Heterogeneous nuclear ribonucleoprotein D0B is a sequence-specific DNA-binding protein

Mate TOLNAY*^{†1}, Lyudmila A. VERESHCHAGINA* and George C. TSOKOS*[†]

*Department of Clinical Physiology, Walter Reed Army Institute of Research, 14th and Dahlia Streets, Bldg. 40, Washington, DC 20307-5100, U.S.A., and [†]Department of Medicine, Uniformed Services University of the Health Sciences, 4301 Jones Bridge Road, Bethesda, MD 20814-4799, U.S.A.

Complement receptor 2 (CR2) is important in the regulation of the B lymphocyte response; the regulation of its expression is therefore of central importance. We recently reported that a 42 kDa heterogeneous nuclear ribonucleoprotein (hnRNP) is involved in the transcriptional regulation of the human CR2 gene [Tolnay, Lambris and Tsokos (1997) *J. Immunol.* **159**, 5492–5501]. We cloned the cDNA encoding this protein and found it to be identical with hnRNP D0B, a sequence-specific RNA-binding protein. By using a set of mutated oligonucleotides, we demonstrated that the recombinant hnRNP D0B displays sequence specificity for double-stranded oligonucleotide defined by the CR2 promoter. We conducted electrophoretic mobility-shift assays to estimate the apparent K_d of hnRNP D0B for the double-stranded DNA motif and found it to be 59 nM. Interestingly, hnRNP D0B displayed affinities of 28 and 18 nM for the sense and anti-sense strands of the CR2 promoter-defined

oligonucleotide respectively. The significantly greater binding affinity of hnRNP D0B for single-stranded DNA than for double-stranded DNA suggests that the protein might melt the double helix. The intranuclear concentration of sequence-specific protein was estimated to be 250–400 nM, indicating that the protein binds to the CR2 promoter *in vivo*. Co-precipitation of a complex formed *in vivo* between hnRNP D0B and the TATA-binding protein demonstrates that hnRNP D0B interacts with the basal transcription apparatus. Our results suggest a new physiological role for hnRNP D0B that involves binding to double- and single-stranded DNA sequences in a specific manner and functioning as a transcription factor.

Key words: complement receptor 2, TATA-binding protein, transcription factor.

INTRODUCTION

Complement receptor 2 (CR2; CD21) binds the C3d, C3dg and iC3b fragments of C3, the third complement component, and is expressed on the surface membrane of mature B lymphocytes. CR2 has been shown to have an important role in the regulation of the response of B cells to antigen (reviewed in [1–3]). Split products of C3, the result of complement activation, when deposited on antigens, are able to cross-link the B-cell antigen receptor and CR2 simultaneously. This dual stimulation can be crucial for the generation of an effective primary immune response and results in greatly enhanced B-cell proliferation and antibody production [4,5]. There is evidence that the level of CR2 expression affects the intensity of the B-cell response. *In vitro*, human spleen B cells with high levels of CR2 expression respond better to antigen, indicating the functional importance of the number of CR2 surface molecules per cell [6]. B cells from transgenic mice that are heterozygous at the disrupted *Cr2* locus and express about 50% of the normal level of receptor, respond to antigen less efficiently than B cells from wild-type animals [7,8].

Although the promoter region of the CR2 gene has been identified [9,10] and it has recently been shown that an intronic silencer regulates the expression of CR2 during B-cell ontogenesis [11,12], little is known about how the expression of the receptor is regulated on mature B cells. We have previously shown in the IM-9 human B lymphoblastoid cell line that CR2 expression is

induced by a cAMP-dependent pathway that involves a cAMP-responsive 42 kDa protein that binds specifically to a sequence in the promoter region of the CR2 gene and promotes transcription [13]. We predicted, on the basis of partial amino acid sequence analysis of purified protein, that the 42 kDa protein was identical with, or closely related to, three highly similar heterogeneous nuclear ribonucleoproteins (hnRNPs): AUF1 [14], hnRNP D0 [15] and E2BP [16]. hnRNP is the collective term for the more than 20 proteins that are found in the complex formed around the nascent pre-mRNA [17]. hnRNP proteins usually bind RNA with limited sequence specificity and are among the most abundant proteins in the nucleus. However, their physiological roles are not well understood. Interestingly, hnRNPs D0, AUF1 and E2BP were all reported to bind nucleic acid with unusually high sequence specificity, but different functions have been suggested for each of them. AUF1 binds to AU-rich motifs present in the 3'-untranslated region of the mRNA of several cytokines, oncoproteins and G-protein-coupled receptors, and selectively accelerates degradation of the mRNA [14,18]. A dimerization domain and structural determinants in AUF1 that are required for RNA binding were described recently [19]. AUF1 was shown to be phosphorylated *in vivo* and such a modification could affect its RNA-binding and/or protein-protein interaction capacity [14]. hnRNP D0 was first identified in HeLa cell nuclear extracts as a protein that binds specifically to single-stranded d(TTAGGG)₄ and r(UUAGGG)₄ oligonucleotides, which correspond to the human telomeric repeat

Abbreviations used: CR2, complement receptor 2; CRE, cAMP-response element; EMSA, electrophoretic mobility-shift assay; hnRNP, heterogeneous nuclear ribonucleoprotein; TBP, TATA-binding protein.

¹ To whom correspondence should be addressed, at the Walter Reed Army Institute of Research (e-mail mtolnay@hotmail.com).

The nucleotide sequence data reported will appear in DDBJ, EMBL and GenBank Nucleotide Sequence Databases under the accession number AF039575.

and the pre-mRNA 3' splice site respectively [20]. Three different isoforms were later cloned [15]. hnRNP D0 was also identified as a component of the protein complex that stabilizes the globin mRNA [21]. It has recently been shown that a single gene encodes all four alternatively spliced isoforms of hnRNP D0/AUF1 proteins [22]. However, no double-stranded DNA-binding activity or ability to function as a transcription factor has previously been suggested for hnRNP D0 or AUF1. In contrast, E2BP, a hepatoma-cell-specific hnRNP, has been shown to bind double-stranded DNA and function as a transcription factor of the second promoter of hepatitis B virus [16].

