Journal of Cellular Biochemistry 103:1013–1036 (2008)

Cellular Nucleic Acid Binding Protein Binds G-Rich Single-Stranded Nucleic Acids and May Function as a Nucleic Acid Chaperone

Pablo Armas, Sofía Nasif, and Nora B. Calcaterra*

División Biología del Desarrollo, IBR (Instituto de Biología Molecular y Celular de Rosario), Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)—Área Biología General, Dpto. de Ciencias Biológicas, Facultad de Ciencias Bioquímicas y Farmacéuticas, Universidad Nacional de Rosario, Suipacha 531, (S2002LRK) Rosario, Argentina

Abstract Cellular nucleic acid binding protein (CNBP) is a small single-stranded nucleic acid binding protein made of seven Zn knuckles and an Arg-Gly rich box. CNBP is strikingly conserved among vertebrates and was reported to play broad-spectrum functions in eukaryotic cells biology. Neither its biological function nor its mechanisms of action were elucidated yet. The main goal of this work was to gain further insights into the CNBP biochemical and molecular features. We studied *Bufo arenarum* CNBP (*b*CNBP) binding to single-stranded nucleic acid probes representing the main reported CNBP putative targets. We report that, although *b*CNBP is able to bind RNA and single-stranded DNA (ssDNA) probes in vitro, it binds RNA as a preformed dimer whereas both monomer and dimer are able to bind to ssDNA. A systematic analysis of variant probes shows that the preferred *b*CNBP targets contain unpaired guanosine-rich stretches. These data expand the knowledge about CNBP binding stoichiometry and begins to dissect the main features of CNBP nucleic acid targets. Besides, we show that *b*CNBP presents a highly disordered predicted structure and promotes the annealing and melting of nucleic acids in vitro. These features are typical of proteins that function as nucleic acid chaperones. Based on these data, we propose that CNBP may function as a nucleic acid chaperone through binding, remodeling, and stabilizing nucleic acids secondary structures. This novel CNBP biochemical activity broadens the field of study about its biological function and may be the basis to understand the diverse ways in which CNBP controls gene expression. J. Cell. Biochem. 103: 1013–1036, 2008.

Key words: cellular nucleic acid binding protein; CCHC zinc knuckle; RGG box; single-stranded nucleic acid binding; guanosine-rich; strand annealing promotion activity; nucleic acid melting promotion activity; zinc finger protein 9 ZNF9

Vertebrate genomes encode large multigene families of nucleic acid binding proteins. Although the main characteristics and common structural motifs have been extensively characterized for double-stranded DNA (dsDNA) binding proteins, no such extensive analysis has been made for single-stranded nucleic acid binding proteins. Even more, despite the identification of a growing number of proteins that bind specifically to single-stranded nucleic acids in vitro, the biochemical basis and biological relevance of the binding of these proteins to their putative cellular targets remain elusive.

Cellular nucleic acid binding protein (CNBP), also known as zinc-finger protein 9 (ZNF9), is a zinc-finger protein that binds single-stranded DNA (ssDNA) [Rajavashisth et al., 1989; Michelotti et al., 1995] and RNA [Yasuda et al., 1995; Pellizzoni et al., 1997; Calcaterra et al., 1999; Armas et al., 2004] molecules. CNBP cDNAs have been cloned from mammals [Michelotti et al., 1995; Yasuda et al., 1995; Flink and Morkin, 1995b; Shimizu et al., 2003], chicken [Ruble and Foster, 1998], fish [Armas et al., 2004; Liu and Gui, 2005], and amphibians

This article contains supplementary material, which may be viewed at the Journal of Cellular Biochemistry website at http://www.interscience.wiley.com/jpages/0730-2312/suppmat/index.html.

Grant sponsor: CONICET; Grant number: PIP 03073; Grant sponsor: ANPCyT; Grant number: PICT 01-8754.

*Correspondence to: Nora B. Calcaterra, IBR—CONICET, Área Biología General, Dpto. de Ciencias Biológicas, Facultad de Ciencias Bioquímicas y Farmacéuticas, Universidad Nacional de Rosario, Suipacha 531, (S2002LRK) Rosario, Argentina. E-mail: calcaterra@ibr.gov.ar

Received 22 December 2006; Accepted 12 June 2007

DOI 10.1002/jcb.21474

[Pellizzoni et al., 1997; Armas et al., 2001]. It is a small protein (≅19 kDa) with a highly conserved structural organization and amino acid sequence [Armas et al., 2001, 2004]. CNBP is made up of seven Cys-Cys-His-Cys (CCHC) zinc knuckle motifs (consensus sequence "C– ϕ –X–C–G– ± –X–G–H–X_3– δ –C", where X = variable amino acid, ϕ = aromatic amino acid, ± = charged amino acid, and δ = carbonyl-containing residue) that exhibit striking sequence and functional similarities with the corresponding structure of retroviral nucleocapsid (NC) proteins [Rajavashisth et al., 1989; McGrath et al., 2003]. CNBP also contains a peculiar arginine/glycine-rich region highly similar to the Arg-Gly-Gly (RGG) box motif located between the first and second CCHC motifs. The RGG box motif is composed of imperfect repeated Gly-Gly dipeptides interspersed with Arg and aromatic residues, and has been proposed as an RNA binding motif and a predictor of RNA binding activity [Burd and Dreyfuss, 1994]. CNBP shows several putative phosphorylation sites, and it was recently reported that zebrafish (*Danio rerio*) CNBP is in vitro phosphorylated by an embryonic cAMP-dependent protein kinase (PKA) activity [Lombardo et al., 2007].

A variety of biological functions have been assessed to CNBP. It has been largely implicated in vertebrate embryogenesis. During amphibian and fish development, CNBP is maternally inherited and remains in the cytoplasm until mid-blastula transition (MBT) becoming nuclear afterwards, once zygotic transcription has started [Armas et al., 2001, 2004]. Consequently, it was implicated in maternal and/or newly synthesized RNA binding, as well as in zygotic transcription modulation [Armas et al., 2004]. Recent works show that CNBP is required for forebrain formation during vertebrate organogenesis. *Cnbp*-null mutant mice are embryonic lethal and show severe forebrain truncation and facial abnormalities due to a lack of proper morphogenetic movements of the anterior visceral endoderm (AVE) during pregastrulation stage [Chen et al., 2003]. In chick embryos, *Cnbp* is expressed in the equivalent tissues of the mouse embryo and, furthermore, CNBP knockdown produces forebrain truncation [Abe et al., 2006]. It was proposed that the CNBP role during vertebrate organogenesis is to regulate the forebrain formation by controlling the expression of a number of rostral head transcription factors, such as *c-Myc* [Chen et al., 2003], *BF-1*, *Six3*, and *Hesx1* [Abe et al., 2006]. Indeed, a recent report proposes that CNBP has an essential role in mediating neural crest expansion by controlling proliferation and cell survival during vertebrate rostral head development [Weiner et al., 2007].

Besides, a broad-spectrum of cellular functions has been assigned to CNBP. Binding to DNA, it was reported both as a negative [Rajavashisth et al., 1989; Flink and Morkin, 1995a; Liu et al., 1998] and positive [Michelotti et al., 1995; Konicek et al., 1998] transcriptional regulator. Interacting with RNA, it was reported as a translational [Pellizzoni et al., 1997, 1998; Calcaterra et al., 1999] and transcriptional modulator [Yasuda et al., 1995]. Two main models were proposed wherein CNBP acts as a single-stranded nucleic acid binding protein. The first model involves CNBP and heterogeneous nuclear ribonuclear protein K (hnRNP K) as transcriptional activators of the *c-Myc* gene through binding to the CT promoter element [Michelotti et al., 1995; Tomonaga et al., 1998; Krecic and Swanson, 1999]. The second model proposes CNBP as a translational inhibitor of ribosomal proteins mRNAs (rp-mRNAs) through binding to their 5′ UTR. In this last model, La autoantigen would be a translational activator while the Ro60 autoantigen would alternatively assist the binding of CNBP or La to rp-mRNAs [Pellizzoni et al., 1998].

Scant information exists about CNBP biochemical activities and mechanism of action, a fact that makes it difficult to explain how a unique protein may act in such a wide range of cellular and biological pathways. With the aim of further understanding the CNBP ways of action, we analyzed CNBP binding to RNA and DNA probes in detail. In this study we show that *Bufo arenarum* CNBP (*b*CNBP) binds to RNA and ssDNA, but not to dsDNA. Noteworthy, ssDNA is bound by CNBP monomeric and dimeric forms, whereas RNA is mainly bound by the CNBP dimer. Moreover, our results suggest that monomer and dimer are non-interconvertible forms. A systematic analysis of variant probes showed that the preferred CNBP targets are single-stranded nucleic acid with unpaired G-rich stretches. Besides, prediction structure analysis in addition to results from annealing and melting assays suggest that CNBP functions as a nucleic acid chaperone. This novel CNBP biochemical activity may explain the disparate functions assigned to this

protein. The biological significance of these results is discussed in the context of the possible CNBP cellular roles, such as the control of ribosomal proteins translation and *c-Myc* protooncogene transcription.

MATERIALS AND METHODS

Recombinant Protein

*b*CNBP cDNA, GenBank Accession Number AF144698 [Armas et al., 2001], was fused to an N-terminal glutathione-S-transferase (GST) tag in pGEX-2T plasmid vector (Amersham Biosciences). GST-*b*CNBP fusion was expressed in *Escherichia coli* DH5α and purified to homogeneity by affinity chromatography using Gluthathione Sepharose (Amersham Biosciences) according to the manufacturer's instructions. Purification quality was determined by SDS–PAGE and silver staining. Fusion protein molar concentration was accurately estimated by densitometric analysis of SDS–PAGE Coomassie Brilliant Blue stained gels.

Disorder Prediction

Prediction of intrinsically unstructured regions was carried out using the DisProt VL3-E online predictor (http://divac.ist.temple.edu/disprot/predictor.php) [Peng et al., 2005]. An amino acid with a disorder score above or equal to 0.5 is considered to be in a disordered environment, while below 0.5 to be ordered.

