The helix-loop-helix transcription factor USF (upstream stimulating factor) binds to a regulatory sequence of the human insulin gene enhancer

Martin L. READ, Andrew R. CLARK and Kevin DOCHERTY*
Department of Medicine, University of Birmingham, Queen Elizabeth Hospital, Birmingham, B15 2TH, U.K.

Two important sequence elements, designated insulin enhancer binding site 1 (IEB1) or NIR and IEB2 or FAR, are involved in regulating expression of the rat insulin I gene. These elements bind a helix-loop-helix transcription factor, insulin enhancer factor 1 (IEF1). The IEB1 site is highly conserved among insulin genes but the IEB2 site is not conserved. To investigate the factors binding at the equivalent IEB1 and IEB2 sites in the human insulin gene enhancer, electrophoretic mobility shift assays were performed using a variety of cell extracts and probes specific for the homologous IEB1 and IEB2 sites. The results indicate that a factor with similar tissue distribution and binding characteristics to those of IEF1 binds to the IEB1 site in the human insulin gene, but that a separate factor, identified as the adenovirus major late transcription factor [MLTF, or upstream stimulating factor (USF)] binds to the IEB2 site.

INTRODUCTION

The regulation of insulin gene expression is dependent on sequences located upstream of the transcription start site (Clark and Docherty, 1992). Two important cis-acting elements, the insulin enhancer binding site 1 (IEB1) or NIR box and the IEB2 or FAR box, have been identified in the rat insulin I gene (Karlsson et al., 1987, 1989). Located at positions −104 (IEB1/NIR) and −233 (IEB2/FAR), these elements share an identical 8 bp sequence, GCCATCTG, which contains a consensus sequence, CANNTG, characteristic of E-box elements (Kingston, 1989). E boxes are present in enhancers from a variety of genes, including immunoglobulin and muscle-specific genes, where they interact with transcription factors containing a helix-loop-helix (HLH) dimerization domain (Murre et al., 1989). The IEB1/NIR and IEB2/FAR boxes of the rat insulin I gene bind a factor, insulin enhancer factor 1 (IEF1), which is present in β cells and other neuroendocrine cells, but absent from non-endocrine cells (Ohlsson et al., 1988; Aronheim et al., 1991). Screening of cDNA expression libraries with probes corresponding to the IEB sequence resulted in the cloning of several factors which were identical to the widely distributed HLH factor E12/47 (Nelson et al., 1990; Shibasaki et al., 1990; Walker et al., 1990). Recent results suggest that IEF1 may represent a heterodimer composed of E47 and a 25 kDa HLH protein, insulin enhancer-specific factor 1 (Park and Walker, 1992).

The IEB1 box is highly conserved among insulin genes, and is thus likely to play an important role in controlling transcription. The IEB2 site is not well conserved; in the rat insulin 2 gene the equivalent sequence is GCCACCCAGGAG, and in the human insulin gene the homologous sequence, which has been previously designated the GC2 box (Boam et al., 1990a), is GCCACCGG.

The aim of the present study was to investigate the factors binding at the IEB1 and IEB2 sites in the human insulin gene enhancer. The results show that a factor with a tissue distribution and binding characteristics similar to those of IEF1 binds to the IEB1 site but not to the IEB2 site. A separate HLH factor present in a wide variety of cells, which was originally described by its ability to bind to and stimulate transcription from the adenovirus major late promoter (MLP) and thus named major late transcription factor (MLTF) or upstream stimulatory factor (USF) (Carthew et al., 1985; Miyamoto et al., 1985; Sawadogo and Roeder, 1985), was shown to bind at the IEB2 site.

EXPERIMENTAL

Chemicals and reagents

Oligonucleotides were purchased from Alta Bioscience, University of Birmingham. An antibody specific for the 43 kDa form of human USF was obtained from Dr. M. Sawadogo, MD Anderson Cancer Center, University of Texas, Houston, TX, U.S.A. Radiochemicals were purchased from Amersham International p.l.c., Slough, Berks., U.K.