Apart from E2BP, the only other hnRNP that has been shown to function as a DNA-binding transactivator is hnRNP K, which binds to the CT-rich element of the *c-myc* promoter and enhances gene expression [23,24]. It binds single-stranded DNA preferentially over double-stranded DNA or RNA [23–25]. It has been shown to be a weak transcription factor that interacts directly with the TATA-binding protein (TBP) [26]. Its DNA-binding activity is regulated by phosphorylation [27]. Moreover, it has been demonstrated that hnRNP K can function as an architectural transcription factor by binding to and stabilizing single-stranded DNA, thus making the DNA more flexible and allowing other transcription factors to align and interact more easily [25,28].

In the present study we cloned the 42 kDa hnRNP that recognized the CR2 promoter-defined sequence; here we present evidence that it is identical with the B isoform of hnRNP D0. Recombinant hnRNP D0B is a highly sequence-specific protein that binds double-stranded DNA with physiological affinity. We also demonstrate that hnRNP D0B interacts with the TFIID basal transcription complex *in vivo*.

EXPERIMENTAL

Cell cultures

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

Cloning of hnRNP D0B

The probe used for library screening was generated as follows: PCR primers were designed on the basis of the sequenced peptides derived from the purified 42 kDa nuclear protein. We used a primer pair (upper primer 5'-CTATCACAGGGCGATCAA-3'; lower primer 5'-CTGTTGCTGATATTGTTTC-3') and cDNA produced from the RNA from IM-9 cells to amplify the region between the N-terminal and C-terminal peptides. We assumed that not only the amino acid sequences of the peptides but also the corresponding nucleotide sequences of the cDNA were identical with those of the three related proteins produced by a similarity search of GenBank. This assumption was based on the fact that the cDNA species for all three related proteins were identical along a 554 bp sequence in the middle of the coding region, where all sequenced peptides were located. We also assumed that the order of the peptides in the 42 kDa protein was the same as in the related proteins. We amplified the region between the N-terminal and C-terminal peptides by using cDNA produced from IM-9 cell RNA. A single 404 bp product was generated. Its sequence was found to be identical with that reported for the corresponding cDNA sequence of the three

homologous proteins. The 404 bp PCR product was labelled with random priming to an approximate specific radioactivity of 1.5×10^9 c.p.m./mg and used to screen a human blood cell cDNA library in the lambda ZAP express[™]/EcoRI vector (purchased from Stratagene, La Jolla, CA, U.S.A.). Approx. 120 000 plaques were screened at a density of 2×10^4 plaque-forming units per 150 mm plate, as suggested by Stratagene. Plaque lifts were done in duplicate. The hybridization solution contained $(1-3) \times 10^5$ c.p.m./ml denatured probe, $6 \times$ SSC buffer, 0.1% SDS, $2 \times$ Denhardt's reagent and 50 mg/ml denatured, sonicated salmon sperm DNA. Hybridization was performed overnight at 68 °C, followed by an extensive wash under stringent conditions. Potential positive plaques were rescreened. Individual plaques were picked from confirmed positives. Cloned inserts were excised and recircularized *in vivo* from the lambda ZAP express[™] vector to form a phagemid containing the cloned insert, as suggested by Stratagene. Expression of the inserts in *Escherichia coli* (strain XL0LR) was induced by 10 mM isopropyl β -D-thiogalactoside for 1 h and proteins were isolated from the bacteria, as suggested by Stratagene. A 2 μ l portion of each sample was tested with electrophoretic mobility-shift assay (EMSA) for the presence of CR2 promoter-defined oligonucleotide-binding activity. Samples containing a 50-fold excess of unlabelled oligonucleotide or 1 μ l of AUF1 antibody were also tested. The insert in clone 5 has been sequenced from both directions by Genomix (Duluth, GA, U.S.A.).

Expression and purification of recombinant hnRNP D0B

We used Novagen's pET expression system. The full coding region of hnRNP D0B with a His₆ tag, in the pET-21a(+) vector, was obtained from Dr. Ishikawa (Yokohama, Japan). We induced the expression of the protein in BL21(DE3)pLysS bacteria with 1 mM isopropyl β -D-thiogalactoside for 4 h. Recombinant hnRNP D0B was purified on Ni²⁺ columns (Novagen), as suggested by the manufacturer. The concentration of the purified recombinant protein was measured by using Bio-Rad protein microassay.

EMSA

Nuclear proteins from IM-9 cells were obtained as described [13]. The protein concentrations of all nuclear extracts were determined spectrophotometrically by measuring A_{280} . To obtain double-stranded oligonucleotides, the complementary strands were mixed at an equal molar ratio, incubated at 95 °C for 5 min and allowed to cool slowly. Oligonucleotides were end-labelled with [³²P]ATP by using a kit from Gibco BRL, purified on Centri-Sep columns (Princeton Separations, Adelphia, NJ, U.S.A.) and used as probes in EMSA. Nuclear extracts, purified cellular protein or recombinant proteins were preincubated in 20 μ l of buffer containing 12% (w/v) Ficoll, 50 mM KCl, 20 mM Hepes, pH 7.9, 1 mM EDTA, 1 mM dithiothreitol, poly(dI-dC)·poly(dI-dC) and 7 μ g of BSA for 10 min to block non-specific binding; the labelled oligonucleotide was then added and incubated at 25 °C for 20 min. We used 0.1 μ g of poly(dI-dC)·poly(dI-dC) per sample for the competition experiments with the recombinant or purified cellular protein, 2 μ g with total nuclear extracts, and no poly(dI-dC)·poly(dI-dC) for the K_d measurements. The protein-DNA complex was separated on a 4% (w/v) non-denaturing acrylamide gel in $0.5 \times$ TBE. Gels were analysed with a PhosphorImager (Molecular Dynamics, Sunnyvale, CA, U.S.A.). Curves were fitted by GraFit software

(Erithacus Software Ltd.). Competition experiments were performed by incubating excess unlabelled oligonucleotide with nuclear extracts for 10 min before the addition of the labelled oligonucleotide. In supershift experiments the AUF1 antibody (gift from Dr. Gary Brewer, Winston-Salem, NC, U.S.A.) was added after preincubation of the nuclear protein with labelled oligonucleotide; samples were then incubated for 10 min on ice.