Embryonic and Egg Extracts Preparation

B. arenarum specimens were collected in the surroundings of Rosario, Argentina, and kept in moist chambers at 15°C until used. Ovulations, fertilizations, and embryo culture were performed as reported elsewhere [Calcaterra et al., 1999]. Eggs and embryos extracts were prepared using a Potter–Elvehjem homogenizer, in four volumes of a solution containing 20 mM Tris-HCl, pH 7.6, 10 mM $MgCl_2$, 50 mM KCl, 100 mM NaCl, 1 mM EDTA, 0.1% Triton X-100, and 100 μM PMSF. Cellular debris was removed by centrifugation at 20,000*g* and 4°C for 30 min. Extracts were centrifugated at 100,000*g* for 1 h, and supernatants were used as total soluble protein extracts.

Nucleic Acid Probes

DNA and RNA oligodeoxinucleotides (see sequences in Table I) were purchased from Invitrogen Corp. and dissolved in distilled water to a final concentration of 100 μM. Labeled ssDNA probes were [^{32}P]-5′ end-labeled DNA oligonucleotides. Long RNA probes (sense, RNA-L4-UTR, and antisense, aRNA-L4-UTR) were synthesized by in vitro transcription of the 5′ UTR from *X. laevis* L4 rp-mRNA sequence using as template a plasmid containing the *X. laevis* L4 cDNA, GenBank Accession Number X05216. For labeled RNA-L4-UTR probe synthesis, the in vitro transcription reaction was performed in the presence of [α-^{32}P] UTP. Labeled ssDNA and RNA probes were purified by 12% and 8% PAGE, respectively, in 1X TBE. Probes were recovered by elution in 0.3 M sodium acetate, 0.1% SDS, for 4 h at 37°C, followed by phenol/chloroform extractions and ethanol precipitation, and finally dissolved in sterile H_2O. Labeled single-stranded probes were denatured by heating to 95°C for 5 min and renatured by slowly cooling to room temperature in a buffer containing 100 mM Tris-HCl, pH 8.0, 1 mM EDTA, 100 mM NaCl. dsDNA probes were obtained by incubating equimolar amounts of the completely or partially complementary purified ssDNA oligonucleotides at 95°C for 5 min, and then slowly cooling the mixture to allow duplex formation. Secondary structures of single-stranded nucleic acids were predicted using *RNAstructure 4.11* and *Mfold* softwares [Mathews et al., 1999; Zuker, 2003].

Electrophoretic Mobility Shift Assay (EMSA)

Binding reactions were performed in 20 mM HEPES, pH 8.0, 100 mM NaCl, 10 mM $MgCl_2$, 1 mM EDTA, 1 mM DTT, 1 μg/μl BSA and 10% glycerol. Labeled probes were added to a final concentration of 2 nM. Purified fusion proteins or total soluble protein extracts were added as indicated. For ssDNA probes, 0.01 μg/μl of non-specific competitor DNA (poly(dI:dC)·(dC:dI) from Pierce Nucleic Acid Technologies) and 0.5 μg/μl Heparin were added. For RNA probes, 1 μg/μl of non-specific competitor RNA (Type VI Torula Yeast RNA from Sigma) and 0.5 U/μl RNasin (Promega) were added. Final reaction volume was 20 μl. Binding reactions were incubated for 30 min at 37°C and then loaded, depending on the probe size, onto 5%, 8%, or 12% polyacrylamide gels containing 5% glycerol in TBE 0.5×. After electrophoresis, gels were dried and exposed to X-Ray films (Kodak

TABLE I. Oligonucleotides Used in This Work

Name	Length (b)	Sequence (5' to 3')
sRNA-L4-UTR	43	CCUUUUCUCUUCGUGGCCGCUGUGGAGAAGCAGCGAGGAGAUG
DNA-L4-UTR	43	CCTTTTCTCTTCGTGGCCGCTGTGGAGAAGCAGCGAGGAGATG
DNA-L4-UTR-1	43	CCTTTTGTGTTGCTGGCCGCTGTGGAGAAGCAGCGAGGAGATG
DNA-L4-UTR-2	43	CCTTTTGTGTTGCTGGCCGCTGTGGAGAAGCAGGCAGCACATG
DNA-L4-UTR-3	43	CCTTTTCTCTTCGTGGCCGCCGTGGAGAAGCAGCGAGGAGATG
DNA-L4-UTR-4	43	CCTTTTCTCTTCGTGGCCGCCGTGGGGAAGCAGCGAGGAGATG
DNA-L4-UTR-5	37	CCTTTTCTCTTCGTGGCCGC_____GAAGCAGCGAGGAGATG
DNA-L4-UTR-6	29	CCTTTTCTCTTCGTGG_____CAGCGAGGAGATG
DNA-L4-UTR-7	33	CCTTTTCTCTTCGT_____CTGTGGAG_____GCGAGGAGATG
DNA-L4-UTR-8	39	CCTTTTCTCTTCGTG__CGCTGTGGAGAA__AGCGAGGAGATG
DNA-L4-UTR-F1	22	_____GTGGCCGCTGTGGAGAAGCAGC_____
DNA-L4-UTR-F2	26	_____GGCCGCTGTGGAGAAGCAGCGAGGAG___
DNA-L4-UTR-1-M1	43	GGTTGGGTGGGGCTGGCCGCTGTGGAGAAGCAGCGAGGAGATG
DNA-L4-UTR-1-M2	43	GGGTGGGTGGGGCGGGCCGCTGTGGAGGAGCAGCGAGGGGGGG
DNA-L4-UTR-1-M3	43	GGGTGAGTGGAGCGGGCCGCTGTGGGGGGGCAGCGAGGGGGGG
CT	45	CACCCTCCCCACCCTCCCCATAAGCGCCCCTCCCGGGTTCCCAAA
CT-12	12	CACCCTCCCCAC_____
Comp-CT	45	TTTGGGAACCCGGGAGGGGCGCTTATGGGGAGGGTGGGGAGGGTG___
Comp-CT-1	45	TTTGGGAACGGGGAGGGGCGCTTATGGGGAGGGTGGGGAGGGTG___
Comp-CT-2	45	TTTGGGAACGGGGAGGGGCGCTTATCCGGAGGGTGGGGAGGGTG___
Comp-CT-M2	45	TTTGAGAACTCGGGGGGGGGTGCTTACGGGGGGGGGTAGGGAAGGTG___
Comp-CT-37	37	____GGAACCCGGGAGGGGCGCTTATGGGGAGGGTGGGGAG_____
Comp-CT-23	23	_____ACCCGGGAGGGGCGCTTATGGGG_____
Comp-CT-12	12	_____GTGGGGAGGGTG___
Comp-CT-18	18	_____AGGGTGGGGAGGGTGGGG
Comp-CT-18-M	18	_____AGGGTGAGGAGAGTGGGG
TBA	15	GGTTGGTGTGGTTGG
T30695	16	GGGTGGGTGGGTGGGT
A-25	25	AAAAAAAAAAAAAAAAAAAAAAAAA
C-25	25	CCCCCCCCCCCCCCCCCCCCCCCCC
G-25	25	GGGGGGGGGGGGGGGGGGGGGGGGG
T-25	25	TTTTTTTTTTTTTTTTTTTTTTTTT
non-specific competitor (nsc)	34	AAGGATCCATGAGCAGCAATGAGTGCTTCAAGTG

Families of nucleotides are aligned and changes of sequences in respect of the first oligonucleotide in each group are signaled: CT elements are boxed, nucleotide replacements are bolded and underlined, and nucleotide deletions are indicated by low dashes. CT and Comp-CT as well as CT-12 and Comp-CT-12 are exactly complementary oligonucleotide pairs.

BioMax MS film) or Storage Phosphor Screen that was subsequently scanned in a STORM 860 PhosphorImager using ImageQuant 5.2 software (Amersham Biosciences). For potassium condition EMSAs the probes were denatured and renatured in a potassium-buffer (100 mM Tris-HCl, pH 8.0, 1 mM EDTA, 100 mM KCl) and NaCl in the reactions was replaced by KCl.

Super-Shift Assays

Binding reactions were performed as detailed in EMSA. Antibody used for super-shift assays was raised in rabbits by immunization with purified recombinant *b*CNBP and affinity purified as reported elsewhere [Armas et al., 2004]. One microgram of purified anti-*b*CNBP antibody

was added before or after binding reactions took place and incubated at 4°C for 30 min. Control reactions were performed using equal volumes of a pre-immune serum subjected to the same affinity antibody purification protocol carried out for obtaining anti-CNBP antibody from immune serum.

Apparent Dissociation Constants Estimation

Apparent dissociation constants (K_d) were estimated from the intensity of radioactive bands in EMSAs, which enables the estimation of accurate equilibrium-binding constants [Carey, 1991]. Relative amounts of bound and unbound probes were quantified from radioactive bands, and the bound/unbound ratio was used as an estimation of the [complex]/[free probe] ratio. Free (unbound) protein was estimated by the total protein concentration considering an excess of protein in respect of total probe higher than tenfold. Apparent K_d was estimated from the slope of the linear regression of the free protein concentration versus the [complex]/[free probe] ratio plot for each EMSA using SigmaPlot 9.0 software (SPSS, Inc.). Five independent experiments were carried out and an average apparent K_d value was estimated.

UV Cross-Linking Reactions

Binding reactions were performed as detailed in EMSA. After incubation, reactions were irradiated for 10 min with UV light (2×10^5 erg/mm^2) using a 254 nm UV lamp at 4°C. Proteins were resolved by native 8% PAGE containing 5% glycerol in TBE 0.5× or by 12% SDS–PAGE, and visualized by staining with Coomassie Brilliant Blue. Gels were dried, exposed, and scanned as described above.

