Volume 222, number 2, 332–336 FEB 05175 October 1987

Tissue-specific binding of a nuclear factor to the insulin gene promoter

Clare E. Sample and Donald F. Steiner

Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, IL 60637 and Howard Hughes Medical Institute, 920 E. 58th St, Chicago, IL 60637, USA

Received 18 August 1987

Using a DNA-binding gel mobility shift assay, a protein-DNA complex was detected using the rat II insulin promoter and a nuclear extract from rat insulinoma cells. This complex was detected using extracts from other insulin-producing cell lines but was not found using nuclear extracts from non-insulin-producing pancreatic cell lines or other cell types. Binding of the tissue-specific protein to the labelled promoter fragment was effectively competed by both rat and human insulin genes but not by the SV40 TATA box. This protein-DNA interaction may be involved in tissue-specific regulation of insulin gene expression.

DNA-protein interaction; DNA binding; Insulin gene; Tissue specificity; (Rat)

1. INTRODUCTION

Eukaryotic promoters and enhancers have been shown to be composed of a variable number of cis-acting elements which serve to regulate appropriately gene expression, e.g. in a tissue-specific manner or in response to signals by hormones or other agents [1]. Such sequences generally include a TATA box which directs RNA polymerase to the transcriptional start site, near upstream regions which regulate the frequency of initiation of transcription, and enhancer elements which can affect gene expression over considerable distances in either orientation. These cis-acting elements appear to be binding sites for nuclear proteins. Some are common to many genes such as the binding region for SP1 [2] or CAAT binding transcription factor [3] while others are specific to a given gene or family of genes such as the octameric sequence found in immunoglobulin genes [4–6]. The actual number and type of these elements combine to efficiently regulate transcription of a particular gene.

Correspondence (present) address: C.E. Sample, Dept of Microbiology and Molecular Genetics, Harvard University, Boston, MA, USA

The insulin gene is expressed in a highly tissue-specific manner. Unlike most species, rats possess two non-allelic insulin genes which are highly homologous within the protein coding sequences; this homology extends approx. 500 bp upstream from the cap site of the messenger RNA [7]. The rat insulin I gene, which has lost one of the two intervening sequences present in the rat insulin II gene, is believed to have arisen via an RNA-mediated transposition event from a transcript which was initiated upstream from the normal initiation site of the rat insulin II gene [7].

In rats, insulin messenger RNA is only present in the β cells of the pancreas [8]. Furthermore, it has been shown that a region of the gene extending approx. 300 bp upstream from the transcriptional initiation site is exposed to DNase I in a transplantable insulinoma but not in other rat tissues [9], possibly reflecting the interaction of this region of the gene with DNA-binding proteins.

Further studies on tissue-specific expression of the rat insulin I gene have shown that DNA can be deleted to within 300 bp of the cap site with no effect on expression assays, whereas deletions to 247 bp abolish activity; for the human insulin gene deletions can be made to within 258 bp but activity

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/87/$3.50

is lost upon deleting sequences to within 168 bp [10]. Further experiments demonstrated that sequences between 103 and 249 bp upstream of the rat insulin I gene are sufficient, in either orientation, to stimulate expression from the herpes simplex virus thymidine kinase (TK) gene promoter in a tissue-specific manner [11]. Other experiments focusing on tissue specificity have involved fusing 660 bp upstream of the rat insulin II cap site to the SV40 large T antigen; large T antigen was found exclusively in the pancreatic β cells of transgenic mice carrying this construct [12]. The upstream tissue-specific enhancer has been further characterized by DNase I footprinting and contains binding sites for both tissue-specific and general cellular factors [13].

To determine whether other sequences contributed to tissue specificity, DNA extending from 113 bp upstream to 51 bp downstream of the rat insulin I gene fused to the CAT gene, which maintained negligible levels of activity alone, were coupled to the MSV enhancer. When transfected into cells, this construct maintained a level of expression in insulin-producing cells 10-fold higher than that obtained in non-insulin-producing cells [11]. This was in contrast to the MSV-tK promoter construct which was equally active in both cell types. This result indicates the presence in this region of another sequence which directs tissue-specific expression of the insulin gene. Here, we demonstrate the tissue-specific binding of protein(s) to this promoter region of the rat insulin II gene.

