Heterogeneous nuclear ribonucleoprotein D0 contains transactivator and DNA-binding domains

Mate TOLNAY*†‡, Lajos BARANYI‡ and George C. TSOKOS*†

*Department of Cellular Injury, Walter Reed Army Institute of Research, Silver Spring, MD 20910, U.S.A., †Department of Medicine, Uniformed Services University of the Health Sciences, Bethesda, MD 20814, U.S.A., and ‡Department of Biochemistry, Walter Reed Army Institute of Research, Silver Spring, MD 20910, U.S.A.

Heterogeneous nuclear ribonucleoprotein D0 (hnRNP D0) is an abundant, ubiquitous protein that binds RNA and DNA sequences specifically, and has been implicated in the transcriptional regulation of the human complement receptor 2 gene. We found that *in vivo* expression of hnRNP D0–GAL4 fusion proteins increased the transcriptional activity of a GAL4-driven reporter gene, providing direct proof that hnRNP D0 possesses a transactivator domain. We found, using truncated hnRNP D0 proteins fused to GAL4, that 29 amino acids in the N-terminal region are critical for transactivation. We established, using a series of recombinant truncated hnRNP D0 proteins, that the tandem RNA-binding domains alone were not able to bind double-stranded DNA. Nevertheless, 24 additional amino acids of the C-terminus imparted sequence-specific DNA binding.

Experiments using peptide-specific antisera supported the importance of the 24-amino-acid region in DNA binding, and suggested the involvement of the 19-amino-acid alternative insert which is present in isoforms B and D. The N-terminus had an inhibitory effect on binding of hnRNP D0 to single-stranded, but not to double-stranded, DNA. Although both recombinant hnRNP D0B and D0D bound DNA, only the B isoform recognized DNA *in vivo*. We propose that the B isoform of hnRNP D0 functions in the nucleus as a DNA-binding transactivator and has distinct transactivator and DNA-binding domains.

Key words: regulation of transcription, transactivation, transcription factor.

INTRODUCTION

Heterogeneous nuclear ribonucleoprotein (hnRNP) D0/AUF1 proteins are present in most, if not all, cell types and are encoded by a single gene [1,2]. Four isoforms are generated by alternative splicing of exon 2, which encodes 19 amino acids, and exon 7, which encodes 49 amino acids [1]. The human and murine proteins are 97% identical, suggesting an evolutionarily conserved function for hnRNP D0 [2]. hnRNP D0 has been implicated in the regulation of mRNA stability by recognizing AU-rich elements present in the 3' untranslated region of several mRNA species [3,4], and also binds the telomeric repeat and the 3' splice site sequence [5]. The B isoform of hnRNP D0 recognizes a motif in the promoter region of the human complement receptor 2 (CR2) gene [6] and co-precipitates with the TATA-binding protein (TBP), suggesting that it may represent a transcription factor [7]. Recently, hnRNP D0B has been identified as a component of the B-cell-specific DNA-binding factor LR1, which binds to the immunoglobulin heavy chain switch region and also activates transcription [8–11].

hnRNP D0 is a member of the 2 × RBD-Gly protein family, members of which contain two highly conserved RNA-binding domains (RBDs) in tandem and a Gly-rich region in the C-terminus [12]. Other members of this family in humans include E2BP [13], hnRNP C-type protein or DL4 [14], JKTBR [15], hnRNP A1 [16,17], hnRNP A2/B1 proteins [16,18] and hnRNP E0 [12]. hnRNP D0 differs from hnRNP A1 and hnRNP A2/B1 proteins in its unique N-terminal domain, and has a C-terminal domain which shows only minor identity with those of hnRNP A1 and hnRNP A2/B1. The RBD is present in a large number

of RNA-binding proteins, and consists of approx. 90 amino acids. It contains a highly conserved ribonucleoprotein consensus sequence of eight amino acids (RNP-1) and a less well conserved sequence of six amino acids (RNP-2) [19,20]. The RBD consists of a four-stranded antiparallel β sheet with two α helices packed against one face of the sheet [21–23]. The RNP-1 and RNP-2 motifs lie in the middle of the two central β sheets and probably provide general RNA-binding activity. Only one RBD of hnRNP D0 is required for the specific recognition of UUAG RNA sequences, and the sequence preference is modified by the presence of the 19-amino-acid insertion [12]. The binding of the shortest isoform of hnRNP D0, which lacks both alternative exons, to the AU-rich element of c-fos mRNA requires both RBDs, as well as parts of the C- and N-termini for highest affinity [3]. Although these studies have provided some information on the mechanism of RNA binding of hnRNP D0, there is no information on its DNA-binding characteristics.

Besides hnRNP D0, of all hnRNP proteins only E2BP [13], the oestrogen response element-binding protein (ERE-BP) [24] and the CARG-binding factor A (CBF-A) [25] are known to bind double-stranded DNA. Quite remarkably, all three proteins display very high identity with hnRNP D0, and all of them have been suspected to regulate gene expression. E2BP, a hepatoma-cell-specific human protein that has a 237-amino-acid-long region completely identical to hnRNP D0B, was found to activate the second enhancer of the hepatitis B virus. The ERE-BP, which was cloned from owl monkeys, is similar to hnRNP D0, because the N-terminal part, with the exception of four amino acid substitutions, is identical to human hnRNP D0B. ERE-BP contains an extended 52-amino-acid-long C-terminus due to a

Abbreviations used: CAT, chloramphenicol acetyltransferase; CBF-A, CARG-binding factor A; CR2, complement receptor 2; CRE, cAMP response element; EMSA, electrophoretic mobility shift assay; ERE-BP, oestrogen response element-binding protein; hnRNP, heterogeneous nuclear ribonucleoprotein; RBD, RNA-binding domain; TBP, TATA-binding protein.