Annealing Assays

CT (sense) and [^{32}P]-5′ end-labeled Comp-CT (antisense, Comp-CT*) oligonucleotides were separately heated at 90°C for 3 min, and then transferred to ice for 5 min. Then, Comp-CT* (5 nM) and CT (10 nM) oligonucleotides were independently incubated in annealing buffer (50 mM Tris-HCl, pH 8.0, 0.1 mM EDTA, 1 mM DTT, 6 mM $MgCl_2$, 80 mM KCl, and 0.1 mM $ZnCl_2$) at 37°C for 2 min in the absence or in the presence of the indicated recombinant protein. Finally, to start the reactions, 20 μl of CT-containing reaction mixture were added to 100 μl of reaction mixture containing Comp-CT* and the mixture was incubated at 37°C. Aliquots of 20 μl were taken at specific time points as indicated and annealing reactions were stopped by adding them to 7.5 μl of stop solution (0.25% bromphenol blue, 0.25% xilene cyanol, 20% glycerol, 20 mM EDTA, 0.2% SDS, and 0.4 mg/ml yeast tRNA). Reactions were incubated in stop solution at 37°C for 1 min before being transferred to ice, and then subjected to 15% native PAGE in TBE 1X. Gels were dried, exposed, and scanned as described above. Radioactive bands were quantified and annealing percentage was determined by dividing the amount of annealed oligonucleotide (ds) by the total amount (annealed plus single-stranded (ss)) in each lane, and multiplying by 100 (% annealing $= 100 \times ds/(ds + ss)$). Statistically significant differences were analyzed using SigmaStat 3.11 software (SPSS, Inc.). Oligonucleotides mobility controls were the not annealed Comp-CT* probe alone, and the completely annealed Comp-CT* probe heated and slowly cooled in the presence of unlabeled CT probe.

Melting Assays

For these assays we used pairs of oligonucleotides that form perfect or partially mismatched duplexes. One of these oligonucleotides was [^{32}P]-5′ end-labeled and the pre-annealed duplexes (10 nM) were mixed in annealing buffer with the indicated amount of protein on ice. Tubes were incubated for 5 min at the indicated temperatures in an Eppendorf Thermocycler (Mastercycler Personal). After the incubation at each temperature, a 5-μl aliquot was removed and mixed with 5 μl of ice-cold stop solution. Electrophoretical analysis of the generated product was done as described above using 15% native PAGE for 45-nucleotide-length duplexes or 25% native PAGE for 12-nucleotide-length duplex. Gels were dried, exposed, and scanned as described above. Radioactive bands were quantified and melting percentage was determined by dividing the amount of melted oligonucleotide (ss) by the total amount (annealed (ds) plus melted (ss)) in each lane, and multiplying by 100 (% melting $= 100 \times ss/(ds + ss)$). Melting temperatures (T_m) were calculated as the temperature corresponding to 50% melting from the sigmoidal adjustment (SigmaPlot 9.0 software, SPSS, Inc.) of the melting curves representing melting percentages versus temperature. Oligonucleotides mobility controls were the not annealed

labeled oligonucleotide alone (ss), and the completely annealed labeled oligonucleotides (ds, T = 4°C).

RESULTS

CNBP Binds to *X. Laevis* L4 rp-mRNA 5′ UTR Forming a Unique Protein–RNA Complex

One of the proposed CNBP functions is the rp-mRNAs translational inhibition through its binding to rp-mRNA 5′ UTRs [Pellizzoni et al., 1998]. Although CNBP binding to RNA targets was previously reported [Yasuda et al., 1995; Pellizzoni et al., 1997, 1998; Calcaterra et al., 1999; Armas et al., 2004], scant information exists concerning CNBP binding behavior and biochemical properties. Therefore, we further studied the patterns and biochemical requirements of *b*CNBP binding to synthetic probes that contain the 5′ UTR sequence from the *X. laevis* L4 rp-mRNA.

EMSAs were carried out using a recombinant GST fusion to *B. arenarum* CNBP (GST-*b*CNBP) and a 90-mer in vitro synthesized RNA (RNA-L4-UTR), which contains the complete (43-nucleotides-long) 5′ UTR from the *X. laevis* L4 rp-mRNA. GST-*b*CNBP was unable to shift RNA-L4-UTR probe in vitro by itself (Fig. 1A), suggesting that some cellular factors could be needed for the binding. As CNBP is

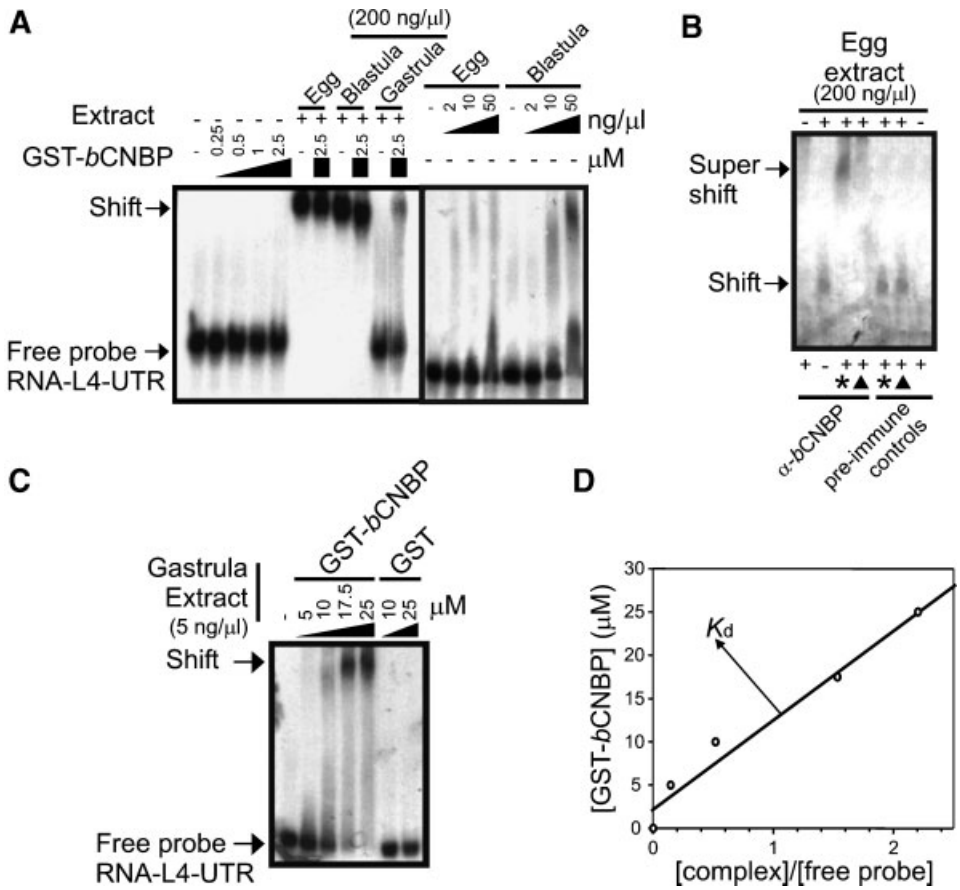

Fig. 1. *b*CNBP binding to L4 5′ UTR RNA. EMSAs using labeled RNA containing the 5′ UTR sequence from *X. laevis* L4 rp-mRNA (RNA-L4-UTR) as a probe. **A**: EMSA carried out with increasing amounts of GST-*b*CNBP. Embryonic extracts were added alone or in combination with the highest GST-*b*CNBP concentration. Embryonic extracts from egg, blastula, and gastrula stages were used in a final concentration of 200 ng/µl (left gel) or 2, 10 and 50 ng/µl (right gel). Free and shifted probes are indicated by arrows at the left of the figure. **B**: Super-shift EMSA performed with egg extract, anti-*b*CNBP antibody (α-*b*CNBP), and both of them. The antibody was pre-incubated with egg extract before carrying out (▲) or after finishing (*) the binding reaction. Shifted and super-shifted probes are indicated by arrows at the left of the figure. Control reactions were performed using equal volumes of an affinity purified pre-immune serum subjected to the same affinity antibody purification protocol carried out for obtaining anti-CNBP antibody from immune serum. **C**: EMSA performed with RNA-L4-UTR probe and increasing amounts of GST-*b*CNBP. GST was used as control and 5 ng/µl of gastrula extract was added to reaction mixtures. Free and shifted probes are indicated by arrows at the left of the figure. **D**: Apparent K_d calculation plot, linear regression of the free protein concentration versus the [complex]/[free probe] ratio plot for the EMSA shown in (C).

required for embryonic development [Chen et al., 2003; Abe et al., 2006; Weiner et al., 2007] and shows a differential subcellular localization depending on the developmental stage [Armas et al., 2001, 2004], embryonic extracts from stages before, during, and after MBT were added to binding reaction mixtures. For egg and blastula extracts, band-shifts were observed even in absence of GST-*b*CNBP (Fig. 1A). A unique complex was observed, even in the presence of lower non-saturating extract concentrations which leave some unshifted probe (Fig. 1A, right side). Super-shift assays were performed to address whether embryonic *b*CNBP was actually forming part of the RNA–protein complex. Figure 1B shows that the complex formed with egg extract was super-shifted by anti-*b*CNBP antibody added at the end of the binding reaction, while complex formation was completely impaired when the anti-*b*CNBP antibody was pre-incubated with the egg extract. In this experiment, the gel was over-run to obtain maximal resolution of the shifted and super-shifted species, thus causing the loss of the free probe. These results confirmed the presence of *b*CNBP in the EMSA complexes formed by embryonic protein extracts with the RNA-L4-UTR probe.