2. MATERIALS AND METHODS

2.1. *Cells and nuclear extracts*

RINm5F, RIND8 and MSLG2 cells were grown in RPMI and 1640 media supplemented with 10% fetal calf serum, 50 μg/ml gentamicin and 2.5 μg/ml amphotericin B. HepG2 cells were grown in DME plus 10% fetal calf serum plus antibiotics. P388D1 cells were grown in DME plus 20% fetal calf serum, antibiotics and 25 mM Mops, pH 7.0. HIT cells were grown in Ham's F12 nutrient media containing 18% horse serum, 3% fetal calf serum and antibiotics. To prepare nuclear extracts, cells were grown in spinner flasks containing Cytodex 3 microcarriers (Pharmacia). Cells were removed from the microcarriers by mild trypsin treatment. Serum was added to inactivate the trypsin and cells were separated from the microcarriers by filtration through a Cellector apparatus (EC Apparatus). Nuclear extracts were prepared following the method of Dignan et al. [14]. Briefly, cells were collected by centrifugation, resuspended and incubated in hypotonic buffer and lysed by homogenization. The nuclei were collected by centrifugation, washed several times in hypotonic buffer, resuspended in a high-salt buffer (0.42 M NaCl) and placed on a rotating apparatus for 60 min. The suspension was centrifuged and the supernatant fraction dialyzed against low-salt buffer. Higher salt concentrations did not improve extraction of binding proteins as determined by the gel mobility shift assay.

2.2. *DNA binding*

DNA binding was carried out at room temperature for 30 min in a 20 μl reaction volume containing 10 μg poly(dI-dC) (Pharmacia), 20 μg nuclear extract protein, 10000 cpm (~0.5 ng) ^{32}P-labelled insulin promoter fragment (DNA was labelled using T_4 polynucleotide kinase and [γ-^{32}P]ATP (Amersham, 5000 Ci/mmol)), 10 mM Hepes, pH 7.9, 100 mM KCl, 10% glycerol, 0.1 mM EDTA, and 0.25 mM DTT [15]. Where indicated, 0.5 μg competitor DNA was included. Bromphenol blue was added to the reaction mixtures which were loaded onto pre-electrophoresed 4.0% polyacrylamide gels. Electrophoresis was carried out at 10 V/cm in 40 mM Tris-borate, 1 mM EDTA for 2 h [16–18].

3. RESULTS

The binding of proteins which may be involved in regulation of expression to the insulin promoter was examined using the DNA-binding gel mobility shift assay [16–18]. The fragment of the rat insulin II gene used in these studies is shown in fig.1 and encompasses nucleotides −86 to +47, relative to the site of transcription initiation (nucleotide +1). The 3′-end is located just within the first intron located in the untranslated region of the messenger RNA. Upon incubating this labelled DNA fragment with a nuclear extract made from rat insulinoma cells, Rinm5F, the gel mobility shift assay revealed a single major band shifted upward from the free DNA indicating the interaction of

protein(s) with the DNA (fig.2). On longer exposure, several minor bands representing other DNA-protein interactions were also seen (not shown). Increasing the amount of non-specific DNA in the binding reaction did not interfere with binding indicating that the binding reflects a specific interaction with sequences in the rat II insulin promoter.

To determine whether this factor might be a general cellular factor present in all cell types or a tissue-specific protein, nuclear extracts were prepared from a variety of cell lines and examined for the presence of proteins able to bind to the insulin promoter (fig.3). The protein-DNA complex seen with the RINm5F cell extracts was also seen in other insulin-producing cells such as the hamster insulinoma, HIT, cells (lane 7) and a clonally selected rat insulinoma cell line, RIND 8 (lane 3). However, no interaction was seen using nuclear extracts from non-insulin-producing cell types such as a hepatoma cell line (Hep G2) (lane 4) or a macrophage cell line (P388D1) (lane 5). Furthermore, the interaction was not seen with extracts from MSLG2 cells (lane 6), a transplantable insulinoma which does not produce insulin in culture but which has restored insulin production when passaged in rats [19], or with extracts from an RIN5F cell line which had been passaged extensively in culture and no longer produced insulin (not shown).

Since the formation of the protein-DNA complex was not only tissue specific but localized to cells actively expressing insulin, competition experiments were performed to determine whether the binding was specific to the insulin promoter or

Fig.1. The rat insulin II promoter fragment used to study protein binding. Numbers are distances in nucleotides relative to the site of transcription initiation indicated by the arrow.

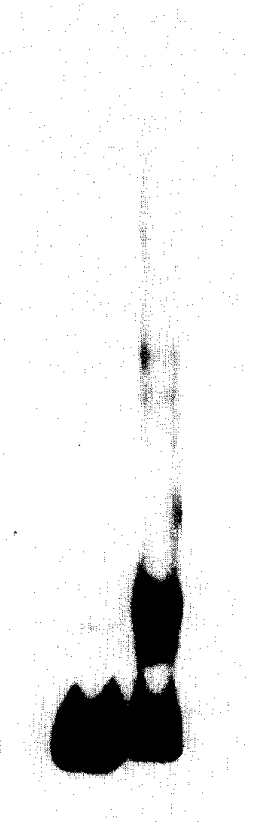

Fig.2. A nuclear protein binds to the insulin gene promoter. The end-labelled fragment shown in fig.1 was incubated in the absence (lane 1) or presence (lane 2) of a nuclear extract from rat insulinoma (Rinm5F) cells as described in section 2.