† To whom correspondence should be addressed: Department of Cellular Injury, Walter Reed Army Institute of Research, 503 Robert Grant Ave., Bldg. 503, Rm. 1A32, Silver Spring, MD 20910-7500, U.S.A. (e-mail mtolnay@hotmail.com).

base insertion at the stop codon corresponding to the human gene. The ERE-BP was found to bind to double-stranded ERE (oestrogen response element) half-sites and to compete with the oestrogen receptor for ERE binding, thus negatively influencing gene expression. CBF-A is a ubiquitous murine protein that binds to double-stranded CArG-box motifs, as well as to single-stranded DNA and RNA, and functions as a transcriptional repressor. As of now, the only hnRNP that has been proved to possess a transactivator domain is hnRNP K, which is involved in the transcription of the *c-myc* gene and preferentially binds single-stranded DNA [26,27].

We have previously cloned the B isoform of hnRNP D0, by virtue of its binding to a human CR2 promoter-defined motif. This motif can independently drive the transcription of a reporter gene and is recognized by hnRNP D0B in a specific manner with high affinity [6]. In addition, hnRNP D0 interacts with the TBP, suggesting that it comes into proximity with the transcription machinery [7]. The experiments reported in the present paper clearly demonstrate that hnRNP D0 has a transactivator domain, and that a 24-amino-acid region located in the C-terminus of the protein mediates its DNA binding.

EXPERIMENTAL

Cloning of GAL4-hnRNP D0 fusion proteins and truncated hnRNP D0B, and purification of recombinant proteins

PCR products corresponding to amino acids 2–305 of human hnRNP D0B, amino acids 2–354 of human hnRNP D0D or truncated forms of hnRNP D0B were cloned directionally into the *Bam*HI and *Kpn*I sites of the pSG424 vector [28] (a gift from Dr Mark Ptashne, Memorial Sloan Kettering Cancer Center, New York, NY, U.S.A.), which contains the coding sequence for the DNA-binding domain of GAL4 (residues 1–147) in a mammalian expression vector. The fusion proteins contained a PQLPGIL linker peptide. Sequences have been verified by DNA sequencing. The GAL4-CAT (chloramphenicol acetyltransferase) reporter plasmid has been generated by cloning a synthetic oligonucleotide containing two (CGGAGGACAGTACTCCG) GAL4 sites into the pBL-CAT8+ plasmid [29]. All plasmids used for transfection were purified on columns from Qiagen (Valencia, CA, U.S.A.).

We used PCR to obtain desired fragments of the coding sequence of hnRNP D0B, and cloned them into the pET-21 vector (Novagen, Madison, WI, U.S.A.), together with a His₆ tag. The following proteins were constructed: hnRNP D0B-(70–306), hnRNP D0B-(1–257), hnRNP D0B-(26–306), hnRNP D0B-(1–281) and hnRNP D0B-(70–257). Sequences were verified by DNA sequencing. Vectors expressing complete hnRNP D0B or hnRNP D0D were obtained from Dr F. Ishikawa (Tokyo Institute of Technology, Yokohama, Japan). Recombinant truncated hnRNP D0B was produced in BL21(DE3)pLysS bacteria (Novagen), and proteins were purified on Ni²⁺ columns (Novagen), as suggested by the manufacturer. Purified proteins were subjected to SDS/PAGE, and the concentration of each protein was estimated by Coomassie Blue staining.

Cell cultures

IM-9 B lymphoblastoid cells (American Type Culture Collection, Rockville, MD, U.S.A.) were cultured in the presence of RPMI 1640 (Gibco BRL, Grand Island, NY, U.S.A.) containing 10% (v/v) fetal bovine serum (Gibco BRL), 2 μ M L-glutamine, 100 units/ml penicillin and 100 μ g/ml streptomycin.

DNA transfection and reporter gene analysis

Plasmid DNA was introduced into IM-9 B lymphoblastoid cells by electroporation, using a Bio-Rad Gene Pulser (Bio-Rad, Hercules, CA, U.S.A.). We used 10 μ g of expression plasmid and 2.5 μ g of reporter plasmid for 5×10^6 cells in 0.25 ml of RPMI 1640 medium containing 10% (v/v) fetal bovine serum, in a 0.4 cm electroporation cuvette at settings of 250 V and 960 μ F (time constants obtained were 38–41). Cells were subsequently washed and cultured for 29 h, then treated with dibutyryl cAMP (Sigma, St. Louis, MO, U.S.A.) and PMA (Sigma) for 16 h, or left untreated. Freeze–thaw lysates were assayed for CAT activity, as described [6].

Electrophoretic-mobility-shift assay (EMSA)

Nuclear proteins from IM-9 B cells were obtained as described [6]. EMSAs were performed as described [7]. Sequences of oligonucleotides (from Gibco BRL) used in EMSAs are the following (only sense strands are listed, except when antisense was used single-stranded): specific oligonucleotide defined by the CR2 promoter, TTCTGTAGAGATTACTTCAAATAAAGA; antisense strand of specific oligonucleotide, TCTT-TATTTTGAAGTAATCTCTACAGAA; 20-base-long specific oligonucleotide, TAGAGATTACTTCAAATAA; mutated oligonucleotide 1 (mutated bases underlined), TAGAGATGACG-TCAAATAA; mutated oligonucleotide 2, TAGAGATTA-GTGCAAATAA; CRE (cAMP response element) oligonucleotide, ATTGCCTGACGTCAGAGAGC. We used 0.1 μ g of poly(dI-dC)·poly(dI-dC) per sample with recombinant proteins, and 0.5 μ g of poly(dI-dC)·poly(dI-dC) per sample with nuclear proteins. In supershift experiments, 1 μ l of anti-peptide serum or serum specific to hnRNP D0 (a gift from Dr Gary Brewer, Bowman Gray School of Medicine, Wake Forest University, Winston-Salem, NC, U.S.A.) was added immediately after addition of labelled oligonucleotide, then incubated for 20 min at 25 °C. All sera had been dialysed extensively using 100000- M_r cut-off filters. Gels were analysed using a PhosphorImager (Molecular Dynamics, Sunnyvale, CA, U.S.A.).