It was noticeable that proteins from gastrula extract were not able to bind RNA-L4-UTR per se but, instead, this embryonic extract enabled GST-*b*CNBP to bind to RNA probe causing a shift. The RNA-L4-UTR shift might be due to an indirect CNBP-RNA interaction through potential cofactor(s) present in gastrula extract. However, gastrula extract did not form a stable complex detectable by EMSA unless exogenous *b*CNBP was added. Taken together, our results suggest that CNBP is not only present in RNA–protein complexes, but also that its presence is essential for complex formation. The effect of gastrula extract was observed still when a 40-fold dilution was used (Fig. S1, Supplementary Material), suggesting that some catalytic cofactor(s) are involved in the complex formation. This finding agrees with a previous report that showed that the Ro60 autoantigen acts catalytically in the alternative interaction of La and CNBP with the 5′ UTR of L4 rp-mRNA [Pellizzoni et al., 1998]. The lowest gastrula extract concentration tested that allowed the shift-effect was 5 ng/μl (Fig. S1, Supplementary Material). Consequently, this concentration was used for the following GST-*b*CNBP RNA-L4-UTR binding experiments. GST-*b*CNBP binding to RNA-L4-UTR probe yielded a unique complex, which increased in intensity along with the increment of GST-*b*CNBP (Fig. 1C). GST failed to bind RNA-L4-UTR probe proving that *b*CNBP itself was involved in the interaction with the nucleic acid probe. The sequence-specificity of GST-*b*CNBP binding was demonstrated by the presence of large excess of nucleic acid competitor in EMSAs (see Materials and Methods Section) and by competition with unlabeled specific and non-specific probes (Fig. S2, Supplementary Material). These facts allowed us to quantify binding affinity by estimation of the apparent dissociation constant (K_d) for the equilibrium-binding reaction (Fig. 1D). Apparent K_d estimated values were completely reproducible for independent experiments. An apparent K_d was calculated for each of five independent EMSAs and the averaged apparent K_d value was 9.5 ± 2.8 μM (mean of $K_d \pm$ SEM, $n = 5$).

CNBP Binds ssDNA Representing *X. Laevis* L4 rp-mRNA 5′ UTR Forming Two Protein–DNA Complexes

The *b*CNBP ssDNA binding capability was further explored by using ssDNA probes representing the 5′ UTR sequence from *X. laevis* L4 rp-mRNA (DNA-L4-UTR; see Table I) in EMSAs. This oligonucleotide had previously been used as a probe to evaluate CNBP nucleic acid binding activity [Pellizzoni et al., 1997; Armas et al., 2004; Lombardo et al., 2007]. GST-*b*CNBP bound to DNA-L4-UTR probe (Fig. 2A) with a two-complex binding pattern and there was no need to add embryonic extracts for the binding, indicating that GST-*b*CNBP was able to bind per se the DNA-L4-UTR probe. As GST did not cause the shift of the labeled DNA-L4-UTR probe, the binding could be assessed to *b*CNBP itself. The sequence-specificity of *b*CNBP binding was demonstrated by the presence of a vast excess of competitor nucleic acid in ssDNA-EMSAs (see Materials and Methods Section), and by competition with unlabeled specific and non-specific probes (Fig. S3, Supplementary Material).

The ssDNA binding capability of embryonic *b*CNBP was analyzed using the DNA-L4-UTR probe. A two-complex binding pattern was also observed when embryonic extracts were analyzed in EMSAs (Fig. 2B, third lane). Both complexes showed higher electrophoretic

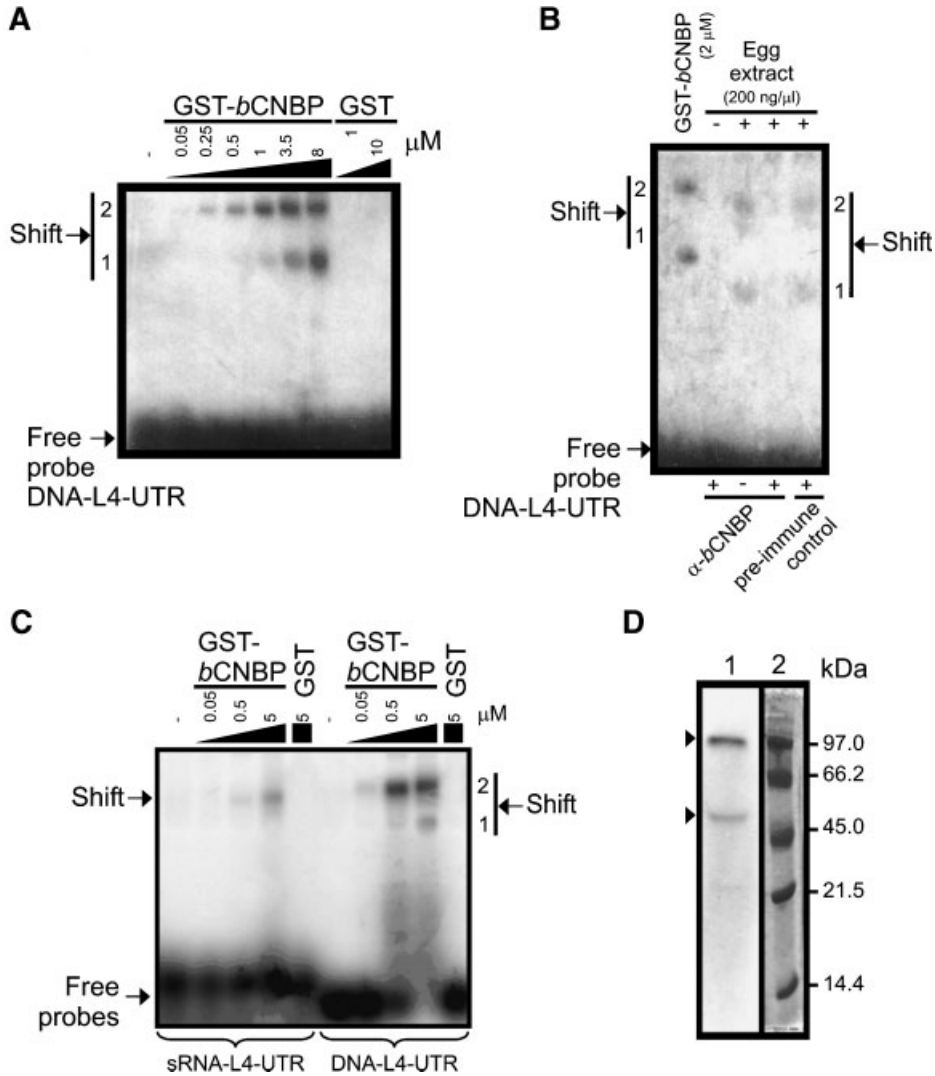

Fig. 2. *b*CNBP binding to L4 5′ UTR ssDNA. **A**: EMSA performed with increasing amounts of GST or GST-*b*CNBP and a labeled ssDNA probe containing the complete sequence of 5′ UTR from *X. laevis* L4 rp-mRNA (DNA-L4-UTR). Shifted bands are indicated as 1 (higher mobility) or 2 (lower mobility). No embryonic extract was added to reaction mixtures. **B**: EMSA carried out with egg extract (**lane 3**), anti-*b*CNBP antibody (α-*b*CNBP, **lane 2**), and both of them (**lane 4**). α-*b*CNBP was pre-incubated with egg extract before the binding reaction was carried out. GST-*b*CNBP binding reaction was used as a mobility control (**lane 1**). Control reaction was performed using an affinity purified pre-immune serum subjected to the same affinity antibody purification protocol carried out for obtaining anti-CNBP antibody from immune serum. Shifted bands are indicated as 1 (higher mobility) or 2 (lower mobility) at the left of the figure for GST-*b*CNBP, and at the right of the figure for embryonic proteins. **C**: EMSA performed with increasing amounts of GST-*b*CNBP and the labeled short RNA probe containing the complete sequence of 5′ UTR from *X. laevis* L4 rp-mRNA (sRNA-L4-UTR) or the labeled DNA-L4-UTR probe (right). Shifted bands are indicated as 1 (higher mobility) or 2 (lower mobility) at the right of the figure. GST was used as a control. No embryonic extract was added to reaction mixtures. **D**: Autoradiography of the SDS–PAGE of the cross-linked binding reaction performed with labeled DNA-L4-UTR probe and 5 μM GST-*b*CNBP (**lane 1**). In this reaction no embryonic extract was added. Radioactive bands are indicated by arrow tips. Molecular mass markers (**lane 2**) were stained with Coomassie Brilliant Blue.

mobilities than those formed by recombinant GST-*b*CNBP (Fig. 2B, first lane), probably due to the presence of the GST tag in the recombinant protein. All embryonic extracts tested were able to shift the probe and the two-complex binding pattern was detected independently of the developmental stage analyzed (Fig. S4, Supplementary Material). The formation of both complexes was impaired by the addition of affinity-purified anti-*b*CNBP antibody (Fig. 2B, fourth lane), suggesting that their formation depended on embryonic *b*CNBP. The combina-

tion of GST-*b*CNBP with embryonic extracts rendered several shift bands coincident with the sum of the complexes formed by recombinant and embryonic *b*CNBP independently (Fig. S5, Supplementary Material).

As RNA-L4-UTR (90-nucleotide-long) probe is longer than DNA-L4-UTR (43-nucleotide-long), it could be possible that the differences in CNBP binding pattern depended on the length of the probes. To test this, we performed EMSAs with ssDNA or RNA probes, but using a short version of RNA-L4-UTR probe (sRNA-L4-UTR, see Table I) that only contains the complete 43-nucleotides-long 5′ UTR sequence from *X. laevis* L4 rp-mRNA. Figure 2C shows that the electrophoretic mobility of the main complex formed with the RNA probe was similar to the CNBP–ssDNA complex that shows the lower electrophoretic mobility.

To further establish the complexes stoichiometry, SDS–PAGE of the cross-linked GST-*b*CNBP and DNA-L4-UTR complexes was carried out. Two bands of apparent M_r of 50,000 and 100,000 were observed, which were coincident with the expected M_r for GST-*b*CNBP monomer and dimer respectively (Fig. 2D). As the binding assay was performed without addition of embryonic protein extract, this result indicates that complexes observed in EMSAs corresponded to GST-*b*CNBP monomer and dimer bound to DNA-L4-UTR. It is worth mentioning that similar results were observed using a recombinant *b*CNBP fused to an His_6 tag (not shown) as well as zebrafish CNBP [Armas et al., 2004].