Cell size estimation

We estimated the size of IM-9 B lymphoblastoid cells microscopically, with nucleus-specific Wright's staining.

Co-precipitation and Western blots

Nuclear protein extracts (100 µg of protein per sample) in 0.5 ml of PBS containing 0.1 % (v/v) Nonidet P40, 1 mM PMSF, 2 µg/ml aprotinin and 0.35 mg/ml BSA were precleared with 5 µl of agarose beads at 4 °C for 30 min. The supernatants were immunoprecipitated with 5 µl of agarose-conjugated rabbit antibody against human TBP (Santa Cruz Biotechnology, Santa Cruz, CA, U.S.A.) or co-precipitated with 5 µl of agarose-conjugated full-length recombinant human TBP (Santa Cruz Biotechnology) at 4 °C for 1 h. Some samples also included 0.5 pmol of long CR2 promoter-defined double-stranded oligonucleotide and 5 µg of poly(dI-dC)·poly(dI-dC). The beads were subsequently washed twice with incubation buffer and twice with PBS. Proteins were eluted by boiling the beads in 50 µl of SDS/polyacrylamide gel loading buffer for 5 min. Eluted proteins, full nuclear extracts and recombinant hnRNP D0 were subjected to SDS/PAGE [10 % (w/v) gel], then transferred to nitrocellulose membranes (Gibco BRL). Membranes were blocked overnight at 4 °C with PBS containing 5 % (w/v) non-fat dried milk, then incubated with AUF1 antibody diluted 1:3000 (gift from Dr. Gary Brewer) in PBS containing 2 % (w/v) non-fat dried milk at 25 °C for 2 h to detect hnRNP D0/AUF1 proteins. The membranes were washed extensively, then incubated with horseradish peroxidase-labelled goat anti-(rabbit IgG) diluted 1:500 (Santa Cruz Biotechnology) in PBS containing 2 % (w/v) non-fat dried milk at 25 °C for 2 h. After extensive washing, the bands were detected with an enhanced chemiluminescence Western detection system (Amersham, Arlington Heights, IL, U.S.A.). Bands were quantified with a laser densitometer (Molecular Dynamics).

Table 1 Sequence of all oligonucleotides used in EMSA

Sense strands only are shown, except when the anti-sense strand was used single-stranded. Mutated nucleotides are underlined.

Oligonucleotide name	Sequence (5' → 3')	Length (bases)
Short CR2 promoter-defined	TAGAGATTACTTCAAATAA	20
Mutated CR2 promoter-defined 1	TAGAGAT <u>G</u> AGTCAAATAA	20
Mutated CR2 promoter-defined 2	TAGAGATTAGTCAAATAA	20
Consensus CRE	ATTGCTGACGTACAGAGGC	20
AP-2, sense	CGTCAGCCCCGAGGCCGCTGCA	23
AP-2, anti-sense	TGCAGGCCGCTCGGGCTGACG	23
Long CR2 promoter-defined, sense	TTCTGTAGAGATTACTTCAAATAAAGA	28
Long CR2 promoter-defined, anti-sense	TCTTTATTTGAAGTAATCTCTACAGAA	28

Figure 1 Recombinant hnRNP D0B runs as a 42 kDa protein on SDS/PAGE and displays sequence preferences similar to that of the purified cellular protein

(A) Purified recombinant hnRNP D0B 2.5 µg was subjected to SDS/PAGE [10 % (w/v) gel] along with molecular mass standards (lane M) and stained with Coomassie Blue. The arrow indicates the position of the recombinant protein. The molecular masses of the standards are indicated (in kDa) at the right. (B) Recombinant hnRNP D0B (2 ng) or 1 µl of cellular protein, diluted 1:1000, that had been purified from IM-9 B lymphoblastoid cells [13] were preincubated with excess unlabelled double-stranded oligonucleotides for 10 min before the addition of 20 fmol of labelled double-stranded short CR2 promoter-defined oligonucleotide and analysed by EMSA. Unlabelled oligonucleotides are indicated above the lanes. Numbers refer to fold molar excess of unlabelled oligonucleotide. Mutated 1 and 2 are mutated CR2 promoter-defined oligonucleotide 1 and 2 respectively.

RESULTS

Molecular cloning of a sequence-specific nuclear protein that recognizes a CR2 promoter-defined double-stranded oligonucleotide

We previously identified a B-cell nuclear protein that recognized a CR2 promoter-defined oligonucleotide (see Table 1 for the sequences of all oligonucleotides used), and suggested that the protein was involved in the regulation of the expression of the CR2 gene [13]. We purified a 42 kDa sequence-specific nuclear protein from B-cell extracts and sequenced three internal peptides. A homology search of GenBank with the peptide sequences (a total of 26 amino acid residues) produced three human hnRNP that matched 100 % with all three peptides: hnRNP D0, AUF1 and E2BP.