Production of anti-peptide sera

The following hnRNP D0-defined peptides were synthesized by Research Genetics (Huntsville, AL, U.S.A.): peptide 1, DEGHSNSSPRHSE; peptide 2, YQQQQWWSRGGF; peptide 3, GRARGRGDQSG; peptide 4, NSQGYGGYGGY-DY. Cocalico Biologicals (Reamstown, PA, U.S.A.) raised anti-peptide sera in rabbits. ELISA titres of all sera were higher than 100000.

Western blots

Western blots were performed as described [7]. Anti-peptide sera were used at 300-fold dilution, whereas the anti-(hnRNP D0) serum was used at 2000-fold dilution.

RESULTS

hnRNP D0 possesses a transactivator domain

We have reported previously that a DNA motif of the human CR2 promoter is recognized by hnRNP D0 and drives transcription of a reporter gene [6], and that hnRNP D0 interacts

Table 1 hnRNP D0 possesses significant transactivating capacity

IM-9 B cells were co-transfected with vectors expressing GAL4-(1–147)–hnRNP D0B-(2–305), GAL4-(1–147)–hnRNP D0D-(2–354) or GAL4-(1–147) alone, and a CAT reporter plasmid driven by two GAL4 sites. After transfection, cells were cultured for 29 h, and then were treated with 1 mM dibutyl cAMP and 100 ng/ml PMA for 16 h, or left untreated (no stimulus). Acetylated products of [¹⁴C]chloramphenicol were separated on TLC plates. Values represent fold increases relative to GAL4-transfected cells without stimulus, and are means ± S.E.M. *P* values are shown for comparisons with the corresponding GAL4-expressing control (stimulated and non-stimulated samples were considered separately), using an unpaired Student's *t*-test, assuming unequal variances and a two-tailed distribution.

Expressed protein	Stimulus	Expression (fold)	<i>P</i>
GAL4–hnRNP D0B	No	2.81 ± 0.75 (<i>n</i> = 7)	0.001
GAL4–hnRNP D0D	No	3.49 ± 1.01 (<i>n</i> = 3)	0.070
GAL4	No	1.0 ± 0.04 (<i>n</i> = 7)	–
GAL4–hnRNP D0B	Yes	9.22 ± 1.94 (<i>n</i> = 4)	0.005
GAL4–hnRNP D0D	Yes	9.17 ± 0.73 (<i>n</i> = 2)	0.077
GAL4	Yes	2.79 ± 0.57 (<i>n</i> = 4)	–

with the TBP *in vivo* [7]. These observations convinced us to propose that hnRNP D0 regulates transcription. Despite the fact that hnRNP D0 is present at a very high concentration (250–400 nM) in B cell nuclei and displays high affinity for its cognate recognition site [7], the binding of another factor to the CR2 promoter-defined DNA motif can be held responsible for inducing transcription. Therefore we performed experiments to prove directly that hnRNP D0 itself possesses a transactivator domain.

Because we were unable to find a cell type that is negative for hnRNP D0, and since expression of hnRNP D0 over the very high level of endogenous protein was not feasible (results not shown), we cloned a chimaeric protein containing the DNA-binding domain of GAL4 and the whole of hnRNP D0, separated by a flexible linker peptide. Increased transcription of a reporter gene driven by GAL4 sites is due to the transactivator domain of the fusion partner, if it has one, because no mammalian factor recognizes the DNA-binding motif of the yeast GAL4 protein. Since B cells express the B and D isoforms of hnRNP D0 at equally high concentrations [7], we cloned two fusion proteins, containing each of these isoforms, into expression vectors. We co-transfected B cells with the expression vector and a CAT reporter plasmid that was driven by two GAL4 sites, and measured the level of CAT expression after 45 h. We found that overexpression of chimaeric proteins containing hnRNP D0B or hnRNP D0D resulted in markedly increased CAT activity (Table 1, unstimulated samples). These data provide direct proof that both the B and D isoforms of hnRNP D0 possess a transactivator domain. Next, we cloned truncated hnRNP D0B proteins fused to GAL4 in order to define the transactivator domain. Removal of the N-terminal 70 amino acids did not reduce the transactivator potential of hnRNP D0B significantly (Table 2, unstimulated samples). Additional truncation of the C-terminal 49 amino acids [hnRNP D0B-(70–257)] still did not affect the ability of the protein to transactivate. We next considered that the 19-amino-acid-long alternative insert that corresponds to amino acids 79–97, which is present in isoforms B and D of hnRNP D0 but not in other isoforms, may encode the transactivator domain. Indeed, removal of amino acids 70–98 of the protein [hnRNP D0B-(99–306)] reduced its ability to transactivate by more than 70%.

We have shown previously that treatment of B cells with the protein kinase A activator cAMP and the protein kinase C

Table 2 The transactivating domain of hnRNP D0 is defined by amino acids 70–98

Cells were stimulated or not with 1 mM dibutyl cAMP and 100 ng/ml PMA. Experiments were performed as described in the legend to Table 1. Values represent percentage activity compared with the corresponding GAL4–hnRNP D0B-transfected cells (stimulated and non-stimulated samples were considered separately), and are means ± S.E.M. *P* values are shown for comparisons with the corresponding GAL4–hnRNP D0B-expressing control, using an unpaired Student's *t*-test, assuming unequal variances and two-tailed distribution.

Expressed protein	Stimulus	Activity (%)	<i>P</i>
GAL4–hnRNP D0B	No	100 ± 7.2 (<i>n</i> = 4)	–
GAL4–hnRNP D0B-(26–306)	No	88.6 ± 16.9 (<i>n</i> = 4)	0.344
GAL4–hnRNP D0B-(70–306)	No	85.5 ± 38.8 (<i>n</i> = 4)	0.570
GAL4–hnRNP D0B-(99–306)	No	28.4 ± 10.6 (<i>n</i> = 4)	0.0002
GAL4–hnRNP D0B-(70–257)	No	99.6 ± 45.8 (<i>n</i> = 3)	0.991
GAL4	No	0 ± 0.9 (<i>n</i> = 4)	–
GAL4–hnRNP D0B	Yes	100 ± 20.9 (<i>n</i> = 4)	–
GAL4–hnRNP D0B-(26–306)	Yes	111.9 ± 31.7 (<i>n</i> = 4)	0.611
GAL4–hnRNP D0B-(70–306)	Yes	70.8 ± 17.1 (<i>n</i> = 3)	0.148
GAL4–hnRNP D0B-(99–306)	Yes	17.4 ± 23.3 (<i>n</i> = 4)	0.004
GAL4–hnRNP D0B-(70–257)	Yes	120.8 ± 50.2 (<i>n</i> = 3)	0.635
GAL4	Yes	0 ± 2.8 (<i>n</i> = 4)	–