When DNA-L4-UTR was used as molecular probe, the dimeric complex was detected at lower GST-*b*CNBP concentrations than those needed to detect the monomeric complex (Fig. 2A), suggesting that the GST-*b*CNBP dimer may have higher affinity for DNA-L4-UTR probe than the monomer. However, the apparent K_d values calculated for GST-*b*CNBP monomer and dimer binding to DNA-L4-UTR did not show significant differences ($K_d1 = 11 \pm 1.9$ μM for the monomer and $K_d2 = 4.9 \pm 1.4$ μM for the dimer; results are expressed as a mean of the $K_d \pm$ SEM, $n = 5$). Moreover, these K_d values were similar to those calculated for the RNA probes.

Results presented here indicate that CNBP possess similar affinities for RNA and ssDNA with identical sequences, but different nucleic acid binding pattern according to the nature of the target molecule. The differential behavior may be due to the fact that RNA and ssDNA probes show identical theoretically predicted secondary structures but different calculated stabilities (Fig. 3A). To analyze this possibility, EMSA competition assays were performed to compare RNA and ssDNA binding features. The binding of GST-*b*CNBP to labeled DNA-L4-UTR probe was competed with increasing amounts of different unlabeled probes, and the disappearance of shifted species was evaluated (Fig. 3B). Addition of 1,000-fold excess of the unlabeled DNA-L4-UTR probe competed for the GST-*b*CNBP monomer binding to the labeled DNA-L4-UTR, whereas little or no effect was observed on the dimer complex formation. This fact suggests that the *b*CNBP dimer may be present at higher amounts. On the other hand, addition of 1,000-fold excess of unlabeled RNA-L4-UTR or sRNA-L4-UTR probes competed for the GST-*b*CNBP dimer binding to the labeled DNA-L4-UTR probe, with little or no consequence in the monomer complex formation. The addition of excess of unlabeled non-specific probes did not change the binding pattern. In the case of the competition with RNA-L4-UTR probe, 5 ng/μl of gastrula extract was added to the reaction mixtures in order to allow GST-*b*CNBP binding to this probe. This result clearly shows the sequence-specificity of the binding and confirms that the main complex observed in RNA-EMSAs corresponds to GST-*b*CNBP dimer bound to the RNA-L4-UTR probe. Taking into account that *b*CNBP dimer binding to DNA-L4-UTR was more effectively competed by RNA-L4-UTR than by DNA-L4-UTR, it seems possible that the *b*CNBP dimer prefers RNA instead of ssDNA targets. Moreover, these results set out the possibility that GST-*b*CNBP dimers and monomers are non-interconvertible forms that bind nucleic acid targets as preformed species.

CNBP Preferentially Targets Single-Stranded Nucleic Acids Containing G-Rich Unpaired Stretches Constrained by an Organized Sequence Environment

A systematic analysis of different oligonucleotide probes was performed to elucidate the main structural features of CNBP targets. Several oligonucleotides were designed by nucleotide replacement or deletion mutagenesis of the DNA-L4-UTR target. Two base-replacement mutant probes were designed to disrupt

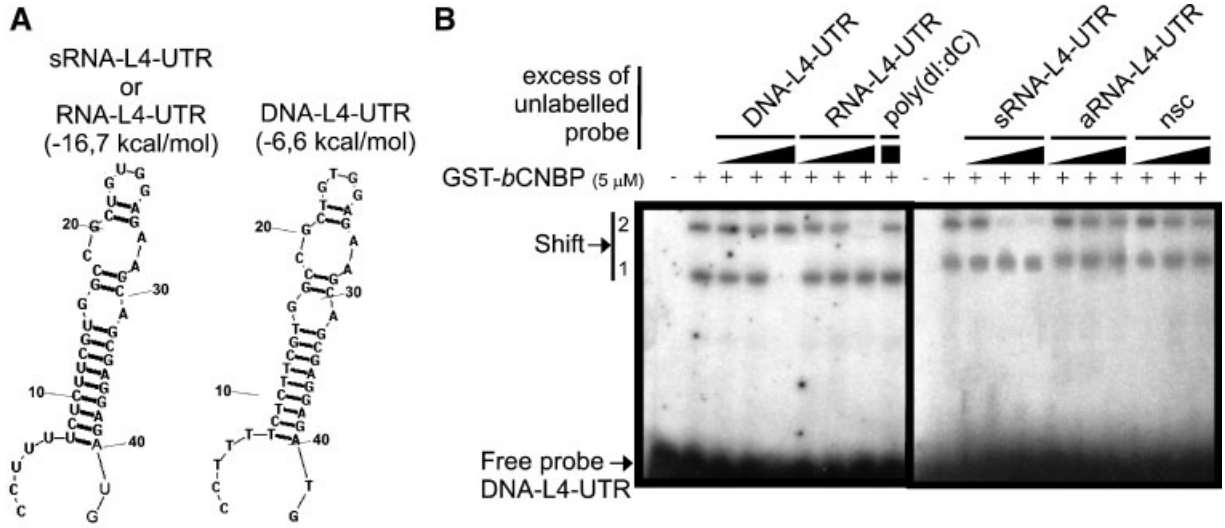

Fig. 3. Comparison of *b*CNBP binding to L4 5′ UTR ssDNA and RNA. **A**: Predicted secondary structures and stabilities of RNA and ssDNA probes representing the 5′ UTR from *X. laevis* L4 rp-mRNA. **B**: Competition EMSA using labeled DNA-L4-UTR probe and 5 μM GST-*b*CNBP. Competitions were done with 25-, 100-, and 1,000-fold excess of unlabeled probes: DNA-L4-UTR, RNA-L4-UTR, poly (dI:dC), sRNA-L4-UTR, antisense RNA-L4-UTR (aRNA-L4-UTR), and non-specific competitor (nsc) as indicated. In the case of unlabeled RNA-L4-UTR probe, 5 ng/μl of gastrula extract was added to the reaction mixture. Shifted bands are indicated by arrows as 1 (monomer) or 2 (dimer) at the left of the figure.

stretches of secondary structure (DNA-L4-UTR-1 and 3; see Table I and Fig. 4A), while other two mutants were designed to compensate them and restore the original secondary structure (DNA-L4-UTR-2 and 4 respectively; see Table I and Fig. 4A). In addition, four deletion mutant probes were designed to eliminate single-stranded loops and bubbles deduced from DNA-L4-UTR secondary structure (DNA-L4-UTR-5 to 8; see Table I and Fig. 4B). As the DNA-L4-UTR-1 mutant probe showed a novel secondary structure (Fig. 4C), two additional deletion mutant probes were designed in order to eliminate the single-stranded ends present in it (DNA-L4-UTR-F1 and -F2; see Table I and Fig. 4C). DNA-L4-UTR-1 was the only probe bound by GST-*b*CNBP with a significant higher affinity (tenfold lower K_d values) than DNA-L4-UTR (Figs. 4A and 5). The compensating mutation in the DNA-L4-UTR-2 probe restored the original binding affinity. DNA-L4-UTR-5 and -6 were bound by GST-*b*CNBP with significant lower affinities (tenfold higher K_d values) than DNA-L4-UTR (Figs. 4B and 5). DNA-L4-UTR-F1 and -F2 were not bound by GST-*b*CNBP (Fig. 4C). A direct relationship between the CNBP binding affinities and the number of unpaired guanosines in the probes was detected (Fig. 5, top chart). To further explore this, a set of artificial probes was designed with the aim of increasing the number of unpaired guanosines (DNA-L4-UTR-1-M1, -M2, and -M3; see Table I). It is important to note that these probes differ from the DNA-L4-UTR-1 only in the number of unpaired guanosines, conserving the secondary structure and stability. EMSAs showed that the rise in the number of unpaired guanosines was directly related with a statistically significant increment of *b*CNBP binding affinity, reaching a plateau above 18 unpaired guanosines (Figs. 4D and 5).

On the other hand, to test whether the guanosine nucleobase preference was dependent on the sequence complexity and/or secondary structure, a set of four homo-oligonucleotidic probes (A-25, C-25, G-25, T-25, see Table I) was assayed. It is worth mentioning that these probes did not show stable predicted secondary structures. EMSAs showed very weak and probably unspecific interactions with C-25, G-25, and T-25 probes. There was no detectable interaction with the A-25 probe (Fig. S6, Supplementary Material). These results indicate that the *b*CNBP preference for unpaired guanosine nucleobases depends on the sequence and structural complexity of the target. Taken together, it would be possible to propose that the observed binding affinities differences were due to a strong CNBP binding to guanosine-rich single-stranded stretches, which

