Identification of direct-repeat-binding protein 1 (DRP-1), a DNA-binding protein that binds specifically to the 'malic' enzyme gene promoter direct repeat element

Kevin G. FORD, David P. HORNBY* and Wafaa S. AL HARRASY†

Krebs Institute, Department of Molecular Biology and Biotechnology, University of Sheffield, Sheffield S10 2TN, U.K.

The 'malic' enzyme (ME) gene promoter contains three main regulatory regions. One of these, the direct repeat element (DRE), contains tandem degenerate Sp1-binding sites separated by a 3 bp intervening sequence. We now show that a previously unreported 95 kDa protein, which we have designated DRP-1, binds strongly to the DRE region in a highly specific manner. Western-blot analysis confirms that this protein is not Sp1, which has been shown to bind to similar degenerate sites. Competitive binding assays using purified DRP-1 further reveal that neither

non-specific nor Sp1-consensus-site-containing oligonucleotides can displace those complexes formed between DRP-1 and the DRE sequence, thus confirming sequence-specific binding by this protein. SDS/PAGE analysis of DRE–protein complexes isolated by direct excision and transplantation from retardation gels confirms the presence of the 95 kDa protein and, in addition, suggests that more than one binding site exists for this protein within the DRE. This is in accord with the repeated nature of the DRE DNA sequence which contains two CACC box motifs.

INTRODUCTION

Cytosolic 'malic' enzyme [ME; L-malate–NADP⁺ oxidoreductase (decarboxylating) EC.1.1.1.40] catalyses the NADP⁺-dependent oxidative decarboxylation of malate to pyruvate and CO₂ and further generates NADPH from NADP⁺. ME has long been believed to be the major generator of NADPH for lipogenesis (reviewed in [1]). A number of studies have shown that ME gene expression can be modulated at the transcriptional level by the action of the thyroid hormone tri-iodothyronine [2–6] and at the post-transcriptional level in response to carbohydrate- or fat-related dietary conditions [2,3,7,8].

The expression of ME during development has been widely studied by either direct assay of enzyme activity or immunoreactivity of the protein as a function of developmental status [9–11]. The results obtained from monitoring steady-state levels of rat liver ME mRNA, both during development and in response to hormonal or dietary changes at the suckling/weaning transition, suggest that the changes in rat hepatic ME expression during development are controlled before translation and that diet and hormones are likely to influence expression of ME during development [11,12].

Previous studies have demonstrated the ability of the rat ME gene promoter to direct transcription in HeLa cells and in a wide variety of other tissues. It is therefore likely that those factors necessary for transcription of this gene are present not only in HeLa-cell extracts, but in other tissues also. Characterization of both the rat ME gene and the gene promoter region has been reported [13,14]. The latter was found to resemble the promoter regions of housekeeping genes in that no TATA or CCAAT box motifs could be identified. The rat ME gene promoter has an 83% GC content and contains nine consensus binding sites for Sp1 (5'-CCGCCC-3') distributed widely through the core promoter, intron and untranslated regions. The promoter also contains a tandem perfect repeat of 10 bp, known as the direct

repeat element (DRE). Each 10 bp repeat contains a non-consensus Sp1-binding site, CCACCC, which houses the CACC box. The importance of the CACC box as a transcriptional control element has been demonstrated [5,15–18]. Deletion of this region in the rat ME gene promoter shows that it is vital for efficient transcription; its removal lowers transcriptional efficiency to 17% of maximal promoter activity in an *in vitro* transcription assay [5]. Recent evidence suggests that the general transcription factor Sp1 differentially interacts with CACC boxes on a number of promoters. For example, Sp1 has been shown to bind the CACC boxes of the ϵ -, β - and γ -globin promoters, although surprisingly not all of these lie within recognized non-consensus Sp1-binding sites [19]. Detailed analysis of Sp1 binding to consensus and non-consensus sites reveals that binding to 5'-CCACCC-3' is only 6-fold weaker than binding to the consensus sequence [20], suggesting that Sp1 may play a role *in vivo* in binding to these sequences.

We now report the discovery of a factor present in HeLa-cell nuclear extract, the direct-repeat-binding protein 1 (DRP-1), which binds to the DRE CACC-box-containing sequence *in vitro* in a sequence-specific manner.