To characterize the nuclear protein completely, we screened a human blood cell cDNA library with a 404 bp PCR product that

Figure 2 hnRNP D0B binds to double-stranded and single-stranded DNA in a sequence-specific manner

Samples (2 ng) of recombinant hnRNP D0B were preincubated with excess unlabelled double-stranded (A) or single-stranded (B) oligonucleotides for 10 min before the addition of 20 fmol of labelled double-stranded short CR2 promoter-defined oligonucleotide and analysed by EMSA. Percentages of bound labelled oligonucleotide were calculated and plotted against the fold molar excess of unlabelled oligonucleotide. The following unlabelled oligonucleotides were used. (A) All double-stranded: ● and solid line, long CR2 promoter-defined; ○ and broken line, short CR2 promoter-defined; ■ and dotted line, mutated CR2 promoter-defined 1; □ and dot-dashed line, mutated CR2 promoter-defined 2; × and double-dot-dashed line, consensus CRE. (B) All oligonucleotides were single-stranded, except the long double-stranded CR2 promoter-defined (● and solid line), which we show for comparison; ○ and broken line, long CR2 promoter-defined, sense strand; × and dotted line, long CR2 promoter-defined, anti-sense strand; ■ and dot-dashed line, AP-2, anti-sense strand. Abbreviation: oligo, oligonucleotide.

was amplified by using primers based on the sequenced peptides (see the Experimental section). The probe corresponded to the identical middle part of hnRNP D0, AUF1 and E2BP. A total of 25 plaques were found to be positive after rescreening. To determine which clones contained an insert encoding the specific protein, expression of the inserts was induced and protein samples were tested by using EMSA for the presence of CR2 promoter-defined oligonucleotide-binding activity. Two clones were identified that encoded a protein recognizing the CR2 promoter-defined oligonucleotide. The interactions were sequence-specific because they were efficiently inhibited by a 50-fold molar excess of unlabelled CR2 promoter-defined oligonucleotide. Furthermore an antibody recognizing both hnRNP D0 and AUF1, but no other hnRNP, supershifted both expressed proteins, con-

Table 2 Summary of inhibitory capacities of oligonucleotides, apparent K_d values for hnRNP D0B binding and corresponding relative transcriptional activities

The fold excess of unlabelled oligonucleotide (oligo) is defined as that required in EMSA to inhibit the interaction of recombinant hnRNP D0B with labelled short double-stranded CR2 promoter-defined oligonucleotide. For K_d determination, triplicate samples were analysed by EMSA (mean $\pm$ S.E.). * $P < 0.05$, ** $P < 0.01$, compared with the double-stranded CR2 promoter-defined oligonucleotide, with Student's unpaired t test. Relative activities in the reporter gene assay are summarized from our previous study ([13], and M. Tolnay, L. A. Vereshchagina and G. C. Tsokos, unpublished work); + + +, strong ability of the inserted sequence to drive the transcription of a CAT reporter gene in IM-9 cells; +, transcriptional rate similar to that of a CAT construct without insert. Abbreviations: ds, double-stranded; n.d., not determined.

Oligo name	Excess unlabelled oligo needed for 50% inhibition (fold)	K_d (app) for hnRNP D0B binding (nM)	Relative activity in reporter gene assay
Short CR2 promoter-defined, ds	4.2	n.d.	+ + +
Long CR2 promoter-defined, ds	2.3	59 ± 4	+ + +
Long CR2 promoter-defined, sense	9.7	$28 \pm 1^{**}$	
Long CR2 promoter-defined, anti-sense	1.2	$18 \pm 2^{**}$	
Mutated CR2 promoter-defined 1, ds	24	n.d.	n.d.
Mutated CR2 promoter-defined 2, ds	70	n.d.	+
Consensus CRE, ds	> 128	n.d.	
AP-2, ds	> 128	n.d.	
AP-2, sense	> 128	n.d.	
AP-2, anti-sense	> 128	$164 \pm 4^*$	

firmed that they are related hnRNPs (results not shown). The insert in clone 5 has been sequenced. The 1741 bp insert contained a sequence that was identical with the reported coding sequence of hnRNP D0B (clone cDx7), one of the three reported isoforms of hnRNP D0 [15]. Our clone contained a novel extended downstream non-coding sequence of 494 bp. A single base substitution was noted in the non-coding sequence, 242 bp after the stop codon.

Molecular mass and DNA binding of recombinant hnRNP D0B

His₆-tagged hnRNP D0B inserted into a bacterial expression vector, which was obtained from Dr. Ishikawa, was used to produce purified recombinant hnRNP D0B [15]. The recombinant protein ran as a 42–43 kDa protein on SDS/PAGE (Figure 1A), further supporting the notion that the 42 kDa cellular protein that we have purified from B cells is identical with hnRNP D0B. The His₆ tag on the recombinant protein added approx. 1 kDa to the molecular mass and thus did not substantially modify the molecular mass of the native protein. As reported previously [15], the mobility of hnRNP D0B during SDS/PAGE is considerably less than the expected mobility, calculated from the molecular mass, owing partly to the basic nature of the protein. The smaller contaminating peptide that was present in the sample was a proteolytic fragment of the recombinant protein that did not constitute more than 15% of the total protein concentration. As expected, hnRNP D0B could be supershifted with an antibody that was specific to hnRNP D0 and AUF1 (results not shown), similarly to the cellular protein.

To test whether the recombinant protein could bind to the CR2 promoter-defined double-stranded oligonucleotide, we performed EMSA. Recombinant hnRNP D0B indeed produced a shifted band (Figure 1B) with a mobility similar to that of the

Figure 3 Binding affinity of recombinant hnRNP D0B for the CR2 promoter-defined oligonucleotide

Apparent K_d values for the binding of recombinant hnRNP D0B to DNA substrates were estimated by using EMSA. Labelled oligonucleotide (2 fmol) was incubated with increasing amounts of recombinant hnRNP D0B in a total volume of 20 μl in triplicate. Samples were analysed by EMSA. Representative gels are shown on the left; corresponding plots of percentage of free oligonucleotide (oligo) against hnRNP D0B concentration are shown on the right. The concentrations of hnRNP D0B are shown (in nM) above the gel lanes. The arrow indicates the position of the shifted band. The calculated K_d value is shown for each oligonucleotide.