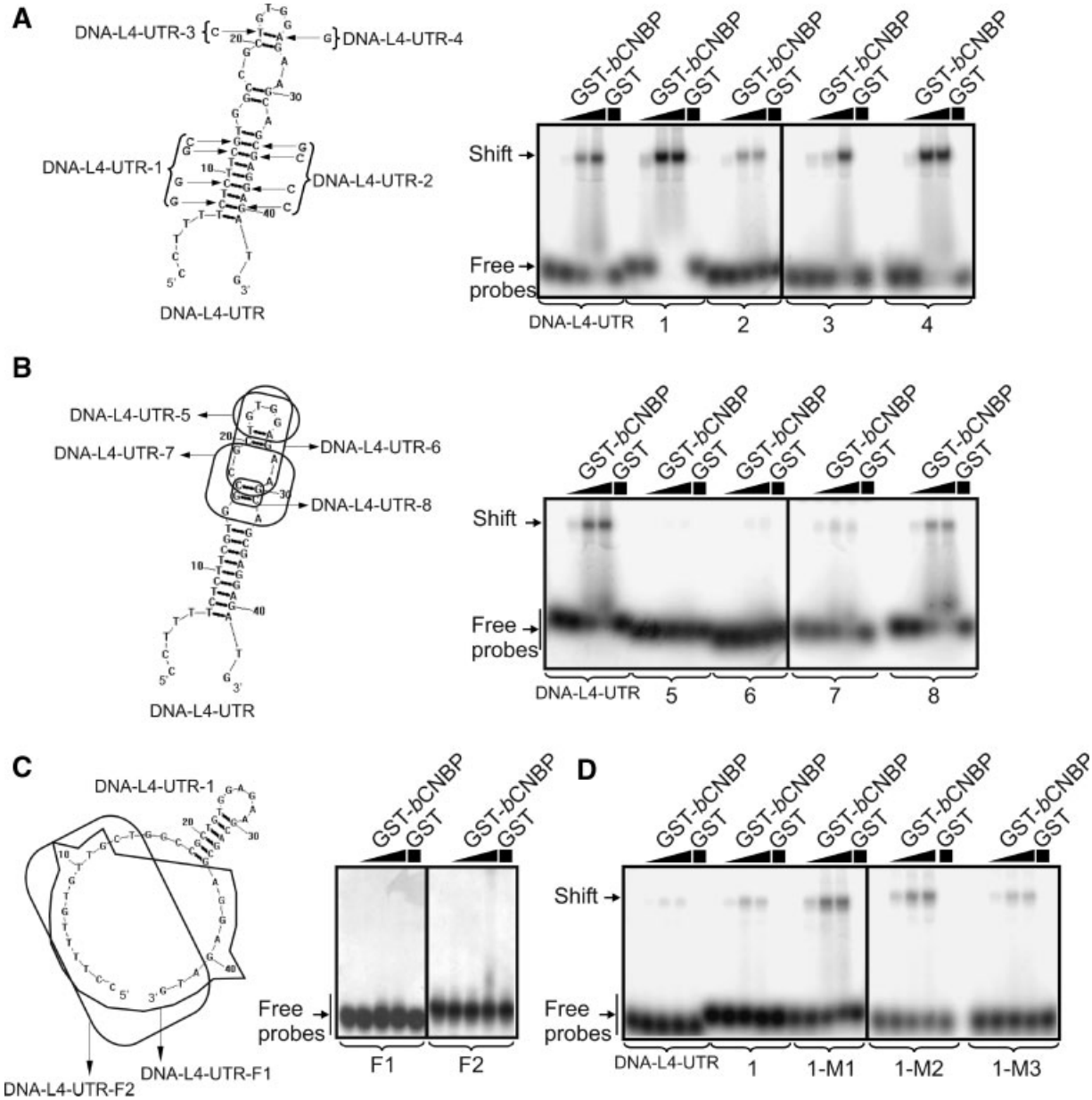

Fig. 4. *b*CNBP binding to L4 5′ UTR mutant probes. **A**: **Left**, DNA-L4-UTR base-replacement mutant probes. Arrows indicate the replaced nucleotides in each mutant probe. Base-replacements in DNA-L4-UTR-1 and DNA-L4-UTR-3 disrupt stretches of secondary structure while mutations in DNA-L4-UTR-2 and DNA-L4-UTR-4 respectively compensate the previous ones and restore the original predicted secondary structure. The secondary structures of the mutant probes are not shown since they were as theoretically expected except for the DNA-L4-UTR-1 probe, which showed a new alternative secondary structure (shown in item (C) of this figure). **Right**: EMSA performed with each of the left represented labeled probes and increasing amounts of GST-*b*CNBP. **B**: **Left**, DNA-L4-UTR deletion mutant probes. The nucleotides removed in each deletion mutant probe (DNA-L4-UTR-5 to 8) are enclosed. **Right**: EMSA performed with each of the left represented labeled probes and increasing amounts of GST-*b*CNBP. **C**: **Left**, DNA-L4-UTR-1 predicted secondary structure. The nucleotides removed in the deletion mutant probes DNA-L4-UTR-F1 and -F2 are enclosed. **Right**: EMSA performed with the left represented labeled probes and increasing amounts of GST-*b*CNBP. **D**: EMSA performed with the DNA-L4-UTR, DNA-L4-UTR-1, DNA-L4-UTR-1-M1, -M2, and -M3 labeled probes and increasing amounts of GST-*b*CNBP. The predicted secondary structures for DNA-L4-UTR-1-M1, -M2, and -M3 probes are identical to the one predicted for DNA-L4-UTR-1 and are equally stable. EMSAs shown in (A–C) were performed using 0.15, 0.5, and 1.5 μM of GST-*b*CNBP while GST was used at 1.5 μM as control. EMSA shown in (D) was performed using 1.5, 5, and 15 nM of GST-*b*CNBP while GST was used as control at 15 nM.

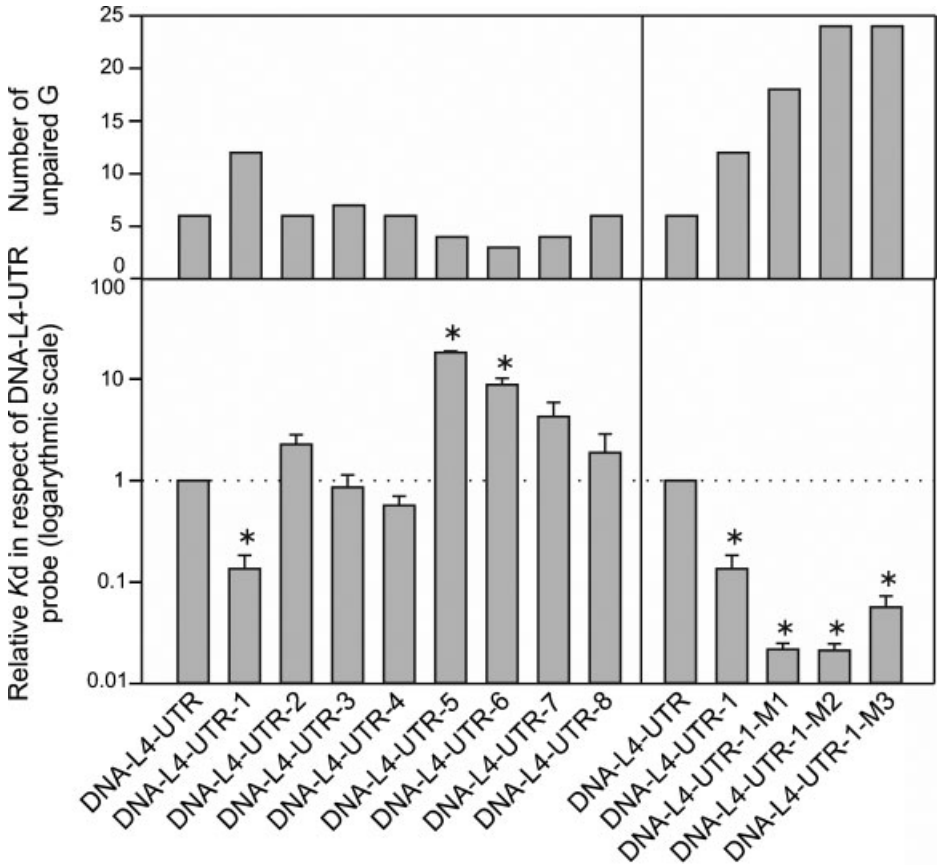

Fig. 5. Correlation of *b*CNBP binding affinities to L4 5′ UTR mutant probes and the number of unpaired guanosines. Bar charts representing the total number of unpaired guanosines (G) in the predicted secondary structure of each analyzed probe (top chart), and the relative GST-*b*CNBP apparent K_d for each of the analyzed probes in respect of the apparent K_d for DNA-L4-UTR probe (bottom chart, means ± SEM, n = 3, logatythmic scale). The left charts show the analysis of the L4 5′ UTR mutant probes that were designed to disrupt secondary structure or to eliminate single-stranded regions (DNA-L4-UTR-1 to -8). The right charts show the analysis of the DNA-L4-UTR-1-M1 to -M3 mutant probes designed to increase the unpaired guanosines number in the DNA-L4-UTR-1 structure. * Indicates $P < 0.05$ in respect of the apparent K_d for DNA-L4-UTR probe (One-way ANOVA, Student–Newman–Keuls Method).

might be favored to interact through a spatial position constrained by the target secondary structure.

CNBP Binds the Purine-Rich Strand of *c-Myc* CT Element Mainly as a Monomer

CNBP had been reported acting as an up-regulator of the human *c-Myc* protooncogene through binding to the CT promoter element [Michelotti et al., 1995]. Thus, the second group of tested probes was conformed by a pair of complementary oligonucleotides that form a perfect duplex representing the CT element (CT and Comp-CT, see Table I and Fig. 6A). These oligonucleotides span the region −135 to −91 upstream the *c-Myc* P1 promoter and include three of the five almost perfect tandem repeats of the CT-element (CCCTCCCC) [Des-Jardins and Hay, 1993]. Both oligonucleotides were labeled and used separately as ssDNA probes or annealed as dsDNA probe for EMSAs. GST-*b*CNBP was able to bind to the Comp-CT strand but not to the CT strand (Fig. 6B) nor to the dsDNA CT probe (Fig. 6C), even when 10 μM GST-*b*CNBP was added (not shown). These data are in agreement with the previously proposed action of CNBP as an ssDNA binding transcription factor that recognizes the purine-rich strand opposite to the pyrimidine-rich strand bound by hnRNP K [Michelotti et al., 1995]. CNBP inability to bind the dsDNA confirms the single-stranded preference previously suggested [Rajavashisth et al., 1989; Flink and Morkin, 1995a]. There was no need to add embryonic extract to achieve GST-*b*CNBP binding to the Comp-CT probe, resembling the DNA-L4-UTR