activator PMA increased the transcription of a reporter gene driven by the CR2-defined motif [6]. Therefore we asked whether the transactivating activity of hnRNP D0 is regulated by protein kinases. We treated B cells transfected with plasmids expressing the GAL4–hnRNP D0 fusion proteins with PMA and a cell-permeable cAMP analogue, dibutyl cAMP, and observed that such treatment clearly increased CAT activity over that of unstimulated cells (Table 1). However, GAL4-expressing control cells were stimulated as well. Removal of the transactivator domain of hnRNP D0B [hnRNP D0B-(99–306)] greatly reduced its ability to transactivate in cells stimulated with PMA and dibutyl cAMP (Table 2, stimulated samples).

Amino acids 258–281 of hnRNP D0B are critical for DNA binding

In order to identify regions of hnRNP D0B that are involved in binding to double-stranded DNA, we cloned several truncated hnRNP D0B proteins with parts of the N- and/or C-termini removed. In preliminary experiments we found that the tandem RBD [hnRNP D0B-(70–257)] is not able to bind DNA, and therefore we focused on regions outside the RBD. Furthermore, we did not remove any part of the RBD, because each RBD forms a compactly folded structural unit, and the disturbance of these protein domains could abort DNA binding even if the amino acids removed are not directly involved in DNA binding. The RBD, however, was found to be thermodynamically stable [12,23].

A schematic representation of the truncated hnRNP D0B constructs used in the present study is shown in Figure 1(A). Recombinant truncated hnRNP D0B proteins with a His₆ tag on the C-terminus were produced in bacteria and purified. We compared by EMSA the ability of truncated proteins to bind to double-stranded DNA with that of full recombinant hnRNP D0B by using a synthetic double-stranded oligonucleotide recognized specifically by hnRNP D0B [7]. We found that removal of the N-terminal 25 or 69 amino acids had no effect on DNA binding of hnRNP D0B (Figure 1B). Removal of the C-terminal 25 amino acids partly reduced the ability of the protein to bind

Figure 1 Amino acids 258–281 of hnRNP D0B are critical for binding to double- and single-stranded DNA

(A) Schematic representation of cloned truncated hnRNP D0B showing the removed parts of the N-terminus (N-term.) and/or the C-terminus (C-term.). (B) Binding to double-stranded DNA of purified recombinant truncated hnRNP D0B or full-length hnRNP D0B was tested by EMSA, using 20 fmol of specific double-stranded oligonucleotide as labelled probe. Amounts of recombinant protein (ng) used in the samples are indicated above the lanes. The positions of the shifted bands are indicated on the left. (C) Recombinant truncated hnRNP D0B or full-length hnRNP D0B were pre-incubated with a 50-fold molar excess of unlabelled double-stranded oligonucleotides for 10 min before the addition of 20 fmol of labelled 20 bp double-stranded specific oligonucleotide, and analysed by EMSA. Mutated oligonucleotides 1 and 2 contain two base changes as compared with the specific oligonucleotide. CRE was used as a non-specific oligonucleotide. (D) Binding to single-stranded DNA of purified recombinant truncated hnRNP D0B or full-length hnRNP D0B was tested by EMSA, using 20 fmol of the antisense strand of the specific oligonucleotide as labelled probe. Amounts of recombinant protein (ng) used in the samples are indicated above the lanes. (E) The region found to be critical for DNA binding (amino acids 258–281) is hatched.

DNA. However, we could not detect any DNA binding with a truncated protein lacking the full C-terminus [hnRNP D0B-(1–257)].

These findings suggest that the first half of the C-terminus, i.e. amino acids 258–281, is critical for binding of hnRNP D0B to double-stranded DNA. However, it is possible that additional amino acids are important, or complement the sequence-specific interaction with DNA. Therefore we tested whether regions other than amino acids 258–281 are needed for sequence-specific DNA binding. We used several unlabelled oligonucleotides and compared their ability to block the interaction between labelled specific oligonucleotide and full hnRNP D0B or truncated proteins. We used two mutated oligonucleotides, which had two bases changed and had been shown previously to compete less efficiently than the specific oligonucleotide [7], along with a non-specific oligonucleotide that contained the consensus CRE. We found that the mutated and non-specific oligonucleotides were clearly less efficient than the specific oligonucleotide in competing for DNA binding of the full hnRNP D0B protein, as well as that of all the truncated proteins that could bind DNA (Figure 1C).

These results suggest that no region of the N- or C-terminus, other than amino acids 258–281, is needed for highly sequence-specific recognition of double-stranded DNA.

We have shown previously that hnRNP D0B can also bind single-stranded DNA [7]. Hence we asked whether the same region of hnRNP D0B that is involved in binding double-stranded DNA is also responsible for binding single-stranded DNA. Using the antisense strand of the specific oligonucleotide as a labelled probe, we detected efficient binding of truncated proteins lacking half of the N-terminus [hnRNP D0B-(26–306)], the full N-terminus [hnRNP D0B-(70–306)] or the last 25 amino acids of the C-terminus [hnRNP D0B-(1–281)] (Figure 1D). No binding to single-stranded DNA was detected with hnRNP D0B lacking the full C-terminus [hnRNP D0B-(1–257)], suggesting that, outside the conserved RBD, the only region of hnRNP D0B that is critical for binding to double- or single-stranded DNA maps to amino acids 258–281 (Figure 1E). Interestingly, proteins lacking half of the N-terminus [hnRNP D0B-(26–306)] or the entire N-terminus [hnRNP D0B-(70–306)] could clearly bind single-stranded DNA more efficiently than could the whole

Figure 2 Production of antibodies that are specific for hnRNP D0-derived peptides