purified cellular hnRNP. While performing EMSA with recombinant hnRNP D0B we noticed that the shifted band frequently split to form a close doublet. The use of higher protein concentrations resulted in the appearance of a third shifted band that migrated more slowly (results not shown). Similar observations have been reported with other recombinant or purified hnRNP proteins and might indicate their potential to form homodimers [19]. When we preincubated the samples with unlabelled oligonucleotides, we found that recombinant hnRNP D0B and purified cellular protein produced strikingly similar inhibition patterns with different unlabelled oligonucleotides (Figure 1B), clearly

indicating comparable sequence specificity. Two mutated double-stranded oligonucleotides that had two bases changed were about 10% as inhibitory as the original CR2 promoter-defined oligonucleotide with both the recombinant hnRNP D0B and the cellular protein. A sample with cellular protein and unlabelled 'mutated oligonucleotide 1' was not included in the experiment shown in Figure 1(B) but was tested in other experiments. A 50-fold molar excess of mutated oligonucleotides displayed inhibition patterns similar to that of a 5-fold excess of unmodified oligonucleotide. A consensus cAMP-response element (CRE) oligonucleotide, which we used as a non-specific control, dis-

played only minimal (8–20%) inhibitory capacity at 50-fold molar excess. These experiments demonstrate that recombinant hnRNP D0B and the cellular protein have similar molecular masses and display indistinguishable DNA-binding specificities.

Sequence specificity of hnRNP D0B

To explore further the degree of sequence specificity of the recombinant protein, we compared the ability of several unlabelled oligonucleotides to inhibit the interaction between the protein and labelled specific oligonucleotide in EMSA. With this approach we could directly detect binding preferences that hnRNP D0B displayed for the specific labelled oligonucleotide and the unlabelled competitor. In all experiments shown, we used labelled double-stranded 20 bp CR2 promoter-defined oligonucleotide. We added the unlabelled competitor oligonucleotides at a wide range of molar excesses and calculated the fold excesses that were required to produce 50% inhibition. As unlabelled inhibitors we tested two double-stranded CR2 promoter-defined oligonucleotides that differed in their lengths. Although each of them could effectively inhibit the binding of hnRNP D0B to the labelled oligonucleotide, the 28-mer oligonucleotide, which contained additional flanking sequences from the CR2 promoter, was more effective than the 20-mer (Figure 2A and Table 2). We also used two mutated double-stranded oligonucleotides, which had had two bases changed, as competitors. Mutated oligonucleotide 1 was approx. 1/6 as inhibitory, and mutated oligonucleotide 2 was 1/17 as inhibitory, as the unmutated oligonucleotide of the same length. Two unrelated double-stranded oligonucleotides (consensus CRE and AP-2), when added at molar excesses of up to 128-fold, inhibited the binding of the protein by less than 10% (results not shown for AP-2). We examined whether the recombinant hnRNP D0B also binds to single-stranded DNA. Both the sense and anti-sense strands of the CR2 promoter-defined oligonucleotide inhibited the binding of hnRNP D0B to labelled double-stranded oligonucleotide (Figure 2B and Table 2). The anti-sense strand inhibited the binding of the protein more efficiently. The sense and anti-sense strands of AP-2, which were used as controls, inhibited the binding by 20% and 30% respectively when used at 128-fold excess (results not shown for the sense strand).

Binding affinity of hnRNP D0B for double-stranded and single-stranded DNA

The binding affinity of hnRNP D0B for the CR2 promoter-defined oligonucleotide reflects the potency of the protein to bind to the CR2 promoter *in vivo* and increase its transcription rate. We therefore estimated by EMSA the binding affinity of hnRNP D0B for selected oligonucleotides by following described approaches [29]. Samples containing the same amount of labelled oligonucleotide and increasing amounts of recombinant protein were analysed and apparent K_d values were determined as the protein concentration at which 50% of the DNA was bound. The apparent K_d for the 28 bp double-stranded CR2 promoter-defined oligonucleotide was measured as 59 nM (Figure 3). By using the sense and anti-sense strands of the same oligonucleotide we detected significantly lower K_d values: 28 and 18 nM respectively. The K_d for the anti-sense strand of AP-2 (negative control) was estimated as 164 nM.

Intranuclear molar concentration of hnRNP D0

To determine whether the measured K_d values were in agreement with the binding of the protein to its DNA target *in vivo*, we

Figure 4 Number of specific oligonucleotide-binding hnRNP D0 molecules per nucleus of IM-9 B lymphoblastoid cells

(A) Nuclear protein (0.025 μ g), obtained from approx. 1200 IM-9 B lymphoblastoid cells, was incubated with increasing amounts of labelled double-stranded long CR2 promoter-defined oligonucleotide and analysed by EMSA. (B) A standard curve was obtained by running on the gel known amounts of labelled oligonucleotide to SDS/PAGE without protein. The number of specific proteins per nucleus was calculated by assuming a 1:1 stoichiometry for the protein–oligonucleotide interaction.

Figure 5 Total number of hnRNP D0 molecules per nucleus of IM-9 B lymphoblastoid cells

Different amounts of nuclear protein obtained from a known number of cells (18 μ g of nuclear protein was obtained from 10^6 IM-9 cells) were subjected to SDS/PAGE along with known amounts of recombinant hnRNP D0B as a standard, then transferred to nitrocellulose and analysed by Western blot analysis with anti-hnRNP D0/AUF1 antibody. Bands were quantified with a laser densitometer. The number of hnRNP D0 molecules per nucleus was calculated from a standard curve obtained with recombinant hnRNP D0B. A representative of three experiments is shown.

estimated the molar concentration of hnRNP D0 in the nucleus of IM-9 B lymphoblastoid cells by using two methods. We used IM-9 cells to correlate the concentration with functional data