Oligonucleotides

Single-stranded complementary oligonucleotides were annealed as previously described (Boam et al., 1990b) and used as double-stranded probes or competitors in electrophoretic mobility shift assays. The following oligonucleotides, corresponding to the indicated sequences in the human insulin gene, were used: oligonucleotide J, which contains the human IEB1 site (underlined) (−122 to −93; 5′ CTCAGCCCCCAGCCATCTG-CCGACCCCCCC 3′); oligonucleotide JR1, which contains the rat insulin I gene IEB1 sequence (−123 to −94; 5′ CAG-CCCCTCTCGCCATCTGCCTACCTACCC 3′); oligonucleotide C (−200 to −171, 5′ TCCTGAGGAAGAGGTGCTG-ACGACCAAGGA 3′); and oligonucleotide D, which contains the human IEB2 site (underlined) (−248 to −219; 5′ ACAGGTCTGGCCACCGGGCCCCTGGTTAAG 3′). Oligonucleotide USF contained the USF binding sequence from the adenovirus MLP (5′ GTAGGCCACGTGACCGGGT 3′).

Abbreviations used: IEB1 and IEB2, insulin enhancer binding sites 1 and 2; MLTF, major late transcription factor; USF, upstream stimulating factor; MLP, major late promoter; HLH, helix-loop-helix; IEF1, insulin enhancer factor 1; IUF1, insulin upstream factor 1.

* To whom correspondence should be addressed.

Electrophoretic mobility shift assays

Electrophoretic mobility shift assays were performed as previously described (Boam et al., 1990b). Nuclear extracts, prepared by the method of Schreiber et al. (1989) (2–4 μg of protein), were incubated with the ^{32}P-labelled probe for 20 min at room temperature in buffer containing 10 mM Tris/HCl, pH 7.5, 50 mM KCl, 5 mM dithiothreitol, 1 mM EDTA and 5% (v/v) glycerol. For competition experiments, the nuclear extract was incubated with 2.5 pmol (approx. 50-fold excess) of unlabelled oligonucleotide for 20 min at room temperature before addition of the probe. In some experiments immune or preimmune serum was added to the nuclear extract and the sample left at room temperature for 20 min before addition of the probe.

RESULTS

To investigate the binding of proteins to the IEB1 and IEB2 sites in the human insulin gene enhancer, electrophoretic mobility shift assays were performed using oligonucleotide probes containing the IEB1 (oligonucleotide J) and IEB2 (oligonucleotide D) sequences, and a variety of cell nuclear extracts. Using oligonucleotide J as a probe and an HIT T15 nuclear extract, a major slow-mobility band was observed (Figure 1a, complex J1). Formation of this complex was abolished by prior incubation of the extract with a 50-fold molar excess of oligonucleotide J, but competition was not observed with oligonucleotide D or with oligonucleotide C. Using a variety of different cell extracts, complex J1 was observed only with extracts from β cell lines such as HIT T15 (Figure 1b) and βTC (results not shown), but was not observed with nuclear extracts from BHK cells, HeLa cells and the mouse pituitary corticotrophic cell line AtT20 (Figure 1b), nor with the rat kidney cell line NRK, the rat liver cell line Φ1 and the human colonic cell line HT29 (results not shown). To investigate the relationship between factor J1 and factors binding at the IEB1 sequence in the rat insulin I gene, an oligonucleotide containing the IEB1 sequence of the rat insulin I gene was used. A single major complex with a similar mobility to that of the J1 complex was observed (Figure 2). Competition for formation of this complex was observed with oligonucleotides JR1 and J, but not with oligonucleotides D or C. These results indicate that a similar factor, probably IEF1, binds at the IEB1 sites in the human and rat insulin I genes.

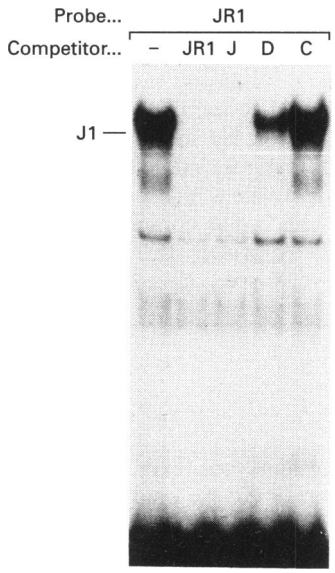

Figure 2 Binding of HIT T15 nuclear extract proteins to oligonucleotide JR1

An electrophoretic mobility shift assay is shown using oligonucleotide JR1 as probe and an HIT T15 nuclear extract. The indicated unlabelled competitor oligonucleotides were used in approximately 50-fold excess. The position of the major complex J1 is indicated.