(A) Amino acid sequence of the N- and C-termini of hnRNP D0B; selected peptides are underlined. The 19-amino-acid insert and part of the D-isoform-specific 49-amino-acid insert are shown in **bold**. In the presence of the 49-amino-acid insert, amino acid D (Asp) at the insertion site changes to G (Gly). The RNP-2 motif (MFIGGL) of RBD1 is shown in *italics*. Arrows indicate the first or last amino acids of truncated hnRNP D0 proteins (see Figure 1a). (B) A portion of 0.1 μ g of recombinant hnRNP D0B or hnRNP D0D, or 10 μ g of nuclear protein obtained from IM-9 B cells, was subjected to SDS/PAGE, transferred to nitrocellulose and analysed by Western blot analysis using anti-(hnRNP D0) serum, obtained by immunizing with the full protein, or with anti-peptide sera. Anti-P1b refers to peptide 1, rabbit b; and so on. Anti-P1b, anti-P4a and anti-P4b sera did not work in Western blots.

protein, suggesting that the first 25 amino acids of hnRNP D0B have a negative effect on binding to single-stranded DNA.

Mapping of DNA-binding regions of hnRNP D0B using anti-peptide antibodies

The use of truncated proteins to identify functionally important regions of a protein always involves uncertainty about the proper folding and conformation of the truncated protein. Removal of one part of a protein may affect the thermodynamic stability of the protein or alter the conformation of more distant regions. We decided to use synthetic peptides representing selected segments of hnRNP D0B to raise anti-peptide antibodies and to determine whether these can interfere with DNA binding of the full recombinant protein. This approach eliminates concerns about proper folding and conformation of proteins that were truncated without information on their three-dimensional structure. We would like to point out that only peptides with certain physico-chemical characteristics are suitable for synthesis and immunization, and thus we were restricted in our selection of hnRNP D0-derived peptides. The designed peptides are shown in Figure 2(A).

Two rabbits (a and b) were immunized with each peptide, and the specificities of sera were tested in Western blots. Sera anti-P1a (peptide 1, rabbit a), anti-P4a and anti-P4b did not recognize hnRNP D0 proteins in Western blots (results not shown), but were effective in EMSAs, suggesting that they preferentially recognize non-denatured protein. The results obtained with the

Figure 3 Anti-peptide sera interfere with DNA binding of recombinant hnRNP D0 or nuclear protein

A portion of 2 ng of recombinant hnRNP D0B (A) or hnRNP D0D (B), or 0.3 μ g of nuclear protein obtained from IM-9 B cells (C), was mixed with 20 fmol of specific labelled double-stranded oligonucleotide, then 1 μ l of anti-peptide rabbit serum or pre-immune serum was added as indicated, followed by analysis by EMSA. Anti-P1a refers to peptide 1, rabbit a; and so on.

remaining sera and a control serum, which was obtained by immunizing a rabbit with the full protein, are shown in Figure 2(B). The B and D isoforms of hnRNP D0 are the dominant alternatively spliced forms in the B cell nucleus and they are present in similar quantities (see also [7]). As expected, anti-P1b, anti-P2a and anti-P2b sera recognized both the B and D isoforms of hnRNP D0. Anti-P3a and anti-P3b sera, produced against peptide 3, which is discontinuous in the longer D isoform, recognized only hnRNP D0B.

Following the verification of the specificity of the sera, we performed EMSAs to determine the ability of sera to interfere with DNA binding of hnRNP D0B. We found that all sera raised against peptides 1, 2 and 3, defined by hnRNP D0B, diminished the formation of the shifted protein–DNA complex, indicating that they recognized hnRNP D0B (Figure 3A). However, only the addition of anti-P3a and anti-P3b sera, produced against peptide 3, resulted in the appearance of a supershifted antibody–protein–DNA complex, suggesting that binding of an antibody molecule to the epitope corresponding to peptide 3 still allows interaction of hnRNP D0B with DNA, whereas binding of antibody to epitopes corresponding to peptides 1 or 2 interferes with DNA binding of the protein. Similar results were obtained with anti-P1 and anti-P2 sera using recombinant hnRNP D0D (results not shown). Unexpectedly, all pre-immune sera contained a protein other than hnRNP D0 proteins that recognized the

specific oligonucleotide (data are shown for pre-immune serum anti-P4a in Figure 3A, two right lanes). As expected, sera produced against peptide 4 (anti-P4a and anti-P4b), which is present only in the longer D isoform of hnRNP D0, did not interfere with DNA binding of hnRNP D0B. Nevertheless, anti-P4 sera supershifted the hnRNP D0D–DNA complex, indicating the specificity of the sera and suggesting that the region corresponding to peptide 4 is not involved in DNA binding (Figure 3B). Anti-P3 sera did not affect DNA binding of hnRNP D0D (data for serum anti-P3a are shown in Figure 3B), demonstrating that the anti-P3 and anti-P4 sera are specific for the B and D isoforms respectively of hnRNP D0. Using these sera, we tested which isoform of hnRNP D0 recognizes the specific oligonucleotide from total B cell nuclear extracts. Remarkably, we found that the B isoform of hnRNP D0 preferentially recognized the oligonucleotide from B cell nuclear extracts (Figure 3C).

DISCUSSION

hnRNP D0 is an abundant, ubiquitous protein that can bind to RNA as well as to double- and single-stranded DNA sequences in a specific manner. Its role in regulating the half-life of mRNA species containing an AU-rich element is well characterized [3,4]. Although it is more abundant in the nucleus than in the cytoplasm [30], no nuclear function has been assigned to hnRNP D0. We recently demonstrated that hnRNP D0 binds to double-stranded DNA in a sequence-specific manner, and suggested that it regulates gene expression [6,7]. In the present study we have determined that hnRNP D0 possesses transactivating activity that is linked to an N-terminal region, and we have identified a unique C-terminal region in the protein that is critical for DNA binding.