Figure 6 hnRNP D0 and TBP form a complex *in vivo*

Nuclear extracts of IM-9 cells that had been precleared with agarose beads were incubated with either anti-TBP antibody conjugated to agarose beads (lanes 2 and 3) or recombinant human TBP conjugated to agarose beads (lanes 4 and 5). The samples in lanes 3 and 5 also contained 0.5 pmol of CR2 promoter-defined oligonucleotide. After an extensive washing, co-precipitated proteins were eluted and subjected to SDS/PAGE along with 15 μ g of total nuclear extract (lane 1) and 0.2 μ g of recombinant hnRNP D0B and hnRNP D0D (lanes 6 and 7). Proteins were transferred to nitrocellulose and analysed by Western blot analysis with anti-(hnRNP D0/AUF1) antibody. Arrows indicate the bands corresponding to hnRNP D0B and hnRNP D0D. The other two bands in lanes 2 and 3 were due to rabbit antibody that leached off the agarose beads. The presence or absence of CR2 promoter-defined oligonucleotide is indicated by + or – respectively.

that we had obtained *in vivo* in our previous studies with these cells [13]. To calculate the intranuclear concentration of the protein, we estimated the number of molecules per nucleus and the volume of the nucleus. To estimate the number of protein molecules that can bind the specific oligonucleotide, we performed EMSA with nuclear extracts purified from a known number of IM-9 cells. We incubated the same amount of nuclear protein with increasing amounts of labelled specific oligonucleotide until saturation was reached (Figure 4). Assuming a 1:1 stoichiometry for the protein–oligonucleotide interaction, we calculated the number of protein molecules present in the shifted band at saturation from a standard curve obtained with free oligonucleotide. By using this method we estimated there to be approx. $(3\text{--}5) \times 10^5$ oligonucleotide-binding proteins per nucleus. The diameter of IM-9 lymphoblastoid cells is approx. 17 μ m and the nucleus constitutes approx. 80 % of the total cell volume, as estimated from nucleus-specific Wright's staining. By using these values the molar concentration of the CR2 promoter-defined oligonucleotide-specific protein was found to be approx. 250–400 nM. That number indicates that the protein binds to 50 % of DNA motifs for which its K_d is 250–400 nM. This also means that the protein saturates under physiological conditions the double-stranded CR2 promoter-defined motif, for which it displays a K_d of 59 nM.

To estimate the total number of hnRNP D0 protein molecules per nucleus we performed Western blots with nuclear extracts purified from a known number of cells (Figure 5). The membranes were immunoblotted with an antibody specific for hnRNP D0/AUF1. As a standard we used recombinant hnRNP D0B. We detected two protein bands in nuclear extracts, corresponding to the B and D isoforms of hnRNP D0. Calculations with the densitometric reading of the bands indicated that there were approx. $(2.5\text{--}3.5) \times 10^6$ hnRNP D0B and hnRNP D0D molecules per nucleus, corresponding to an intranuclear protein concentration of 2000–2800 nM.

hnRNP D0 proteins interact with TFIID *in vivo*

If hnRNP D0B is a transcription factor, it is likely to interact with the RNA polymerase II transcription apparatus. At the centre of this apparatus is TFIID, a multiprotein complex composed of the TBP and a set of TBP-associated factors (TAF_{II}) [30]. To determine whether hnRNP D0B associates with the TBP *in vivo*, nuclear proteins from IM-9 cells were immunoprecipitated with a TBP-specific antibody and the precipitated proteins were analysed by immunoblot analysis with antibody specific for hnRNP D0. hnRNP D0B and hnRNP D0D were immunoprecipitated with the TBP-specific antibody (Figure 6, lanes 2 and 3). In addition, both isoforms of hnRNP D0 co-precipitated with recombinant TBP conjugated to agarose beads, further supporting the interaction between hnRNP D0 and TFIID (Figure 6, lanes 4 and 5). The addition of an oligonucleotide containing a single recognition site for hnRNP D0 did not cause a clear change in the amount of hnRNP D0 that was precipitated with either TBP-specific antibody or recombinant TBP.

DISCUSSION

hnRNP proteins are generally considered to be RNA-binding proteins that display limited sequence specificity. Here we cloned a sequence-specific DNA-binding hnRNP with high affinity for a CR2 promoter-defined oligonucleotide. Recombinant hnRNP D0B displayed a significantly higher affinity for single-stranded DNA than for double-stranded DNA. We have shown that hnRNP D0B displays features similar to that of the nuclear protein that we had previously purified from IM-9 B lymphoblastoid cells and we had predicted to be an hnRNP D0/AUF1 protein [13]. The molecular masses of both the recombinant and cellular proteins were found to be approx. 42 kDa by SDS/PAGE and both proteins displayed identical binding preferences for a set of oligonucleotides. The His₆ tag on the recombinant protein is not likely to result in substantial change of mobility because the molecular mass of the additional amino acid residues is approx. 1 kDa and the native protein is basic. hnRNP D0B translated *in vitro* was found to run as a 40–42 kDa protein by SDS/PAGE [21]. Therefore we conclude that we have molecularly cloned the cDNA for the 42 kDa B-cell-derived nuclear protein that we previously purified and partly sequenced.

