order to characterize the proteins responsible for the specific interaction with RNA probes C and D, the RNA probes were incubated with a cytoplasmic extract and proteins crosslinked to the RNA by u.v. irradiation. In preliminary experiments, the effect of duration of exposure to u.v. radiation was examined. Using a cytoplasmic extract from COS-7 cells, maximum crosslinking of proteins to RNA probes A–E was observed following exposure to u.v. radiation for 10–15 min (Fig. 8). Under these conditions, a single protein of 50 kDa was crosslinked to probe A, two proteins of 30 kDa and 50 kDa were crosslinked to probe B, proteins of 25 kDa, 30 kDa, 35 kDa, 40 kDa and 50 kDa were crosslinked to probe C, proteins of 30 kDa, 50 kDa and 69 kDa...
were crosslinked to probe D and proteins of 30 kDa and 55 kDa were crosslinked to probe E.

A similar pattern of crosslinking was observed with a cytoplasmic extract from HIT T15 cells (Fig. 9), in so far as more proteins were crosslinked to probe D (25 kDa, 46 kDa, 58 kDa, 69 kDa, 97 kDa, 110 kDa and 160 kDa) than to the other probes. Ultraviolet crosslinking using the IRE probe gave a band of 100 kDa, which is the expected size of the IRE-binding protein (Rouault, Tang, Kaptain et al. 1990). The absence of any of the ubiquitous proteins crosslinked to the IRE may indicate that these proteins are single-stranded RNA-binding proteins, since none of the small RNA probes have a secondary structure.

The crosslinking of a 50 kDa COS-cell protein or a 46 kDa HIT-cell protein to all the probes may explain the non-specific protein–RNA interaction observed with probes A, B and E in the mobility shift study. On the other hand, the observation that a number of proteins were specifically crosslinked to probes C and D confirms the specific interaction of proteins with these RNA probes in the mobility shift experiments.

**DISCUSSION**

Previous studies have shown that conserved sequences and structural motifs in the UTRs of eukaryotic mRNAs regulate their activity. Examples include: (i) four short open reading frames in the 5′ UTR of the GCN4 mRNA, which modulate GCN4 expression in *Saccharomyces cerevisiae* (Hinnebusch, 1988); (ii) the IRE in the 5′ UTR of ferritin mRNA which affects translation initiation (Theil, 1990); (iii) multiple copies of the IRE in the 3′ UTR of transferrin mRNA which affect the stability of the mRNA (Theil, 1990); (iv) an AU-rich motif in the 3′ UTR which modulates the stability of c-fos, c-myc and GM-CSF mRNAs (Bohjansen, Petryniak, June et al. 1991); (v) sequences in the 5′ UTR of acetyl coenzyme A carboxylase mRNA (Lopez-Casillas & Kim, 1991) and (vi) the 3′ polyadenyl tail found in most eukaryotic mRNAs which is involved in stability and influences translation (Jackson & Standart, 1990). In most of these examples regulation is facilitated by a specific RNA–protein interaction.

The identification of conserved sequences and secondary structures within the 5′ UTR of preproinsulin mRNA prompted the present study to determine whether specific RNA–protein interactions occurred within this region of the mRNA. Using mobility shift assays, protein–RNA complexes were observed with RNA probes corresponding to the complete 5′ UTR and with short probes encoding overlapping sequences within the UTR. The interaction of proteins with the complete 5′ UTR sequence may be dependent on the presence of a secondary stem–loop structure. Thus, the smaller RNA probes A–E, which cannot form a stem–loop structure, will not compete while the IRE RNA oligonucleotide, which can form a stem–loop structure, will compete. It is unlikely that identical proteins recognize the preproinsulin 5′ UTR and the IRE, but rather that proteins which bind to the 5′ UTR are recognizing the stem–loop structure in a non-specific manner. In support of this, when the IRE was used in a mobility shift assay self-competition was observed, but no competition was observed with RNA probes A–E (data not shown). In addition to the complex formed with the complete 5′ UTR probe, a sequence-specific RNA–protein in-

---

*Figure 4.* Tissue distribution of proteins binding to the 5′ untranslated region (UTR) of preproinsulin mRNA – mobility shift assay using the 32P-labelled 5′-UTR RNA and cytoplasmic protein extracts from RINm5F (RIN; 25 μg), HIT T15 (HIT; 25 μg), COS-7 (COS; 20 μg), HEP G2 (HEP; 10 μg) and NIH/3T3 (3T3; 10 μg) cells. The position of the free probe is indicated (FP). The smear at the bottom of the gel in samples treated with cytoplasmic extracts represents RNA degradation products.
Figure 5. Tissue distribution of proteins binding to sequences in the 5' untranslated region (UTR) of preproinsulin mRNA. The mobility shift assay used 32P-labelled RNA probes corresponding to sequences within the 5' UTR, i.e. A (-1 to -20), B (-11 to -30), C (-21 to -40), D (-31 to -50) and E (-41 to -59), and cytoplasmic protein extracts from COS-7 (COS; 20 μg), HIT T15 (HIT; 25 μg), RINm5F (RIN; 25 μg), NIH/3T3 (3T3; 20 μg) and HEP G2 (HEP; 20 μg) cells. In this experiment, the free probe was permitted to run off the bottom of the gel. The visible bands represent RNA–protein complexes.