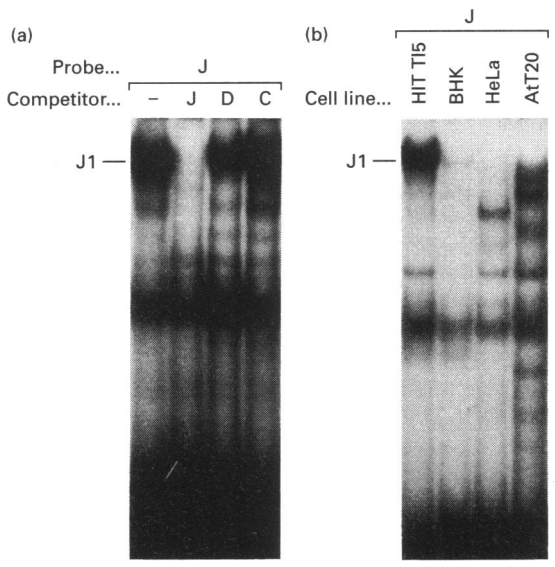

Figure 1 Binding of nuclear extract proteins to oligonucleotide J

(**a**) Electrophoretic mobility shift assay using oligonucleotide J as probe and an HIT T15 nuclear extract. The indicated unlabelled competitor oligonucleotides were used in approximately 50-fold excess. The position of the major complex J1 is indicated. (**b**) Electrophoretic mobility shift assay using oligonucleotide J as a probe and the indicated nuclear extracts.

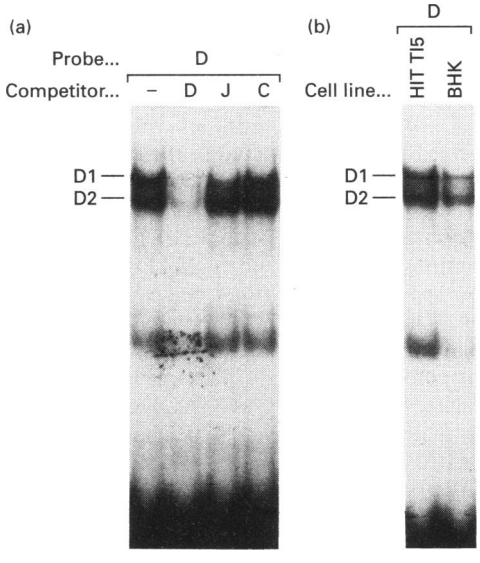

Figure 3 Binding of nuclear extract proteins to oligonucleotide D

(**a**) Electrophoretic mobility shift assay using oligonucleotide D as probe and an HIT T15 nuclear extract. The indicated unlabelled competitor oligonucleotides were used in approximately 50-fold excess. The positions of the two major complexes D1 and D2 are indicated. (**b**) Electrophoretic mobility shift assay using oligonucleotide D as a probe and the indicated nuclear extracts.

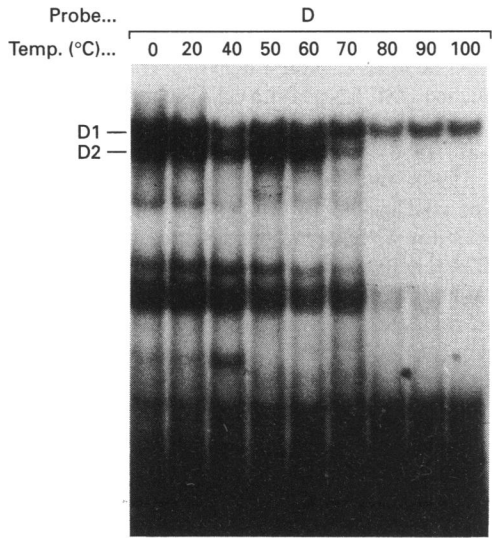

Figure 4 Effect of temperature on the binding of HIT T15 nuclear extract proteins to oligonucleotide D

An electrophoretic mobility assay is shown using oligonucleotide D as probe and an HIT T15 nuclear extract. The nuclear extract (5 μg of protein) was incubated at the indicated temperature for 10 min before addition of the probe. The positions of the major complexes D1 and D2 are shown.