MATERIALS AND METHODS

Oligonucleotides

All oligonucleotides used in this study were synthesized on an Applied Biosystems model 476A DNA synthesizer, precipitated with ethanol and used without further purification. Typically, 0.1 μ g of oligonucleotide was incubated with 1 unit of T4 DNA polynucleotide kinase (BRL), 3 μ l of [γ -³²P]ATP (3000 Ci/mmol; Amersham), 0.1 μ g of BSA in a buffer comprising 50 mM Tris/HCl, pH 7.6, 10 mM MgCl₂, 1 mM dithiothreitol, 0.1 mM spermidine and 0.1 mM EDTA in a final volume of 30 μ l. Incubation was for 1 h at 37 °C. After one round of phenol extraction, unincorporated radionucleotides were removed by

Abbreviations used: DRE, direct repeat element; ME, 'malic' enzyme; DRP-1, direct repeat binding protein 1.

* To whom correspondence should be addressed.

† Present address: Royal Oman Police, Directorate General of Criminal Intelligence and Investigation, Oman P.O. Box 446, Muscat code 113.

the spun-column procedure [21] through Sephadex G-25 superfine resin (Pharmacia). Oligonucleotide sequences used in this study are given below. Sp1 and DRE oligonucleotides are derived from regulatory regions of the ME gene promoter. Oligonucleotides were: non-specific, 5'-AGCATTCCGGAA-TGCT-3'; Sp1, 5'-GGCCGCCGCCGCCGC-3'; DRE, 5'-GCCTCGCCACCCCTCTCGCCACCCAC-3'. They were annealed with their complementary sequences (self-annealed for the non-specific oligonucleotide) in equimolar amounts, by heating to 90 °C for 5 min then cooling to room temperature at the rate of 1 °C/min.

Gel-retardation analysis

Typically gel-retardation assay reactions contained 1 ng of end-labelled oligonucleotide, 1 µg of poly(dI-dC) (Pharmacia) and 0.5 µg of protein in a buffer comprising 25 mM Hepes, pH 7.6, and 1 mM EDTA in a final reaction volume of 40 µl. Samples were incubated on ice for 30 min, then subjected to electrophoresis on 4% (19:1) polyacrylamide gels in 0.5 × TBE (1 M Tris, 0.8 M boric acid, 25 mM EDTA) at 100 V for 2–3 h in 0.5 × TBE electrophoresis buffer. After electrophoresis, the gel was transferred to Whatman 3MM filter paper, dried under vacuum and subjected to autoradiography overnight at –70 °C with intensifying screens.

Western-blot analysis

Western blotting was carried out essentially as previously described [22]. Sp1 polyclonal antiserum was a gift from S. Jackson and A. Bannister (Wellcome/CRC Institute, Cambridge, U.K.) and was used in conjunction with the ECL detection system (Amersham) according to the manufacturer's instructions.

Protein purification

HeLa-cell extract (1 ml), prepared as described elsewhere [23], was applied to a 5 ml Affigel-heparin cartridge (Bio-Rad) pre-equilibrated with 5 column vol. of buffer A (10 mM Hepes, pH 7.9, 1 mM EDTA). Fractions were collected over a gradient of 0–1 M NaCl in buffer A and adjusted to 10% glycerol before storage. They were assayed both for their ability to specifically retard DRE probe and by SDS/PAGE analysis. Active fractions were dialysed at 4 °C overnight against 2000 vol. of buffer B (10 mM Hepes, pH 7.9, 1 mM EDTA, 10% glycerol, 1 mM dithiothreitol and 1 mM PMSF). These dialysed fractions (1 ml volumes) were loaded on to a Mono Q HR5/5 anion-exchange column (Pharmacia), pre-equilibrated with buffer A. Fractions were eluted in 1 ml volumes over a linear gradient of 0–1 M NaCl in buffer A. Samples were assayed as detailed above. Active fractions were pooled, dialysed against buffer B and, when required, concentrated by reverse dialysis against solid poly(ethylene glycol) (15–20 kDa).

SDS/PAGE analysis of retarded complex

The binding reaction mixture contained 250 µl of purified DRP-1 (Mono Q fraction), 20 µg of poly(dI-dC), 20 ng of labelled DRE and 10 µg of BSA in gel-retardation buffer in a final volume of 0.5 ml. After non-denaturing PAGE the gel was exposed directly to X-ray film and the retarded bands were identified and excised. The excised bands were then loaded directly into preformed recesses in an SDS/polyacrylamide gel [24] and electrophoresed at constant voltage (200 V) for 3–4 h. After electrophoresis the proteins were visualized by staining with silver stain plus (Bio-Rad).