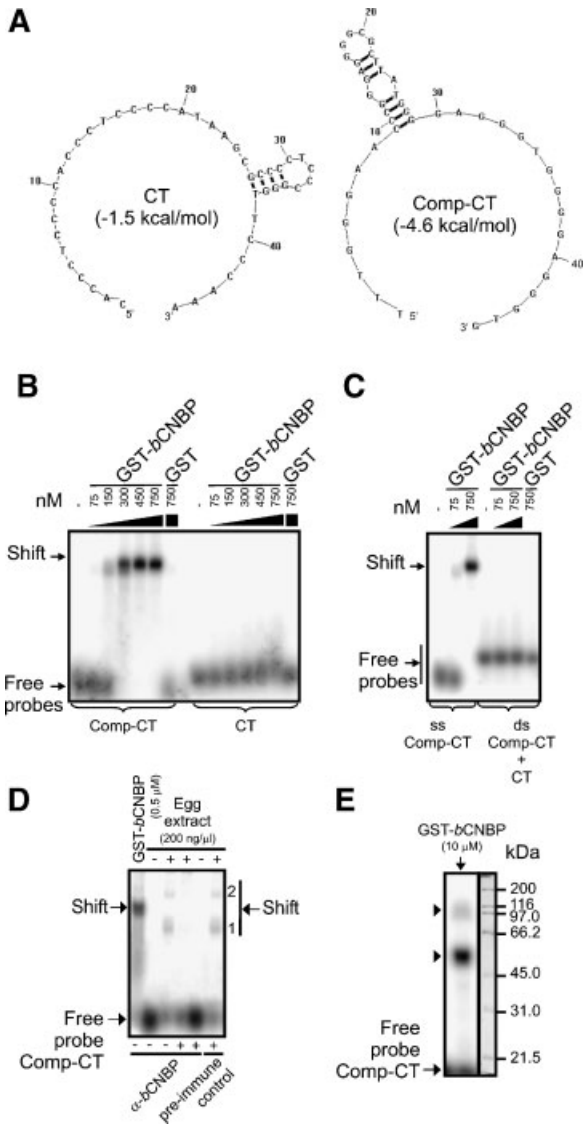

Fig. 6. *b*CNBP binding to *c-Myc* promoter CT element. **A**: Predicted secondary structures and stabilities of ssDNA probes representing sense and antisense strands of the CT promoter element of the human *c-Myc* protooncogene (CT and Comp-CT). **B**: EMSAs using increasing amounts of GST-*b*CNBP and both labeled oligonucleotides separately as ssDNA probes and (**C**) annealed as dsDNA probe. GST was used as control. **D**: EMSA carried out with egg extract (**lane 3**), anti-*b*CNBP antibody (α-*b*CNBP, **lane 5**), and both of them (**lane 4**). α-*b*CNBP was pre-incubated with egg extract before the binding reaction was carried out. GST-*b*CNBP binding reaction was used as a mobility control (**lane 1**). Control reaction was performed using an affinity purified pre-immune serum subjected to the same affinity antibody purification protocol carried out for obtaining anti-CNBP antibody from immune serum. Putative monomeric (1) and dimeric (2) complexes are signaled. **D**: Autoradiography of the SDS–PAGE of the cross-linked binding reaction performed with 10 μM GST-*b*CNBP and labeled Comp-CT probe (**left lane**). In this reaction no embryonic extract was added. Radioactive bands are indicated by arrow tips. Molecular mass markers (**right lane**) were stained with Coomassie Brilliant Blue.

binding requirements. Yet, only one main complex was observed with a calculated apparent $K_d1 = 0.43 \pm 0.18$ μM, tenfold lower than K_ds for DNA-L4-UTR.

Comp-CT probe was bound by proteins from embryonic extracts in a two-complex pattern compatible with the monomer and dimer binding profile described for DNA-L4-UTR (Fig. 6D, third lane). The formation of both complexes was impaired by the addition of affinity-purified anti-*b*CNBP antibody (Fig. 6D, fourth lane), suggesting that their formation depended on embryonic *b*CNBP. The main detectable complex formed by GST-*b*CNBP presented an intermediate electrophoretic mobility between the complexes observed with embryonic extracts. Considering the pattern observed for DNA-L4-UTR binding, this result suggests that the monomer is the major GST-*b*CNBP form that interacts in vitro with the Comp-CT probe. In order to test this, the stoichiometry of the cross-linked complexes was determined by SDS–PAGE (Fig. 6E). Two complexes were observed with apparent M_r of 50,000 and 100,000, which were coincident with the expected M_r for the cross-linked GST-*b*CNBP monomer and dimer, respectively. The main complex was the one corresponding with the molecular mass of the monomer. Therefore, the results reported here not only confirm that CNBP binds to the purine-rich strand of the duplex assayed, that is, Comp-CT probe, but also show that CNBP binds this ssDNA mainly as a monomer. Dimeric complex formation was observed when higher CNBP concentrations were assayed.

CNBP Binding to the Purine-Rich Strand of the *c-Myc* CT Element Depends on the Number of Unpaired Guanosines

The CT element controls up to 75–95% of the *c-Myc* transcription (Liu and Levens, 2006; Yang and Hurley, 2006). Due to the high sensitivity of this DNA segment to DNase I and S1 nuclease it is termed the nuclease-hypersensitive element III_1 (NHE III_1) [Siebenlist et al., 1984; Yang and Hurley, 2006]. The purine-rich strand of NHE III_1 adopts a very stable intrastrand fold-back DNA tetraplex, called G-quadruplex or G-quartet, which requires potassium ions for stability in vitro and was proposed as a critical component of the *c-Myc* transcriptional silencing mechanism [Simonsson et al., 1998; Siddiqui-Jain et al.,

2002; Seenisamy et al., 2004; Yang and Hurley, 2006]. In view of these data, we analyzed the structural features of the CT element taking into account not only the predicted secondary structure, but also the atypical and stable G-quadruplex.

To test the influence of G-quadruplex on CNBP nucleic acid binding behavior, a set of oligonucleotides which had been previously analyzed [Siddiqui-Jain et al., 2002; Seenisamy et al., 2004; Yang and Hurley, 2006] were used as probes in EMSA. Probes used were a 18-nucleotide-long probe corresponding to the minimal sequence involved in the *c-Myc* NHE III_1 G-quadruplex (Comp-CT-18, see Table I) and a mutant probe that has two adenine instead of guanosine residues, thus, impairing the G-quadruplex formation (Comp-CT-18-M, see Table I). Furthermore, the G-quadruplex control probes thrombin-binding aptamer (TBA, see Table I), and HIV-integrase-binding oligonucleotide (T30695, see Table I) were analyzed. None of these four probes showed stable predicted secondary structures. EMSAs were performed in sodium or potassium conditions with identical results (Fig. 7A, only sodium condition EMSA is shown). Compared with the Comp-CT probe, GST-*b*CNBP binding affinities for the G-quadruplex forming probes were significantly lower (Fig. 7A,E). GST-*b*CNBP binding affinities for Comp-CT-18-M and for Comp-CT-18 were not significantly different, suggesting that binding affinities are not directly dependent on G-quadruplex formation. In addition to this, the affinity of *b*CNBP for the DNA-L4-UTR-1-M2 probe, which has the consensus motif for stable G-quadruplex formation $(G_{3\text{-}5}\ N_{1\text{-}7})_4$ [Huppert and Balasubramanian, 2005], and for the DNA-L4-UTR-1-M3 probe, which is theoretically incapable of forming stable G-quadruplex, were not significantly different (see Figs. 4D and 5), even when tested in potassium condition EMSAs (not shown).

From the analysis of the G-quadruplex forming probes it was noticeable that those probes with lower affinities (i.e., statistically significant higher K_ds) than Comp-CT probe, contained lower number of unpaired guanosines (Fig. 7E). This finding was further explored by designing a systematic group of Comp-CT mutant probes that were tested by EMSA. A base-replacement mutant probe disrupted the secondary structure and its compensating mutant probe restored the original secondary structure (Comp-CT-1 and -2, respectively; see Table I and Fig. 7B,C). Two deletion mutant probes were designed in order to eliminate the single-stranded ends present in the Comp-CT predicted secondary structure (Comp-CT-37 and Comp-CT-23; see Table I and Fig. 7B,D). Another tested probe consisted in a single CT element repeat and no stable predicted secondary structure (Comp-CT-12; see Table I and Fig. 7D). Among these probes, only Comp-CT-12 and Comp-CT-23 were bound by GST-*b*CNBP with significant differences in affinities compared to the Comp-CT probe, showing more than 10-fold increases in K_d values (Fig. 7D,E). It is worth noticing that these probes, which have significantly higher K_d values, contain less unpaired guanosines than Comp-CT (Fig. 7E).

Taken together, the data from the simultaneous analysis of CNBP apparent K_ds and the number of unpaired guanosines in the probes support the hypothesis that CNBP would prefer binding to guanosine-rich single-stranded stretches in a spatially constrained environment.