Fusion protein constructs containing the DNA-binding domain of GAL4 and the B or D isoforms of hnRNP D0 increased the transcriptional activity of a CAT reporter gene driven by GAL4 recognition sites. Because no mammalian protein binds to GAL4 recognition sites, determination of the potential transactivating activity of the fused partner is possible. This experimental system has been used extensively to determine directly whether a protein can influence transcription [26,28,31]. We found that hnRNP D0 was able to transactivate the GAL4-driven reporter 3–4-fold. The fold increase in transactivation detected in our experiments is similar to that observed with E2BP [13] and hnRNP K [26], the other two known hnRNPs with positive transactivator potential. We fused truncated forms of hnRNP D0B to GAL4 in order to locate the transactivator domain. We found that the tandem RBD [hnRNP D0B-(70–257)] still possessed full transactivating ability. However, removal of the N-terminal 29 amino acids of RBD1 [hnRNP D0B-(99–306)] reduced the transactivating capacity to 28% of that of the full protein. The region that we found to be critical for transactivation includes the 19-amino-acid alternative insert present in isoforms B and D. It is attractive to consider this alternative insert as a modular unit that lends the ability to transactivate to those isoforms of hnRNP D0 that contain it. The transactivating ability of hnRNP D0 is likely to depend on its reported interaction with the TBP [7]. The TBP is the DNA-binding subunit of the TFIID multiprotein complex, which recruits RNA polymerase II and initiates transcription. Several transcription factors which have an RBD, among them hnRNP K [32] and hTAF_{II}68 [33], have been shown to interact with the TBP.

We had shown previously that protein kinase A and protein kinase C activators increased the transcription of a reporter gene driven by the CR2-defined motif to which hnRNP D0 binds specifically [6]. It is unclear, however, if phosphorylation by

protein kinase A and protein kinase C influences the transactivating ability of hnRNP D0. Although stimulation of these signalling pathways markedly increased reporter gene expression in cells expressing GAL4–hnRNP D0 (Table 1), stimulation increased reporter gene expression of control samples as well. Protein kinase C, protein kinase A and casein kinase II phosphorylate hnRNP A1 *in vitro*, and it has been shown that its phosphorylation state affects its ability to interact with RNA [34,35]. hnRNP A2 is also phosphorylated by casein kinase II [36]. Finally, the DNA-binding activity of LR1, a nucleolin–hnRNP D0B heterodimer present in B cells, is phosphorylation-dependent [11].

The present study is the first to dissect the features of binding to double-stranded DNA of any hnRNP. We found, using bacterially expressed truncated proteins, that the tandem RBDs of hnRNP D0B (amino acids 70–257) do not bind double-stranded DNA in the absence of an additional 24-amino-acid region (amino acids 258–281) of the C-terminus adjacent to the RBDs. This segment of the protein is sufficient to impart DNA binding that is highly sequence-specific. The N-terminal 69 amino acids of hnRNP D0B are not involved in binding to double-stranded DNA. In most studies investigating the RNA binding of hnRNP D0B [12,23] and other homologous proteins [22,37], one RBD was found to be sufficient for sequence-specific binding. Two independent rabbit sera produced against a peptide comprising residues 263–275, located at the centre of the 24-amino-acid region, blocked the DNA binding of hnRNP D0B, supporting data obtained with truncated proteins (Figure 1). Sera produced against amino acids 277–289 did not interfere with DNA binding, further refining the boundaries of the DNA-binding domain. The D isoform of hnRNP D0, which contains a 49-amino-acid alternative insert after amino acid 284, could bind DNA efficiently, indicating that the 49-amino-acid insert does not interfere with DNA binding. Supporting that observation, sera produced against the 49-amino-acid insert did not block DNA binding. We also produced sera against the 19-amino-acid alternative insert that was found to alter the sequence specificity of hnRNP D0 for RNA [1] and which is necessary for DNA binding of LR1 [8]. We found these sera to interfere with the DNA binding of hnRNP D0B, suggesting that the 19-amino-acid alternative insert plays a role in DNA binding, possibly due to its close proximity to the conserved RNP-2 motif. Notably, we found the 19-amino-acid insert to be involved in transactivation as well, raising the possibility that it may link DNA binding to transactivation.

A homology search of GenBank using the 24-amino-acid region (amino acids 258–281) that we identified to be critical in DNA binding found a number of proteins. E2BP [13] and the ERE-BP [24] are completely identical, whereas CBF-A [25] is 58% identical, with hnRNP D0B in the 24-amino-acid region. Notably, these three proteins are the only hnRNP proteins, besides hnRNP D0, that have been reported to bind specifically to double-stranded DNA sequences. The rat hnRNP C-type protein [38], which has been reported to bind double-stranded DNA non-specifically, also displays complete identity with hnRNP D0B in the 24-amino-acid region. Notably, differences between the human and murine hnRNP D0 proteins, with the exception of two amino acid substitutions in the 19-amino-acid insert, are restricted to the N-terminal region that we found not to be involved in binding to double-stranded DNA. In conclusion, the 24-amino-acid region is unique to a small group of hnRNPs that have been reported to bind double-stranded DNA; the 19-amino-acid insertion is present in all of these proteins.

We have shown previously that hnRNP D0 can bind the sense and antisense strands of its cognate recognition motif [7]. Thus

we investigated which regions of the protein are involved in binding to single-stranded DNA. Using truncated recombinant proteins, we found that binding to single-stranded DNA requires the same protein domains as binding to double-stranded DNA. Interestingly, however, the N-terminal 25 amino acids exerted a negative effect on binding to single-stranded DNA, but had no effect on binding to double-stranded DNA. Proteins interacting with the N-terminus of hnRNP D0 may mask this negative domain, thus increasing the affinity of hnRNP D0 for single-stranded DNA. We have reported that recombinant hnRNP D0 displayed a higher affinity for the sense and antisense strands of its cognate recognition motif than for double-stranded DNA, and predicted, on a theoretical basis, that this will result in melting of the double helix [7]. Binding of a regulatory protein to the N-terminus of hnRNP D0 would result in facilitated melting of the double helix, which in turn may contribute to the creation of a flexible single-stranded DNA region, in a fashion similar to that proposed for hnRNP K [39].