RESULTS

DRE-binding proteins can be resolved from Sp1 by heparin–Sephacrose chromatography

In an effort to characterize proteins that might bind the DRE sequence, crude HeLa-cell extract was fractionated on a heparin–Sephacrose column as described in the Materials and methods section. The fractionation profile is shown in Figure 1(a). The ability of each fraction to retard the DRE oligonucleotide was examined and the results shown in Figure 1(b). A strong retardation signal is seen for fractions 24–27 (lanes 5–8) similar to that seen using crude extract (lane 1), suggesting that a number of DNA-binding proteins are associated with

Figure 1 Purification of DRP-1

(a) Elution profile of HeLa-cell nuclear extract by heparin–Sephacrose chromatography. ———, A 0–1 M NaCl gradient used to elute the protein fractions. (b) Gel-retardation analysis of relevant fractions derived from heparin–Sephacrose chromatography. All reaction mixtures contained 1 ng of labelled DRE and 25 µg/ml poly(dI-dC). Lanes 3–25 contained 50 µl of each protein fraction, corresponding directly to fractions 22–44 in (a). Lane 1 contained 10 µl of crude HeLa-cell nuclear extract, and lane 2 contained 10 µl of column-flow-through material. (c) Immunoblot analysis of protein fractions derived from heparin–Sephacrose fractionation, using Sp1 antibody. Lanes 1–4, fractions 24–27 in (a); lanes 5–10, fractions 33–38. ns, Non-specific signal.

Figure 2 SDS/PAGE showing purified DRP-1 sample

This sample was used in subsequent experiments. Lane 1, molecular mass; lane 2, 50 μ l of purified DRP-1 after Mono Q ion-exchange chromatography.

these fractions. Weaker retardation signals are seen for fractions 34–41 (lanes 13–22), including retarded matter in the wells of the gel. In the light of the evidence for Sp1 binding to degenerate sites and the CACC box sequence in general, an immunoblotting experiment was performed on active fractions using Sp1 polyclonal antisera. The main concentration of Sp1 detected by the antibody (both isoforms) was centred around fractions 33–38 (lanes 5–10, Figure 1c). There was no evidence of Sp1 being present in fractions 24–27 (lanes 1–4). An additional signal, seen in all fractions examined, results from non-specific interaction and is indicated in Figure 1(c). Those fractions that specifically retarded the DRE probe but did not contain Sp1 protein as judged by Western analysis were pooled, dialysed and applied to a Mono Q ion-exchange column. Gel-retardation analysis of eluted fractions revealed specific binding of the DRE by fractions eluted at around 0.3 M NaCl concentration (results not shown). SDS/PAGE analysis of these pooled active fractions revealed a major band corresponding to a protein with an approximate molecular mass of 95 kDa (Figure 2). A minor band at 50 kDa was also present at very low level.

Purified DRP-1 binds specifically to the DRE oligonucleotide in the presence of non-specific and Sp1-site-containing competitor DNA

The purified DRP-1 sample was subsequently used in gel-retardation experiments designed to demonstrate fully its sequence specificity. Binding of the DRP-1 fraction to the DRE probe results in the formation of three main retarded species, designated A, B and C in Figure 3. Gel-retardation analysis of DRP-1 bound to the DRE in the presence of non-specific (lanes 1–3) or Sp1-site-containing (lanes 4–6) competitor oligonucleotides is shown in Figure 3(a). None of the retarded complexes are abolished even in the presence of a 500-fold molar excess of competitor DNA. The results of a competitive titration of the DRE–DRP-1 complex with the DRE oligonucleotide is shown in

Figure 3 Competitive binding assay of DRE oligonucleotide

(a) All lanes contained 1 ng of labelled DRE oligonucleotide, 25 μ g/ml poly(dI-dC), 30 μ l (0.4 μ g) of purified protein in addition to non-specific competitor oligonucleotide at 100, 300 and 500 ng (lanes 1–3 respectively), or Sp1 competitor oligonucleotide at 100, 300 and 500 ng (lanes 4–6 respectively). (b) All lanes contained 1 ng of labelled DRE oligonucleotide, 30 μ l (0.4 μ g) of purified protein, 25 μ g/ml poly(dI-dC), 0.5 μ g of BSA in addition to 50, 100, 150, 200, 250, 300, 350, 400 and 500 ng (lanes 2–10 respectively) of competitor DRE oligonucleotide.