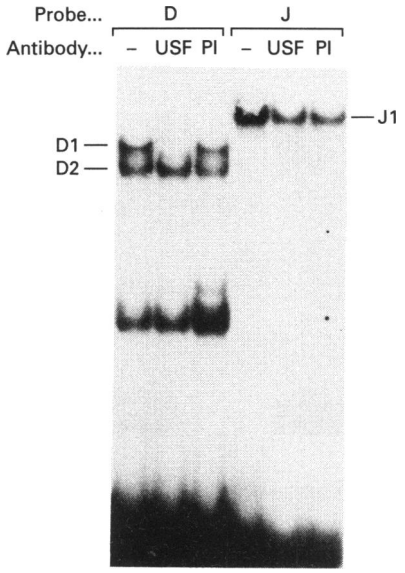

Figure 6 An anti-USF antibody competes for formation of complex D1 but does not affect formation of J1

An electrophoretic mobility shift assay is shown using oligonucleotide D (**a**) or oligonucleotide J (**b**) as probe and an HIT T15 nuclear extract. Nuclear extracts were preincubated for 20 min at room temperature, with no addition (–), 1 μl of anti-USF antibody (USF), or 1 μl of a preimmune antibody (PI). The positions of the major complexes D1, D2 and J1 are shown.

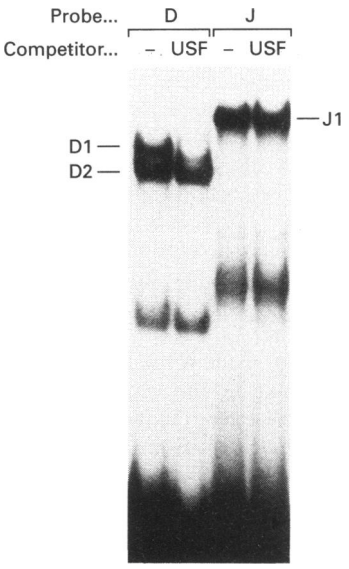

Figure 5 The USF oligonucleotide competes for binding of proteins to oligonucleotide D but not to oligonucleotide J

Oligonucleotide D or J was used in an electrophoretic mobility shift assay with an HIT T15 nuclear extract. The indicated competitor oligonucleotide USF was used in an approximately 50-fold excess.

Two major complexes (D1 and D2) were observed using oligonucleotide D as a probe and an HIT T15 nuclear extract (Figure 3a). Competition for formation of both complexes was observed with oligonucleotide D, but not with oligonucleotides J or C. In addition to HIT T15 cells the factors responsible for the formation of complexes D1 and D2 were present in BHK cells (Figure 3b) and in all other cell lines tested (results not shown). Treatment of the HIT T15 nuclear extract at various temperatures up to 100 °C prior to incubation with labelled oligonucleotide D did not affect the formation of complex D1, while formation of complex D2 was eliminated at temperatures above 60 °C (Figure 4). A candidate for such a widely distributed heat-stable factor was USF.

Confirmation that USF bound at the IEB2 site was obtained using an oligonucleotide containing the USF binding site from the adenovirus MLP. Oligonucleotide USF competed for formation of complex D1 with oligonucleotide D, but did not compete for formation of the tissue-specific complex J1 with oligonucleotide J (Figure 5). Further proof that USF bound at the IEB2 site was provided using an anti-USF antibody. Thus the anti-USF antibody competed for formation of complex D1 using HIT T15 (Figure 6a) and BHK (results not shown) nuclear extracts, while non-immune serum did not compete. The antibody did not compete for formation of complex J1 using oligonucleotide J (Figure 6b). In further experiments it was demonstrated that USF was present in the HIT T15 nuclear extract. Using oligonucleotide USF as a probe, two complexes designated U1 and U2 were observed (Figure 7a). Both complexes were abolished by prior incubation of the nuclear extract with excess unlabelled oligonucleotide USF. Oligonucleotide D competed for formation of complex U2 but not U1, while oligonucleotides J and C did not compete with either complex (Figure 7a). The factors responsible for the formation of complex U1 were β-cell-specific, whereas complex U2 was observed with nuclear extracts from BHK, HeLa, HT29 and AtT20 cells lines (Figure 7b). The anti-USF antibody competed for formation of the U2 complex using the HIT T15 nuclear extract, but did not affect formation of the U1 complex (Figure 7c).