From the results of experiments using isoform-specific sera, we conclude that only the B isoform of hnRNP D0, and not the D isoform, recognizes DNA *in vivo*. However, both isoforms are present in the B cell nucleus, as determined by Western blot analysis (Figure 2 and [7]), and in recombinant form the B and D isoforms bind to oligonucleotides with similar affinity [7]. We speculate that the D isoform is modified secondarily, possibly on its extra 49-amino-acid insert that is very rich in tyrosine residues, and that the modified form cannot bind DNA. The 49-amino-acid insert is immediately adjacent to the 24-amino-acid region that we found to be critical in DNA binding. It is possible that the different isoforms of hnRNP D0 participate in separate cellular processes. Primarily, the isoforms lacking the 19-amino-acid insert were shown to regulate mRNA stability *in vivo* [40], whereas we demonstrated the involvement of the 19-amino-acid insert in transactivation. We propose that the B isoform, which has the 19-amino-acid insert involved in DNA binding and transactivation but lacks the 49-amino-acid insert, participates in transcriptional regulation.

hnRNP D0 has been demonstrated to interact with nucleolin [8], TBP [7], α CP1 [4] and heat-shock protein 70 [41]. Considering the multiple protein and nucleic acid contacts of hnRNP D0, together with its abundance, hnRNP D0 may play a central role in balancing transcription, RNA stability and possibly other cellular functions.

We gratefully acknowledge Dr Mark Ptashne for providing the pSG424 vector, Dr Fuyuki Ishikawa and Dr Yasuko Kajita for the hnRNP D0 expression vector, and Dr Gary Brewer for the hnRNP D0-specific antibody. We thank Dr David Levens for helpful discussions, Dr Lyudmila Vereshchagina for her suggestions and Dr Alan Hubbs for his help with sequencing. This work was supported by National Institutes of Health Grant AI42782. L.B. is a fellow of the National Research Council.

REFERENCES

- Wagner, B. J., DeMaria, C. T., Sun, Y., Wilson, G. M. and Brewer, G. (1998) Structure and genomic organization of the human AUF1 gene: alternative pre-mRNA splicing generates four protein isoforms. *Genomics* **48**, 195–202
- Dempsey, L. A., Li, M., DePace, A., Bray-Ward, P. and Maizels, N. (1998) The human HNRPD locus maps to 4q21 and encodes a highly conserved protein. *Genomics* **49**, 378–384
- DeMaria, C. T., Sun, Y., Long, L., Wagner, B. J. and Brewer, G. (1997) Structural determinants in AUF1 required for high affinity binding to A+U-rich elements. *J. Biol. Chem.* **272**, 27635–27643
- Kiledjian, M., DeMaria, C. T., Brewer, G. and Novick, K. (1997) Identification of AUF1 (heterogeneous nuclear ribonucleoprotein D) as a component of the alpha-globin mRNA stability complex. *Mol. Cell. Biol.* **17**, 4870–4876
- Ishikawa, F., Matunis, M. J., Dreyfuss, G. and Cech, T. R. (1993) Nuclear proteins that bind the pre-mRNA 3' splice site sequence r(UUAG/G) and the human telomeric DNA sequence d(TTAGGG)n. *Mol. Cell Biol.* **13**, 4301–4310
- Tolnay, M., Lambris, J. D. and Tsokos, G. C. (1997) Transcriptional regulation of the complement receptor 2 gene: role of a heterogeneous nuclear ribonucleoprotein. *J. Immunol.* **159**, 5492–5501
- Tolnay, M., Vereshchagina, L. A. and Tsokos, G. C. (1999) Heterogeneous nuclear ribonucleoprotein D0B is a sequence-specific DNA-binding protein. *Biochem. J.* **338**, 417–425
- Dempsey, L. A., Hanakahi, L. A. and Maizels, N. (1998) A specific isoform of hnRNP D interacts with DNA in the LR1 heterodimer: canonical RNA binding motifs in a sequence-specific duplex DNA binding protein. *J. Biol. Chem.* **273**, 29224–29229
- Brys, A. and Maizels, N. (1994) LR1 regulates c-myc transcription in B-cell lymphomas. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 4915–4919
- Bullone-Paus, S., Dempsey, L. A. and Maizels, N. (1995) Host factors LR1 and Sp1 regulate the Fp promoter of Epstein-Barr virus. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 8293–8297
- Williams, M. and Maizels, N. (1991) LR1, a lipopolysaccharide-responsive factor with binding sites in the immunoglobulin switch regions and heavy-chain enhancer. *Genes Dev.* **5**, 2353–2361
- Kajita, Y., Nakayama, J., Aizawa, M. and Ishikawa, F. (1995) The UUAG-specific RNA binding protein, heterogeneous nuclear ribonucleoprotein D0. Common modular structure and binding properties of the 2xRBD-Gly family. *J. Biol. Chem.* **270**, 22167–22175
- Tay, N., Chan, S. H. and Ren, E. C. (1992) Identification and cloning of a novel heterogeneous nuclear ribonucleoprotein C-like protein that functions as a transcriptional activator of the hepatitis B virus enhancer II. *J. Virol.* **66**, 6841–6848
- Lahiri, D. K. and Thomas, J. O. (1986) A cDNA clone of the hnRNP C proteins and its homology with the single-stranded DNA binding protein UP2. *Nucleic Acids Res.* **14**, 4077–4094
- Kamei, D., Tsuchiya, N., Yamazaki, M., Meguro, H. and Yamada, M. (1999) Two forms of expression and genomic structure of the human heterogeneous nuclear ribonucleoprotein D-like JKTBR gene (HNRPD). *Gene* **228**, 13–22
- Beyer, A. L., Christensen, M. E., Walker, B. W. and LeSturgeon, W. M. (1977) Identification and characterization of the packaging of core 40 S hnRNP particles. *Cell* **11**, 127–138
- Cobianchi, F., Karpel, R. L., Williams, K. R., Notario, V. and Wilson, S. H. (1988) Mammalian heterogeneous nuclear ribonucleoprotein complex protein A1. Large-scale overproduction in *Escherichia coli* and cooperative binding to single-stranded nucleic acids. *J. Biol. Chem.* **263**, 1063–1071
- Burd, C. G., Swanson, M. S., Goriach, M. and Dreyfuss, G. (1989) Primary structures of the heterogeneous nuclear ribonucleoprotein A2, B1, and C2 proteins: a diversity of RNA binding proteins is generated by small peptide inserts. *Proc. Natl. Acad. Sci. U.S.A.* **86**, 9788–9792
- Query, C. C., Bentley, R. C. and Keene, J. D. (1989) A common RNA recognition motif identified within a defined U1 RNA binding domain of the 70 K U1 snRNP protein. *Cell* **57**, 89–101
- Merrill, B. M., LoPresti, M. B., Stone, K. L. and Williams, K. R. (1986) High pressure liquid chromatography purification of UP1 and UP2, two related single-stranded nucleic acid-binding proteins from calf thymus. *J. Biol. Chem.* **261**, 878–883
- Zamore, P. D., Zapp, M. L. and Green, M. R. (1990) RNA binding: betas and basics. *Nature (London)* **348**, 485–486
- Nagai, K., Oubridge, C., Jessen, T. H., Li, J. and Evans, P. R. (1990) Crystal structure of the RNA-binding domain of the U1 small nuclear ribonucleoprotein A. *Nature (London)* **348**, 515–520
- Nagata, T., Kurihara, Y., Matsuda, G., Saeki, J., Kohno, T., Yanagida, Y., Ishikawa, F., Uesugi, S. and Katahira, M. (1999) Structure and interactions with RNA of the N-terminal UUAG-specific RNA-binding domain of hnRNP D0. *J. Mol. Biol.* **287**, 221–237
- Chen, H., Hu, B., Gacad, M. A. and Adams, J. S. (1998) Cloning and expression of a novel dominant-negative-acting estrogen response element-binding protein in the heterogeneous nuclear ribonucleoprotein family. *J. Biol. Chem.* **273**, 31352–31357
- Shinji, K. and Takeshi, M. (1992) A protein binding to CArG box motifs and to single-stranded DNA functions as a transcriptional repressor. *Gene* **119**, 229–236
- Tomonaga, T. and Levens, D. (1995) Heterogeneous nuclear ribonucleoprotein K is a DNA-binding transactivator. *J. Biol. Chem.* **270**, 4875–4881
- Tomonaga, T. and Levens, D. (1996) Activating transcription from single-stranded DNA. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 5830–5835
- Sadowski, I. and Ptashne, M. (1989) A vector for expressing GAL4(1–147) fusions in mammalian cells. *Nucleic Acids Res.* **17**, 7539
- Seiler-Tuyns, A., Walker, P., Martinez, E., Merillat, A. M., Givel, F. and Wahli, W. (1986) Identification of estrogen-responsive DNA sequences by transient expression experiments in a human breast cancer cell line. *Nucleic Acids Res.* **14**, 8755–8770
- Nakamaki, T., Imamura, J., Brewer, G., Tsuruoka, N. and Koeffler, H. P. (1995) Characterization of adenosine-uridine-rich RNA binding factors. *J. Cell. Physiol.* **165**, 484–492