Figure 3(b). The two upper bands, A and B, are abolished by a 50- and 100-fold molar excess of the DRE oligonucleotide respectively, confirming sequence-specific binding for complex-formation. Complex C is not abolished until the competing oligonucleotide is at a molar excess of 500-fold.

Involvement of DRP-1 in formation of complex A and B

In order to clarify the role of DRP-1 in the formation of the multiple complexes seen in retardation assays, a scaled-up retardation experiment using 250 μ l of purified protein (1 μ g) and 100 ng of DRE oligonucleotide was performed (results not shown). Retarded complexes (corresponding to bands A, B and C in Figure 3) were excised from the gel and subjected to SDS/PAGE, followed by silver staining. The results are shown in Figure 4. Complexes A and B, which are the most sensitive to specific competition by the DRE oligonucleotide, both contain

Figure 4 SDS/PAGE analysis of retarded complexes

Purified DRP-1 was used in a large-scale gel-shift experiment. Complexes designated A, B and C as in Figure 3 were directly excised from the retardation gel and transplanted into the wells of the protein gel. Lane 1, complex A; lane 2, complex B; lane 3, complex C; lane 4, purified sample used to generate the retarded complexes. The gel was silver stained. The 95 kDa protein DRP-1 is arrowed; molecular-mass marker sizes (kDa) are given on the right.

the 95 kDa protein. Owing to the nature of this experiment, protein bands are not highly resolved and there is a certain amount of unavoidable smearing, most evident in lane 2, which also has the highest protein content. The relative quantities of protein associated with complexes A and B reflect the intensity of their respective retardation signals. SDS/PAGE analysis failed to reveal any defined protein bands for complex C (lane 3).

DISCUSSION

We have presented evidence for the existence of a 95 kDa protein that binds specifically to the DRE of the ME gene promoter. Given that Sp1 also exists as a 95 kDa protein and moreover has been shown to bind the CACC motif that exists within the DRE sequence, it is important to demonstrate that DRP-1 and Sp1 are not the same protein. Sp1 purified from HeLa-cell nuclei is consistently isolated as a mixture of two forms, a 105 kDa and a 95 kDa polypeptide in approximately equal amounts, reflecting differing states of phosphorylation and/or glycosylation [25]. These two forms were detected in our experiments using Sp1 antibody, but not in DRP-1-containing fractions, confirming that DRP-1 and Sp1 are distinct proteins. Moreover, DRP-1 is only purified as a single 95 kDa species, as demonstrated in Figure 2. The inability of Sp1-site-containing oligonucleotides to compete effectively against DRP-1-DRE complex-formation in retardation experiments (compared with DRE competitor oligonucleotides) confirms that Sp1 and DRP-1 are distinct proteins.

Incubation of DRE probe with partially purified DRP-1, derived from a single heparin-Sepharose chromatographic step, gives rise to a single retarded species in gel-retardation assays, identical in mobility with that obtained using crude extract (Figure 1), and reflects the presence of additional DNA-binding proteins in this partially purified sample. Retardation patterns obtained when a more purified DRP-1-containing sample is complexed with DRE probe reveal three distinct retarded complexes designated A, B and C in Figure 3. Analysis of the protein content of complexes A, B and C by SDS/PAGE of gel slices containing the retarded complexes reveals that the 95 kDa DRP-1 species is present in complexes A and B, but not in C. This suggests that complex A, the complex of least mobility, contains more than one DRP-1 protein bound to the DRE probe. This concept is in line with the repeat nature of the DRE sequence, which offers the opportunity for two DNA-binding events in tandem. Footprinting studies on the ME gene promoter using rat liver nuclear extracts show that both copies of the repeat sequence in the DRE are bound by protein, separated by a dominant DNase I-hypersensitive site corresponding to the 3 bp region that separates the direct repeat sequences [5]. The presence of complex A in gel-retardation experiments may thus be indicative of both target sites being accessible to the protein, i.e. internally located, not situated at the extreme ends of the sequence. Whether the CACC boxes fulfil this positional requirement in the DRE oligonucleotide is not readily apparent. The identification of the CACC box as a major transcriptional element does, however, suggest itself as a candidate for DRP-1

binding. The identity of the protein(s) responsible for complex C formation is at this time unknown. The inability to detect any protein may result from its migration off the end of the gel (< 15–20 kDa), a natural low abundance or imprecise excision of the retarded complex from the retardation gel. The relatively high electrophoretic mobility of the retarded complex would suggest that the protein is of low molecular mass however. The identity of this protein is currently under investigation.