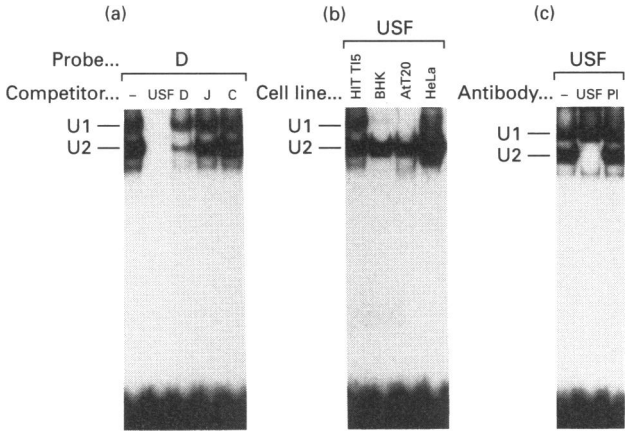

Figure 7 Electrophoretic mobility shift assays using oligonucleotide USF as a probe

(**a**) HIT T15 nuclear extract. The indicated competitor oligonucleotides were used in approximately 50-fold excess. The two complexes U1 and U2 are indicated. (**b**) Tissue distribution of oligonucleotide USF binding proteins. (**c**) HIT T15 nuclear extract. Nuclear extracts were preincubated for 20 min at room temperature, with no addition (—), 1 μl of anti-USF antibody (USF) or 1 μl of a preimmune antibody (PI).

DISCUSSION

This study has demonstrated that an HLH factor with properties similar to those of IEF1 binds to the IEB1 site in the human insulin gene 5′ regulatory region. A separate HLH factor, identified as USF, binds to the IEB2 site. Several lines of evidence support the conclusion that USF binds to the IEB2 site. First, an oligonucleotide containing the USF binding site from the adenovirus MLP efficiently competed for binding at the IEB2 site in oligonucleotide D. Secondly, the heat stability for formation of complex D1 using oligonucleotide D was similar to that for the binding of USF to the adenovirus MLP (Sawadogo and Roeder, 1985). Thirdly, an anti-USF antibody specifically competed for formation of complex D1 using oligonucleotide D. USF has previously been shown to bind along with two other factors, one of which is immunologically related to E12/E47, to a fragment of the rat insulin II gene that contains the IEB1 sequence (Cordle et al., 1991; Shieh and Tsai, 1991). In the present study we found no evidence for USF binding at the IEB1 site, and if it does so it is only with weak affinity.

Complex D2 did not involve binding at the variant E box ('IEB2') sequence, since competition for formation of this complex was observed with oligonucleotide Dm1 (Boam et al., 1990a), which contained a critical mutation within the E box, i.e. CACCGG to AACCGG (results not shown). Factor D2 may be similar to a recently identified factor that binds to the adjacent sequence GGGCCC (Reibel et al., 1993), although D2 appears to have a more widespread tissue distribution.

USF was initially identified by its ability to bind to and activate the adenovirus MLP (Carthew et al., 1985; Miyamoto et al., 1985; Sawadogo and Roeder, 1985). The presence of the factor in a wide variety of uninfected cell types suggested that it might play a role in the activation of cellular genes. USF has been shown to bind to MLP-related sequences in the rat γ fibrinogen promoter (Chodosh et al., 1987), the mouse metallothionein 1 promoter (Carthew et al., 1987), the human growth hormone promoter (Peritz et al., 1988; Lemaigre et al., 1989), the factor IIIa gene from *Xenopus* (Scotto et al., 1989; Kaulen et al., 1991), the liver-type pyruvate kinase promoter (Vaulont et al., 1989), the mouse immunoglobulin λ2 chain promoter (Chang et al., 1992), and the rat class I alcohol dehydrogenase gene (Potter et al., 1991). Purified USF from HeLa cells is composed of 43 and 44 kDa polypeptides that appear to bind independently to DNA (Sawadogo et al., 1988). The 43 kDa protein has been cloned (Gregor et al., 1990), and shown to be a member of the myc-related family of HLH proteins (Kingston, 1989).