Fig.3. Binding of a nuclear factor to the insulin gene promoter is tissue-specific. The end-labelled insulin promoter fragment was incubated with nuclear extracts from the following cells: none (lane 1), Rinm5F (lane 2), RinD8 (lane 3), HepG2 (lane 4), P388D1 (lane 5), MSLG2 (lane 6), or HIT (lane 7).

could interact with other promoters (fig.4). For these experiments, a large excess of unlabelled putative competitor DNA was included in the binding reaction. Inclusion of pBR322 DNA had no effect on binding (not shown). The free DNA is shown in lane 1 and control binding with no added competitor is shown in lane 4. When plasmids containing either the rat II insulin promoter (lane 2) or the entire human insulin gene (lane 3) were included in the binding reaction, the unlabelled DNA competed effectively for binding of the protein(s) involved in the complex. The plasmid $pA_{10}CAT_2$ [20] which contains the SV40 TATA box and the 21 bp repeats containing the SP1 binding sites attached to the CAT gene did not compete for binding (lane 5). The plasmid pSV_2neo [21], which contains the SV40 TATA box, 21 bp repeats and enhancer coupled to the neomycin resistance gene, competed slightly (lane 6) but to a much lesser extent than DNA containing the insulin promoter.

4. DISCUSSION

The DNA-binding gel mobility shift assay provides a useful method for analyzing protein-DNA interactions using unfractionated nuclear extracts. By including varying amounts of non-specific DNA in the binding reactions, conditions may be established such that specific protein-DNA complexes are resolved from non-specific interactions. Care must be taken in interpreting the results of footprinting experiments in which these conditions have not been determined. Using this technique, we have demonstrated the presence of a protein(s) able to bind to the rat insulin II promoter in insulin-producing cells. This factor is either not present in non-insulin-producing cell lines or is modified in some way that renders it unable to bind to DNA. The absence of binding not only in other cell types but also in insulinoma cells which have lost the ability to express insulin suggests that this factor may play a role in insulin gene expression. Furthermore, the lack of competition by the SV40 21 bp repeats and TATA box indicates that the factor does not bind to any promoter sequence such as a TATA box binding factor [22]. It may in fact be specific to the insulin gene but this conclusion cannot be made without further experimentation. The fact that the presence of the SV40 enhancer resulted in a slight competition may indicate the presence of a partially homologous sequence or a weak binding site for this factor. In addition to competition by the rat insulin II promoter, the human insulin gene was also an effective competitor. Since 30% of the proximal promoter sequence in the human gene is non-homologous to the rat insulin II gene [7], this may indicate that the binding sequence has been conserved during evolution as would be expected for a sequence with an important role in regulation of insulin gene transcription.

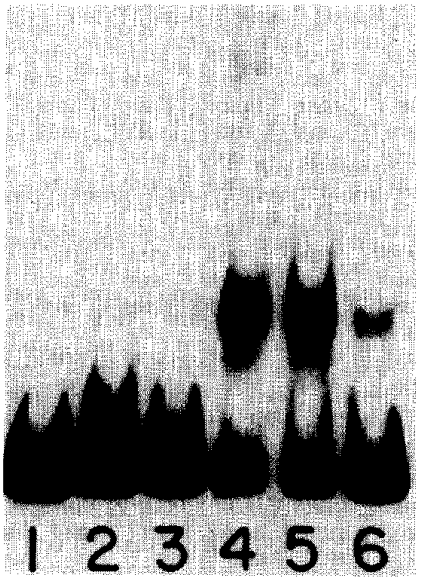

Fig.4. Competition analysis of protein binding by other DNAs. The free DNA probe is shown in lane 1. The binding reactions included the following DNAs as competitors: rat insulin promoter (lane 2), human insulin gene (lane 3), none (lane 4), $pA_{10}CAT_2$ (lane 5), or pSV_2neo (lane 6).