- 31 Sadowski, I., Ma, J., Triezenberg, S. and Ptashne, M. (1988) GAL4-VP16 is an unusually potent transcriptional activator. *Nature (London)* **335**, 563–564
- 32 Michelotti, E. F., Michelotti, G. A., Aronsohn, A. I. and Levens, D. (1996) Heterogeneous nuclear ribonucleoprotein K is a transcription factor. *Mol. Cell. Biol.* **16**, 2350–2360
- 33 Bertolotti, A., Lutz, Y., Heard, D. J., Chambon, P. and Tora, L. (1996) hTAFII68, a novel RNA/ssDNA-binding protein with homology to the pro-oncoproteins TLS/FUS and EWS is associated with both TFIID and RNA polymerase II. *EMBO J.* **15**, 5022–5031
- 34 Idriss, H., Kumar, A., Casas-Finet, J. R., Guo, H., Damuni, Z. and Wilson, S. H. (1994) Regulation of in vitro nucleic acid strand annealing activity of heterogeneous nuclear ribonucleoprotein protein A1 by reversible phosphorylation. *Biochemistry* **33**, 11382–11390
- 35 Hamilton, B. J., Burns, C. M., Nichols, R. C. and Rigby, W. F. C. (1999) Modulation of AUUUA response element binding by heterogeneous nuclear ribonucleoprotein A1 in human T lymphocytes. *J. Biol. Chem.* **272**, 28732–28741
- 36 Pancetti, F., Bosser, R., Krehan, A., Pyerin, W., Itarte, E. and Bachs, O. (1999) Heterogeneous nuclear ribonucleoprotein A2 interacts with protein kinase CK2. *Biochem. Biophys. Res. Commun.* **260**, 17–22
- 37 Gorlach, M., Burd, C. G. and Dreyfuss, G. (1994) The determinants of RNA-binding specificity of the heterogeneous nuclear ribonucleoprotein C proteins. *J. Biol. Chem.* **269**, 23074–23078
- 38 Sharp, Z. D., Smith, K. P., Cao, Z. and Helsel, S. (1990) Cloning of the nucleic acid-binding domain of the rat hnRNP C-type protein. *Biochim. Biophys. Acta* **1048**, 306–309
- 39 Tomonaga, T., Michelotti, G. A., Libutti, D., Uy, A., Sauer, B. and Levens, D. (1998) Unrestraining genetic processes with a protein-DNA hinge. *Mol. Cell* **1**, 759–764
- 40 Loflin, P., Chen, C. and Shyu, A. (1999) Unraveling a cytoplasmic role for hnRNP D in the in vivo mRNA destabilization directed by the AU-rich element. *Genes Dev.* **13**, 1884–1897
- 41 Larola, G., Cuesta, R., Brewer, G. and Schneider, R. J. (1999) Control of mRNA decay by heat shock-ubiquitin-proteasome pathway. *Science* **284**, 499–502

Received 1 December 1999/21 February 2000; accepted 8 March 2000