In the present study we have not demonstrated that binding of USF to the IEB2 site activates transcription of the gene. However, there are several lines of evidence indicating that the binding site is an important regulatory region. In the rat insulin I gene the equivalent sequence (the FAR box) has been shown to function as a component of a mini-enhancer (the FF mini-enhancer) composed of the FAR box and the adjacent FAR-linked AT-rich region (FLAT) element, which act synergistically to confer tissue-specific transcriptional activity of a heterologous promoter (German et al., 1992a,b). The FAR box binds the HLH factor IEF-1. In the human insulin gene the analogous sequences to the FF mini-enhancer are the IEB2 box and the CT2 box which binds a β-cell-specific factor, IUF-1 (Boam and Docherty, 1989; Scott et al., 1991). It has been proposed that the IEB2 and CT2 boxes also function as a proto- or mini-enhancer (Boam et al., 1990a). Some preliminary data suggest that, as for the FF mini-enhancer, the IEB2 and CT2 boxes individually exhibit weak transcriptional activity, but can act synergistically to produce substantial transcriptional activation of a heterologous promoter (M. L. Read, A. R. Clark and K. Docherty, unpublished work). This is in keeping with the results of studies showing that deletion through this region of the gene has a dramatic effect on transcription (Boam et al., 1990a).

USF exhibits a widespread tissue distribution. It has been shown to interact co-operatively with the TATA box binding factor TFIID to stimulate transcription of the adenovirus MLP (Sawadogo and Roeder, 1985). Like other members of the HLH family of transcription factors it may interact with DNA as a dimer, with dimerization occurring via leucine repeats or HLH domains (Gregor et al., 1990). It is possible that homodimers of the 43 kDa polypeptide, or heterodimers composed of the 43 kDa and the 44 kDa polypeptides, may interact with other constitutive or β-cell-specific HLH or homeodomain proteins to control insulin gene expression.

A comparison can be made between the binding of USF at the IEB2 box of the human insulin gene and at the L4 site in the liver-type pyruvate kinase gene (Vaulant et al., 1989). The binding sequence in both cases differs from the USF binding site in the adenovirus MLP, i.e. CACGTG in the MLP, CACGGG (a half site) in the L4 binding site, and CACCGG in the IEB2 box. Like the insulin gene, the liver-type pyruvate kinase gene is transcriptionally regulated by glucose. The L4 binding site acts as a negative regulatory element in the absence of glucose, but in the presence of glucose/insulin it interacts with an adjacent element, L3, to activate transcription (Bergot et al., 1992). The L3 site probably binds a hepatocyte enriched protein, HNF4, which is a member of the steroid hormone receptor superfamily (Sladek et al., 1990). The FLAT element in the rat insulin I gene binds a number of homeodomain proteins (German et al., 1992b), including the hepatocyte-enriched HNF1α (Emens et al., 1992). The FLAT element interacts with the FAR box within a region of the enhancer which is capable of conferring glucose-responsive transcriptional control of the insulin gene (German et al., 1990). One could speculate that USF binding at the IEB2 site and interaction with a homeodomain protein binding at the CT2 site might also confer glucose-responsive control of insulin gene expression.

M.L.R. was supported by a Wellcome Trust Prize studentship and A.R.C. by a Fellowship from the British Diabetic Association.

REFERENCES

Aronheim, A., Ohlsson, H., Park, C., Edlund, T. and Walker, M. D. (1991) Nucleic Acids Res. **19**, 3893–3899

Bergot, M.-O., Diaz-Guerra, M.-J. M., Puzenat, N., Raymondjean, M. and Kahn, A. (1992) Nucleic Acids Res. **20**, 1871–1878

Boam, D. S. W. and Docherty, K. (1989) Biochem. J. **264**, 233–239

Boam, D. S. W., Clark, A. R. and Docherty, K. (1990a) J. Biol. Chem. **265**, 8285–8296

Boam, D. S. W., Clark, A. R., Shennan, K. I. J. and Docherty, K. (1990b) in Peptide Hormone Secretion (Hutton, J. C. and Siddle, K., eds.), pp. 271–307, Oxford University Press, Oxford

Carthew, R. W., Chodosh, L. A. and Sharp, P. A. (1985) Cell **43**, 439–448

Carthew, R. W., Chodosh, L. A. and Sharp, P. A. (1987) Genes Dev. **1**, 973–980

Chang, L. A., Smith, T., Pognonec, P., Roeder, R. G. and Murialdo, H. (1992) Nucleic Acids Res. **20**, 287–293

Chodosh, L. A., Carthew, R. W., Morgan, J. G., Crabtree, G. R. and Sharp, P. A. (1987) Science **238**, 684–688

Clark, A. R. and Docherty, K. (1992) in Insulin, from Molecular Biology to Pathology (Ashcroft, F. M. and Ashcroft, S. J., eds.), pp. 37–63, Oxford University Press, Oxford