We found that the hnRNP D0B shows exceptional sequence specificity for the double-stranded CR2 promoter-defined oligonucleotide. Two oligonucleotides mutated at two bases competed for the binding of the protein 1/6 and 1/17 as efficiently as the unmutated oligonucleotide. Because all four reported isoforms of hnRNP D0/AUF1 as well as E2BP have an identical 184-residue middle part, it is important to consider whether all proteins bind to the CR2 promoter-defined oligonucleotide in a similar fashion. The longest isoform, hnRNP D0D (clone cDx9), is especially interesting because it contains all the amino acids that are present in hnRNP D0B, plus a 49-residue insert close to the C-terminus of the protein. Recombinant hnRNP D0D with a His₆ tag runs as a 46 kDa protein by SDS/PAGE (results not shown), so it is unlikely to be identical with the 42 kDa B-cell-derived protein. However, in EMSA, hnRNP D0D displayed the same sequence specificity as hnRNP D0B and had a similar K_d for the CR2 promoter-defined oligonucleotide (results not shown). These findings indicate that the 49-residue insert does not alter the ability of the protein to recognize the CR2 promoter-defined oligonucleotide. We consistently detected in B-cell nuclear extracts a single protein that recognized the CR2 promoter-defined oligonucleotide and were unable to purify [13], or clone,

a sequence-specific protein with a molecular mass of approx. 46 kDa. However, we could clearly detect the D isoform of hnRNP D0 in IM-9 cells by Western blot analysis. It is possible that hnRNP D0D interacts with proteins and/or nucleic acids owing to its extra insert but is not able to interact efficiently with the CR2 promoter-defined sequence, at least under the conditions of the EMSA. The second isoform, hnRNP D0A (clone cDx4), with the 49-residue insertion but without a 19-residue insertion in the N-terminal RNA-binding domain, is a less abundant isoform. hnRNP D0A has been shown to display a different sequence specificity for RNA from that of hnRNP D0B, suggesting that the 19-residue insertion modifies the sequence preference of hnRNP D0 [15]. AUF1 is another hnRNP that shows great similarity to hnRNP D0B [14]. The recombinant p37 AUF1, which is the shortest isoform lacking both the 49-residue and 19-residue inserts [22], was tested in EMSA with specific oligonucleotide; however, we did not detect any binding (results not shown). We suggest that sequence-specific binding to the CR2 promoter-defined oligonucleotide is not a general property of hnRNP D0/AUF1 proteins, in spite of extensive sequence similarities. It is not clear which are the structural determinants of sequence-specific binding. The highest-affinity binding of the AUF1 protein to RNA required the presence of both RNA recognition motifs and an alanine-rich region at the N-terminus and a glutamine-rich region at the C-terminus (which is also present in hnRNP D0B), suggesting that the interaction of the protein with the nucleic acid is quite complex [19].

We measured an apparent K_d of 59 nM for the interaction of hnRNP D0B with double-stranded CR2 promoter-defined oligonucleotide. We estimated by EMSA the molar concentration of the specific protein in IM-9 cell nuclei to be in the range 250–400 nM. This implies that hnRNP D0B binds to the CR2 promoter *in vivo*. By using Western blots we estimated the total intranuclear concentration of hnRNP D0B to be 2000–2800 nM. Western blots detect all hnRNP D0B molecules, whereas EMSA identifies only those that can bind to the specific oligonucleotide. It is possible that only a fraction of hnRNP D0B molecules are in the proper phosphorylation state to bind DNA. Furthermore, because hnRNP D0/AUF1 proteins were found to bind to several nucleic acid sequences and interact with a number of proteins [14], it is possible that most of the molecules are incorporated into other protein–nucleic acid complexes and are not free to bind to the CR2 promoter-defined oligonucleotide. Indeed, it has been suggested that hnRNP K, an equally abundant protein with multiple functions, co-ordinates RNA processing, signal transduction and transcription by modulating all of these processes, but at limiting concentrations [26]. We have found that hnRNP D0B also binds to the sense and anti-sense strands of the CR2 promoter-defined oligonucleotide with affinities of 28 and 18 nM respectively. Although both strands are rich in A and T (75% AT), we could not find a clear pattern of similarity between the two strands. The affinity of hnRNP D0B for its RNA target sequence, r(UUAGGG)₄, has been estimated to be 34 nM [15]. hnRNP D0 proteins were also found to bind to the human telomeric repeat, d(TTAGGG)₄, but with lower affinities [20]. AUF1 was shown to bind to (AUUU)₅ motifs as well as to AU-rich motifs in the *c-myc* and *c-fos* mRNA species with an apparent K_d of 8–21 nM [29].

We have experimental evidence suggesting that the cellular protein, identified now as hnRNP D0B, functions in B cells as a transcriptional regulator of the human CR2 gene [13]. The sequence preferences of hnRNP D0B support this position. We found that high-affinity binding to DNA is correlated with the ability of the DNA motif to drive the transcription of a reporter gene (Table 2). The native CR2 promoter-defined oligonucleotide

did drive the transcription of a reporter gene construct, whereas the oligonucleotide that contained two bases changed and was 1/17 as inhibitory in EMSA did not. How does hnRNP D0B work? It is likely that hnRNP D0B has its own transactivator domain, as does hnRNP K [24]. Our finding that hnRNP D0B, and also D0D, associates with the TBP *in vivo* strongly suggests that hnRNP D0 affects transcription by interacting with the TFIID complex. In fact, transcription factors usually target TFIID by binding with their transactivating domains to the TBP or one of the TAF_{II}, which in turn influence the assembly of the transcription initiation complex [31]. In addition, the protein might melt the double helix through its even higher affinity for specific single-stranded DNA. On a theoretical basis, a protein that has higher affinity for single-stranded DNA than for double-stranded DNA always causes thermal destabilization of the double helix and shifts the equilibrium toward the single-stranded state [32]. Melting of DNA could facilitate transcription by increasing the flexibility of the otherwise rigid double-stranded helix, which in turn can facilitate the alignment of regulatory proteins that bind to *cis* elements at the opposing sides of the single-stranded region [28,33]. A similar mechanism of action has recently been demonstrated for another transcription factor, hnRNP K [28].

Whereas hnRNP D0B has previously been reported to bind RNA, and with lower affinity single-stranded DNA [20], we propose a new physiological role for the protein. The ability of hnRNP proteins to bind DNA in a sequence-specific manner and to influence gene transcription is important. The abundance of hnRNP proteins might allow them to bind their DNA and RNA motifs at the same time and thus to participate in multiple cellular functions. The unusual features of hnRNP D0/AUF1 proteins might highlight interconnections between transcription, RNA stability and DNA conformation.