The presence of a sequence within the insulin gene promoter to which proteins bind in a tissue-specific manner is of great interest. It has been demonstrated that insulin gene transcription is controlled by two distinct sequences [11]. The first is a tissue-specific enhancer-like element located between -103 and -249 nucleotides upstream. The second element is located in the promoter region studied in this report and leads to cell-specific expression when coupled to non-cell-specific enhancers. Efficient tissue-specific regulation may be achieved by interaction between these

two sites. This is apparently true for the other tissue-specific genes. The mouse prealbumin gene also contains two distinct tissue-specific control elements; as is the case for the insulin gene, these are a promoter sequence and a far upstream enhancer-like element [23]. The promoter sequence only allows transcription in hepatoma cells even in the presence of a non-specific enhancer, and the prealbumin enhancer stimulates expression from a heterologous promoter only in hepatoma cells. Similarly, immunoglobulin genes possess both a cell-specific enhancer and a cell-specific promoter [5]. The light chain enhancer stimulates its own promoter about 20-fold more than a heterologous one [24]. Similar results are seen for the heavy chain promoter and enhancer, and both enhancers stimulate their own promoters to a greater extent than their effect on heterologous promoters. This may in fact be of more general occurrence but has escaped notice, since many studies on enhancer activity use heterologous promoters and fail to compare the activity achieved with that obtained using the homologous promoter.

An attractive rationale for having two distinct areas governing tissue-specific expression could be the tissue-specific binding of proteins to each of the areas followed by an interaction between these proteins leading to enhanced transcription. Such a situation would dictate a tighter control on tissue-specific expression. It is very likely that such a situation exists for the insulin gene. It has been demonstrated previously that both the enhancer and promoter of the insulin gene act in a tissue-specific manner [11] and, with the results reported here, it has been shown that proteins bind in a tissue-specific manner to both the enhancer [13] and the promoter. Additional experiments to determine the precise sequences involved in the interaction described here, followed by a functional assay, are required before it can be conclusively stated that this protein-DNA interaction participates in the tissue-specific regulation of insulin gene expression.

ACKNOWLEDGEMENTS

We thank Florence Rozenfeld for her assistance in preparing this manuscript and Mary Mathieu for skillful technical assistance. This work was supported by NIH grants DK 13914 and DK 20595.

REFERENCES

[1] Serfling, E., Jasnin, M. and Schaffner, W. (1985) Trends Genet. 1, 224–230.
[2] Dynan, W.S. and Tjian, R. (1985) Nature 316, 774–778.
[3] Jones, K.A., Yamamoto, K.R. and Tjian, R. (1985) Cell 42, 559–572.
[4] Parslow, T.G., Blair, D.L., Murphy, W.J. and Granner, D.K. (1984) Proc. Natl. Acad. Sci. USA 81, 2650–2654.
[5] Bergman, Y., Rice, D., Grosschedl, R. and Baltimore, D. (1984) Proc. Natl. Acad. Sci. USA 81, 7041–7045.
[6] Falkner, F.G. and Zachau, M.G. (1984) Nature 310, 71–74.
[7] Soares, M.B., Schorr, E., Henderson, A., Karathanasis, S.K., Cate, R., Zeitlin, S., Chirgwin, J. and Efstratiadis, A. (1985) Mol. Cell. Biol. 5, 2080–2103.
[8] Giddings, S.J., Chirgwin, J. and Permutt, M.A. (1985) Diabetologia 28, 343–347.
[9] Wu, C. and Gilbert, W. (1981) Proc. Natl. Acad. Sci. USA 78, 1577–1580.
[10] Walker, M.D., Edlund, T., Boulet, A.M. and Rutter, W.J. (1983) Nature 306, 557–561.
[11] Edlund, T., Walker, M.D., Barr, P.J. and Rutter, W.J. (1985) Science 230, 912–916.
[12] Hanahan, D. (1985) Nature 315, 115–122.
[13] Ohlsson, H. and Edlund, T. (1986) Cell 45, 35–44.
[14] Dignam, J.D., Lebowitz, R.M. and Roeder, R.G. (1983) Nucleic Acids Res. 11, 1475–1489.
[15] Singh, H., Sen, R., Baltimore, D. and Sharp, P.A. (1986) Nature 319, 154–158.
[16] Fried, M. and Crothers, D.M. (1981) Nucleic Acids Res. 9, 6505–6525.
[17] Garner, M.M. and Reozin, A. (1981) Nucleic Acids Res. 9, 3047–3060.
[18] Straus, F. and Varshavsky, A. (1984) Cell 37, 889–901.
[19] Madsen, O.D., Larsson, L.I., Rehfeld, J.F., Schwartz, T.W., Lernmark, A., Labrecque, A.D. and Steiner, D.F. (1986) J. Cell Biol. 103, 2025–2034.
[20] Laimins, L.A., Gruss, P., Pozzatti, R. and Khoury, G. (1984) J. Virol. 49, 183–189.
[21] Southern, P.J. and Berg, P. (1981) J. Mol. Appl. Genet. 1, 327–341.
[22] Sawadogo, M. and Roeder, R.G. (1985) Cell 43, 165–173.
[23] Costa, R.H., Lai, E. and Darnell, J.E. (1986) Mol. Cell. Biol. 6, 4697–4708.
[24] Garcia, J.W., Bich-Thury, L.T., Stafford, J. and Queen, C. (1986) Nature 322, 383–385.