In vitro u.v. crosslinking has been used to investigate numerous RNA–protein interactions including: (i) the composition of the heterogenous nuclear RNA–protein particle (Greenberg, 1980); (ii) the interaction of the poly(A) binding protein with the poly(A) tail (Moore, Chen & Whoriskey, 1988); (iii) proteins involved in the spliceosome (Wang & Pederson, 1990) and (iv) the interaction between the

Figure 6. Competitor studies on the binding of factors to the preproinsulin 5'-untranslated region (UTR) probe – competition mobility shift assay using a 32P-labelled RNA probe of the full length preproinsulin 5' UTR. Samples were incubated with an approximately tenfold excess of unlabelled RNA probe (A–E) or an unlabelled RNA corresponding to the sequence of the iron response element (IRE) as indicated and 20 μg cytoplasmic protein extract from COS-7 cells. The position of the free RNA probe is indicated (FP). The position of the specific retarded protein–RNA complex is indicated by an arrow.
**Figure 7.** Investigation of sequence-specific interactions within the 5' untranslated region – competition mobility shift assays using \(^{32}\)P-labelled RNA probes A–E. The RNA probes were incubated with a tenfold excess of unlabelled RNA oligonucleotides A–E and 20 \(\mu\)g cytoplasmic protein extract from COS-7 cells. The five gels labelled A to E represent the results for each RNA probe with the competitor RNAs indicated above each track of the gel. The position of the free probe is indicated (FP).

**Figure 8.** In-vitro u.v. crosslinking of a cytoplasmic extract from COS-7 cells to RNA probes. \(^{32}\)P-Labelled RNA probes A–E were crosslinked to 20 \(\mu\)g cytoplasmic extract from COS-7 cells as described in the Materials and Methods. Exposure to u.v. radiation was for 0, 5, 10 or 15 min, as indicated. Protein molecular size markers were run on the track on the left of the gel.
IRE and the IRE-binding protein (Leibold & Munro, 1988). In the present study the crosslinking of proteins to RNA probes A–E reinforced the findings of the mobility shift competition experiments, i.e. sequence-specific interactions occurred within probes C and D. An ubiquitous 50 kDa COS-cell protein and a 46 kDa HIT-cell protein were crosslinked to all the sequences, but specific interactions of a number of proteins with probes C and D were observed.

That RNA transcripts exist in association with protein both in the nucleus and in the cytoplasm is well established (Dreyfuss, 1986). The RNA–protein particle is a dynamic multi-protein–RNA assembly, consisting of core RNA-binding proteins (Pfnol-Roma, Choi, Matunis & Dreyfuss, 1988) and other specialized proteins which modulate the activity of the transcript. The RNA–protein particle acts to package and protect the RNA from RNase degradation, but it also provides a framework facilitating mRNA maturation in the nucleus and its translation in the cytoplasm. In the nucleus, the interactions with the spliceosome complex excise the introns via U1 small nuclear RNA recognition of the heterogenous nuclear RNA splice junction (Guthrie & Patterson, 1988), while the AAUAAA motif in the 3’ UTR enables the correct 3’-end formation and polyadenylation (Wickens, 1990). In the cytoplasm, the mRNA–protein complex provides a framework so that the translation (Theil, 1990) and/or stability (Bohjanen et al. 1991) of particular mRNAs can be differentially regulated by the interaction of specific factors.

Secondary RNA structures may affect the initiation of translation. Following recognition of the mRNA through a process facilitated by the binding of eukaryotic initiation factor-4E to the 5’ cap structure, the ribosome translocates along the 5’ UTR, melting any secondary structure, and initiates protein synthesis at the first AUG which occurs in a favourable context (Kozak, 1991b). The presence of stable stem–loop structures in the 5’ UTR can impede ribosome initiation, either by blocking the binding to or progress along the 5’ UTR (Kozak, 1989). Thus, the stability of the structures in the 5’ UTR can be correlated to the translational efficiency of a particular mRNA (Pelletier & Sonenberg, 1985). The stem–loop structures found in the preproinsulin mRNA 5’ UTR range in free energy from \( \Delta G = -13 \text{kcal/mol} \) for monkey and human to \( \Delta G = -8 \) to \(-11\text{kcal/mol} \) for rat and mouse stem–loops. By comparison with the free energy of structures shown to have initiation blocking potential (Pelletier & Sonenberg, 1985; Kozak, 1989) it is unlikely that the stem–loops in the preproinsulin mRNA 5’ UTR alone would block ribosome initiation. It is possible that the absence of a stable structure in the 5’ UTR may account, in part, for the additional increase in synthesis of insulin over and above the general increase in protein synthesis caused by glucose. The fact that both HIT T15 and RINm5F cells cannot elicit the translational control of insulin synthesis by glucose (Praz, Halban, Wollheim et al. 1983; Gold, Qian & Grodsky, 1988) implies that the secondary structure alone does not provide a sufficient explanation for the sensitivity of preproinsulin mRNA translation to changes in glucose concentration. An additional possible explanation for the conservation of a stem–loop structure in the 5’ UTR of preproinsulin mRNA is that it may stabilize the mRNA after the excision of the 5’ UTR intron, a consequence of which is that the splice junction (Guthrie & Patterson, 1988) is located on the 3’ side of the stem–loop structure.

The combination of the mobility shift and u.v.-crosslinking data has highlighted the region from −21 to −50 of the human preproinsulin mRNA 5’ UTR as a site for sequence-specific interactions. This sequence corresponds to the 5’ side of the stem–loop (Fig. 2). It is possible that such interactions may facilitate the enhanced translation of preproinsulin mRNA. However, before one can comment on the role of such interactions in the

![Figure 9](image_url)
regulation of the translation of preproinsulin mRNA in the β cell, it is necessary to investigate further the tissue distribution of these sequence-specific RNA binding proteins. In addition, further studies may best be performed using isolated islets of Langerhans, which respond strongly to changes in glucose concentration.

ACKNOWLEDGEMENTS

S. W. K. was supported by an MRC research studentship.

REFERENCES