We thank Dr. David Levens for helpful discussions, Dr. Tibor Szentendrei for his help with sequencing, Dr. Fuyuki Ishikawa and Dr. Yasuko Kajita for providing hnRNP D0B and hnRNP D0D in an expression vector, and Gary Brewer for the AUF1-specific antibody and recombinant AUF1. This work was supported by National Institutes of Health grant AI42782.

REFERENCES

- 1 Fearon, D. T. and Carter, R. H. (1995) *Annu. Rev. Immunol.* **13**, 127–149
- 2 Carroll, M. C. and Fischer, M. B. (1997) *Curr. Opin. Immunol.* **9**, 64–69
- 3 Tolnay, M. and Tsokos, G. C. (1998) *Clin. Immunol. Immunopathol.* **88**, 123–132
- 4 Mongini, P. K. A., Vilensky, M. A., Highet, P. F. and Inman, J. (1997) *J. Immunol.* **159**, 3782–3791
- 5 Dempsey, P. W., Allison, M. E. D., Akkaraju, S., Goodnow, C. C. and Fearon, D. T. (1996) *Science* **271**, 348–350
- 6 Boyd, A. W., Anderson, K. C., Freedman, A. S., Fisher, D. C., Slaughenhaupt, B., Schlossman, S. F. and Nadler, L. M. (1985) *J. Immunol.* **134**, 1516–1523
- 7 Molina, H., Holers, V. M., Li, B., Fung, Y., Mariathasan, S., Goellner, J., Strauss-Schoenberger, J., Karr, R. W. and Chaplin, D. D. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 3357–3361
- 8 Takahashi, K., Kozono, Y., Waldschmidt, T. J., Berthiaume, D., Quigg, R. J., Baron, A. and Holers, V. M. (1997) *J. Immunol.* **159**, 1557–1569
- 9 Rayhel, E. J., Dehoff, M. H. and Holers, V. M. (1991) *J. Immunol.* **146**, 2021–2026
- 10 Yang, L. M., Behrens, M. and Weis, J. J. (1991) *J. Immunol.* **147**, 2404–2410
- 11 Hu, H., Martin, B. K., Weis, J. J. and Weis, J. H. (1997) *J. Immunol.* **158**, 4758–4768
- 12 Makar, K. W., Pham, C. T., Dehoff, M. H., O'Connor, S. M., Jacobi, S. M. and Holers, V. M. (1998) *J. Immunol.* **160**, 1268–1278
- 13 Tolnay, M., Lambris, J. D. and Tsokos, G. C. (1997) *J. Immunol.* **159**, 5492–5501
- 14 Zhang, W., Wagner, B. J., Ehrenman, K., Schaefer, A. W., DeMaria, C. T., Crater, D., DeHaven, K., Long, L. and Brewer, G. (1993) *Mol. Cell Biol.* **13**, 7652–7665
- 15 Kajita, Y., Nakayama, J., Aizawa, M. and Ishikawa, F. (1995) *J. Biol. Chem.* **270**, 22167–22175
- 16 Tay, N., Chan, S. H. and Ren, E. C. (1992) *J. Virol.* **66**, 6841–6848

-
- 17 Dreyfuss, G., Matunis, M. J., Pinol-Roma, S. and Burd, C. G. (1993) *Annu. Rev. Biochem.* **62**, 289–321
 - 18 Nakamaki, T., Imamura, J., Brewer, G., Tsuruoka, N. and Koeffler, H. P. (1995) *J. Cell Physiol.* **165**, 484–492
 - 19 DeMaria, C. T., Sun, Y., Long, L., Wagner, B. J. and Brewer, G. (1997) *J. Biol. Chem.* **272**, 27635–27643
 - 20 Ishikawa, F., Matunis, M. J., Dreyfuss, G. and Cech, T. R. (1993) *Mol. Cell Biol.* **13**, 4301–4310
 - 21 Kiledjian, M., DeMaria, C. T., Brewer, G. and Novick, K. (1997) *Mol. Cell. Biol.* **17**, 4870–4876
 - 22 Wagner, B. J., DeMaria, C. T., Sun, Y., Wilson, G. M. and Brewer, G. (1998) *Genomics* **48**, 195–202
 - 23 Takimoto, M., Tomonaga, T., Matunis, M., Avigan, M., Krutzsch, H., Dreyfuss, G. and Levens, D. (1993) *J. Biol. Chem.* **268**, 18249–18258
 - 24 Tomonaga, T. and Levens, D. (1995) *J. Biol. Chem.* **270**, 4875–4881
 - 25 Tomonaga, T. and Levens, D. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 5830–5835
 - 26 Michelotti, E. F., Michelotti, G. A., Aronsohn, A. I. and Levens, D. (1996) *Mol. Cell. Biol.* **16**, 2350–2360
 - 27 Bomsztyk, K., Van Seuningen, I., Suzuki, H., Denisenko, O. and Ostrowski, J. (1997) *FEBS Lett.* **403**, 113–115
 - 28 Tomonaga, T., Michelotti, G. A., Libutti, D., Uy, A., Sauer, B. and Levens, D. (1998) *Mol. Cell* **1**, 759–764
 - 29 DeMaria, C. T. and Brewer, G. (1996) *J. Biol. Chem.* **271**, 12179–12184
 - 30 Shamoo, Y., Abdul-Manan, N. and Williams, K. R. (1995) *Nucleic Acids Res.* **23**, 725–728
 - 31 Burley, S. K. and Roeder, R. G. (1996) *Annu. Rev. Biochem.* **65**, 769–799
 - 32 Herrick, G. and Alberts, B. (1976) *J. Biol. Chem.* **251**, 2133–2141
 - 33 Kahn, J. D., Yun, E. and Crothers, D. M. (1994) *Nature (London)* **368**, 163–166
-

Received 20 July 1998/9 November 1998; accepted 8 December 1998