Cordle, S. R., Henderson, E., Masuoka, H., Weil, P. A. and Stein, R. (1991) Mol. Cell. Biol. **11**, 1734–1738

Emens, L. A., Landers, D. W. and Moss, L. G. (1992) Proc. Natl. Acad. Sci. U.S.A. **89**, 7300–7304

German, M. S., Moss, L. G. and Rutter, W. J. (1990) J. Biol. Chem. **265**, 22063–22066

German, M. S., Moss, L. G., Wang, J. and Rutter, W. J. (1992a) Mol. Cell. Biol. **12**, 1777–1788

German, M. S., Wang, J., Chadwick, R. B. and Rutter, W. J. (1992b) Genes Dev. **6**, 2165–2176

Gregor, P. D., Sawadogo, M. and Roeder, R. G. (1990) Genes Dev. **4**, 1730–1740

Karlsson, O., Edlund, T., Moss, L. B., Rutter, W. J. and Walker, M. D. (1987) Proc. Natl. Acad. Sci. U.S.A. **84**, 8819–8823

Received 23 February 1993/6 May 1993; accepted 14 May 1993

Karlsson, O., Walker, M. D., Rutter, W. J. and Edlund, T. (1989) Mol. Cell. Biol. **9**, 823–827

Kaulen, H., Pognonec, P., Gregor, P. D. and Roeder, R. G. (1991) Mol. Cell. Biol. **11**, 412–424

Kingston, R. E. (1989) Curr. Opin. Cell Biol. **1**, 1081–1087

Lemaigre, L. P., Courtois, S. J., Lafontaine, D. A. and Rousseau, G. G. (1989) Eur. J. Biochem. **181**, 555–561

Miyamoto, N. G., Moncollin, V., Egly, J. M. and Chambon, P. (1985) EMBO J. **4**, 3563–3570

Murre, C., McCaw, P. S. and Baltimore, D. (1989) Cell **56**, 777–783

Nelson, C., Shen, L.-P., Meister, A., Fodor, E. and Rutter, W. J. (1990) Genes Dev. **4**, 1035–1043

Ohlsson, H., Karlsson, O. and Edlund, T. (1988) Proc. Natl. Acad. Sci. U.S.A. **85**, 4228–4231

Park, C. W. and Walker, M. D. (1992) J. Biol. Chem. **267**, 15642–15649

Peritz, L. N., Fodor, E. J. B., Silversides, D. W., Cattini, P. A., Baxter, J. D. and Eberhard, N. L. (1988) J. Biol. Chem. **263**, 5005–5007

Potter, J. J., Cheneval, D., Dang, C. V., Resar, L. M. S., Mezey, E. and Yang, V. W. (1991) J. Biol. Chem. **266**, 15457–15463

Reibel, L., Besnard, C., Lores, P., Jami, J. and Gacon, G. (1993) Nucleic Acids Res. **21**, 1595–1600

Sawadogo, M. and Roeder, R. G. (1985) Cell **43**, 165–175

Sawadogo, M., Van Dyke, M. W., Gregor, P. D. and Roeder, R. G. (1988) J. Biol. Chem. **263**, 11985–11993

Schreiber, E., Matthias, P., Müller, M. M. and Schaffner, W. (1989) Nucleic Acids Res. **17**, 6419

Scott, V., Clark, A. R., Hutton, J. C. and Docherty, K. (1991) FEBS Lett. **290**, 27–30

Scotto, K. W., Kaulen, H. and Roeder, R. G. (1989) Genes Dev. **3**, 651–662

Shibasaki, Y., Sakura, H., Takaku, F. and Kasuga, M. (1990) Biochem. Biophys. Res. Commun. **170**, 314–321

Shieh, S.-Y. and Tsai, M.-J. (1991) J. Biol. Chem. **266**, 16708–16714

Sladek, F. M., Zhong, W., Lai, E. and Darnell, J. E., Jr. (1990) Genes Dev. **4**, 2353–2365

Vaulont, S., Puzenat, N., Levrat, F., Cognet, M., Kahn, A. and Raymondjean, M. (1989) J. Mol. Biol. **209**, 205–219

Walker, M. D., Park, C. W., Rosen, A. and Aronheim, A. (1990) Nucleic Acids Res. **18**, 1159–1166