The inability of non-specific or Sp1-site-containing DNA duplexes to abolish DRE complex-formation with purified DRP-1 confirms the sequence specificity of DRP-1. The precise role of this protein and the DRE sequence motif in the transcriptional regulation of the ME gene has yet to be determined. *In vitro* transcription experiments are currently under way which should help to define the effect of Sp1 and/or DRP-1 binding to the DRE on gene transcription.

We are grateful for the support of both the Wellcome Trust and the YCRC (to K.F.). W.A.-H. was the recipient of a studentship from the Royal Oman Police. The Krebs Institute is an SERC Centre for Molecular Recognition.

REFERENCES

- 1 Frenkel, R. (1975) *Curr. Top. Cell. Regul.* **9**, 157–181
- 2 Tepperman, H. M. and Tepperman, J. (1964) *Am. J. Physiol.* **206**, 357–361
- 3 Towle, H. C., Mariash, C. N. and Oppenheimer, J. H. (1980) *Biochemistry* **19**, 579–585
- 4 Dozin, B., Magnuson, M. A. and Nikodem, V. M. (1985) *Biochemistry* **24**, 5581–5586
- 5 Petty, K. J., Morioka, H., Mitsuhashi, T. and Nikodem, V. M. (1989) *J. Biol. Chem.* **264**, 11483–11490
- 6 Petty, K. J., Desvergne, B., Mitsuhashi, T. and Nikodem, V. M. (1990) *J. Biol. Chem.* **265**, 7395–7400
- 7 Katsurada, A., Iritani, N., Fukuda, H., Noguchi, T. and Tanaka, T. (1987) *Eur. J. Biochem.* **168**, 487–491
- 8 Dozin, B., Magnuson, M. A. and Nikodem, V. M. (1986) *J. Biol. Chem.* **261**, 10290–10292
- 9 Ballard, F. J. and Hanson, R. W. (1967) *Biochem. J.* **102**, 952–958
- 10 Taylor, C. B., Bailey, E. and Bartley, W. (1967) *Biochem. J.* **105**, 717–722
- 11 Barton, C. H. and Bailey, E. (1987) *J. Dev. Physiol.* **9**, 215–224
- 12 Mann, D. J. and Bailey, E. (1991) *Biochem. J.* **279**, 407–412
- 13 Morioka, H., Magnuson, M. A., Mitsuhashi, T., Song, M.-K., Rall, J. E. and Nikodem, V. M. (1989) *Biochemistry* **28**, 4912–4916
- 14 Morioka, H., Tennyson, G. E. and Nikodem, V. M. (1988) *Mol. Cell. Biol.* **8**, 3542–3545
- 15 Schule, R., Muller, M., Otsuka-Murakami, H. and Renkawitz, R. (1988) *Nature (London)* **332**, 87–90
- 16 Mantovani, R., Magaretti, N., Nicolis, S. et al. (1988) *Nucleic Acids Res.* **16**, 4299–4312
- 17 deBoer, E., Antoniou, M., Mignotte, V., Wall, L. and Grosveld, F. (1988) *EMBO J.* **17**, 4203–4212
- 18 Motamed, K., Bastiani, C., Zhang, Q., Bailey, A. and Shen, C.-K. (1993) *Gene* **123**, 235–240
- 19 Yu, C.-Y., Motamed, K., Chen, J., Bailey, A. and Shen, C.-K. (1991) *J. Biol. Chem.* **266**, 8907–8915
- 20 Letovsky, J. and Dynan, W. S. (1989) *Nucleic Acids Res.* **17**, 2639–2653
- 21 Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
- 22 Ford, K., Taylor, C., Connolly, B. and Hornby, D. (1993) *J. Mol. Biol.* **230**, 779–786
- 23 Dignam, J. D., Lebovitz, R. M. and Roeder, R. G. (1983) *Nucleic Acids Res.* **11**, 1475–1489
- 24 Laemmli, U. K. (1970) *Nature (London)* **227**, 680–685
- 25 Briggs, M. R., Kadosh, J. T., Bell, S. P. and Tjian, R. (1986) *Science* **234**, 47–52