CNBP May Function as a Nucleic Acid Chaperone Protein

The unique structural characteristics of CNBP, which has no apparent functional domains other than the CCHC Zn knuckles and the RGG box, suggest that the mechanism of any kind of action of this protein may depend on the biochemical characteristics of these domains. CCHC zinc knuckles as well as the RGG box are often present in proteins that act as nucleic acid chaperones. Indeed, the zinc knuckles from HIV-1 NC protein are essential for virus cycle and they are important components, together with a basic amino acid background, for the in vitro nucleic acid chaperone activity [Rein et al., 1998; Levin et al., 2005]. The RGG box is ubiquitous in proteins involved in diverse aspects of RNA metabolism such as nucleolin, hnRNP A1, and FMRP, and has been proposed as a participant in nucleic acid chaperone activity [Munroe and Dong, 1992; Hanakahi et al., 2000; Gabus et al., 2004]. On the other hand, CNBP was proposed to participate together with La and Ro60 autoantigens in the regulation of the translational fate of the L4 rp-mRNA [Pellizzoni et al., 1998], and with hnRNP K in the binding of the CT sequence element participating in *c-Myc* transcriptional

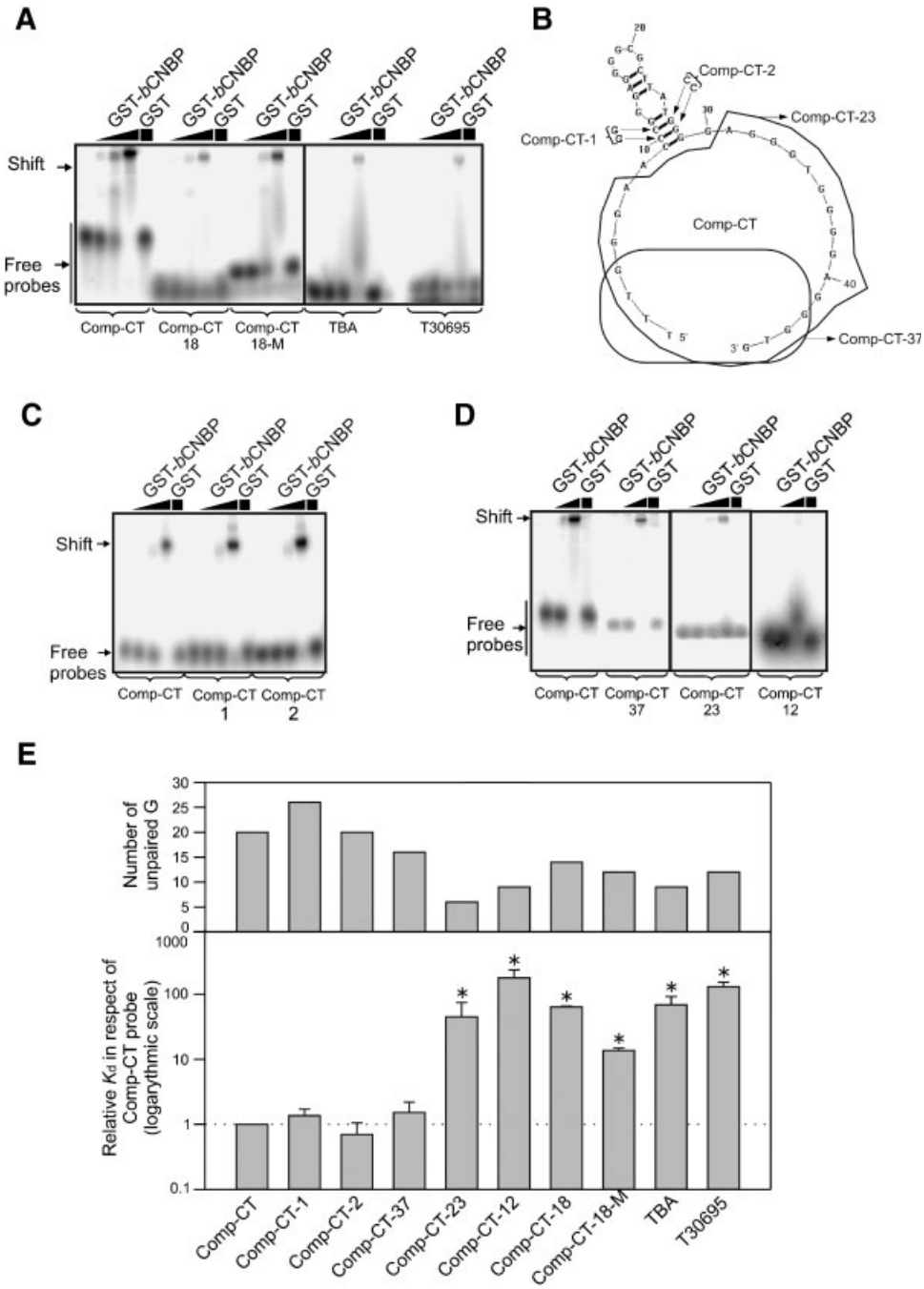

Fig. 7. *b*CNBP binding to Comp-CT mutant probes and G-quadruplex forming probes. **A**: EMSA performed with Comp-CT, Comp-CT-18, Comp-CT-18-M, TBA and T30695 labeled probes and increasing amounts of GST-*b*CNBP. **B**: Comp-CT mutant probes. Arrows indicate the replaced nucleotides in each mutant probe. Base-replacements in Comp-CT-1 disrupt a stretch of secondary structure while mutations in Comp-CT-2 compensate the previous ones and restore the original predicted secondary structure. The nucleotides removed in each deletion mutant probe (Comp-CT-37 and Comp-CT-23) are enclosed. **C**: EMSA performed with Comp-CT, Comp-CT-1 and Comp-CT-2 labeled probes and increasing amounts of GST-*b*CNBP. **D**: EMSA performed with Comp-CT, Comp-CT-37, Comp-CT-23, and Comp-CT-12 labeled probes and increasing amounts of GST-*b*CNBP. EMSAs shown in (A,C,D) were performed using 50, 150, and 500 nM of GST-*b*CNBP while GST was used at 500 nM as control. **E**: Bar charts representing the total number of unpaired guanosines (G) in the predicted secondary structure of each of the analyzed probes (top chart), and the relative GST-*b*CNBP apparent K_d for each analyzed probe in respect of the apparent K_d for Comp-CT probe (bottom chart, means ± SEM, n = 3, logatythmic scale was used). * Indicates $P < 0.05$ in respect of the apparent K_d for Comp-CT probe (One-way ANOVA, Student–Newman–Keuls Method).

activation [Michelotti et al., 1995]. The three proposed CNBP partners are nucleocytoplasmic shuttling RNA-binding proteins that display nucleic acid structure remodeling activities [Tomonaga et al., 1998; Krecic and Swanson, 1999; Fabini et al., 2000; Maraia and Intine, 2002; Ivanyi-Nagy et al., 2005]. To test whether CNBP may act as a nucleic acid chaperone, we analyzed several CNBP structural and biochemical properties. First, we analyzed CNBP primary structure using the DisProt VL3-E online predictor [Peng et al., 2005]. It was proposed that proteins with nucleic acid chaperone functions have the highest frequency of disorganized structure among different protein classes, with more than half of all amino acids residues situated in putative unstructured segments [Tompa and Csermely, 2004]. These intrinsically unstructured regions provide a platform for interaction with a variety of nucleic acid targets, thus making these proteins suitable for integrating gene expression control signals in the cell through chaperone activity [Ivanyi-Nagy et al., 2005]. Figure 8 shows large amino acid segments predicted as highly intrinsically unstructured in the CNBP structure.

Second, as nucleic acid chaperones promote the annealing of complementary strands [Munroe and Dong, 1992; Rein et al., 1998; Hanakahi et al., 2000; Martin and Bushman, 2001; Levin et al., 2005], we assayed CNBP annealing promotion using CT and Comp-CT oligonucleotides (Fig. 9A) and increasing amounts of GST-*b*CNBP. The assayed molar protein/probe ratios were from 1:10 to 30:1, observing faster annealing rates than controls above 3:1 protein/probe molar ratios. The highest annealing rate was observed using molar ratios of 10:1 (tenfold excess of fusion protein over probes), reaching approximately 55% annealing at 30 min (Fig. 9B). This value was statistically significant and was near the maximum obtained in the assay, as shown in the completely annealed control (Fig. 9C). Molar ratios higher than 10:1 caused a decrease in annealing promotion

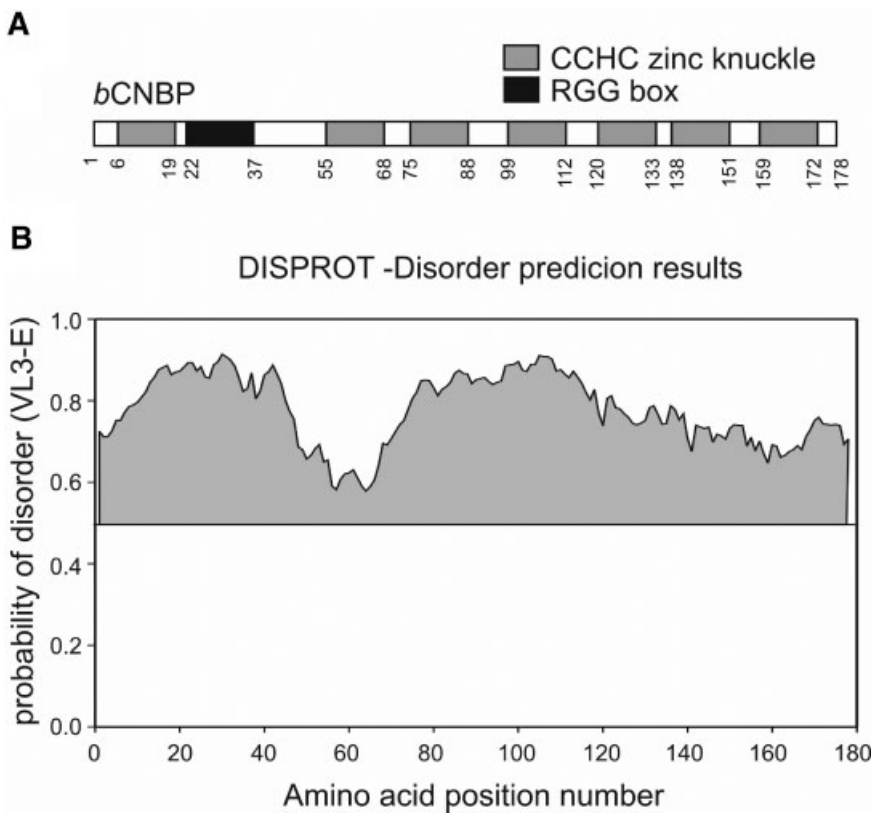

Fig. 8. Prediction of *b*CNBP intrinsically unstructured regions. **A**: *b*CNBP domain structure scheme. Numbers below the scheme indicate the amino acid position in the protein sequence. **B**: Prediction of *b*CNBP intrinsically unstructured regions using the DisProt VL3-E predictor. A score above or equal to 0.5 predicts an amino acid in a disordered environment.

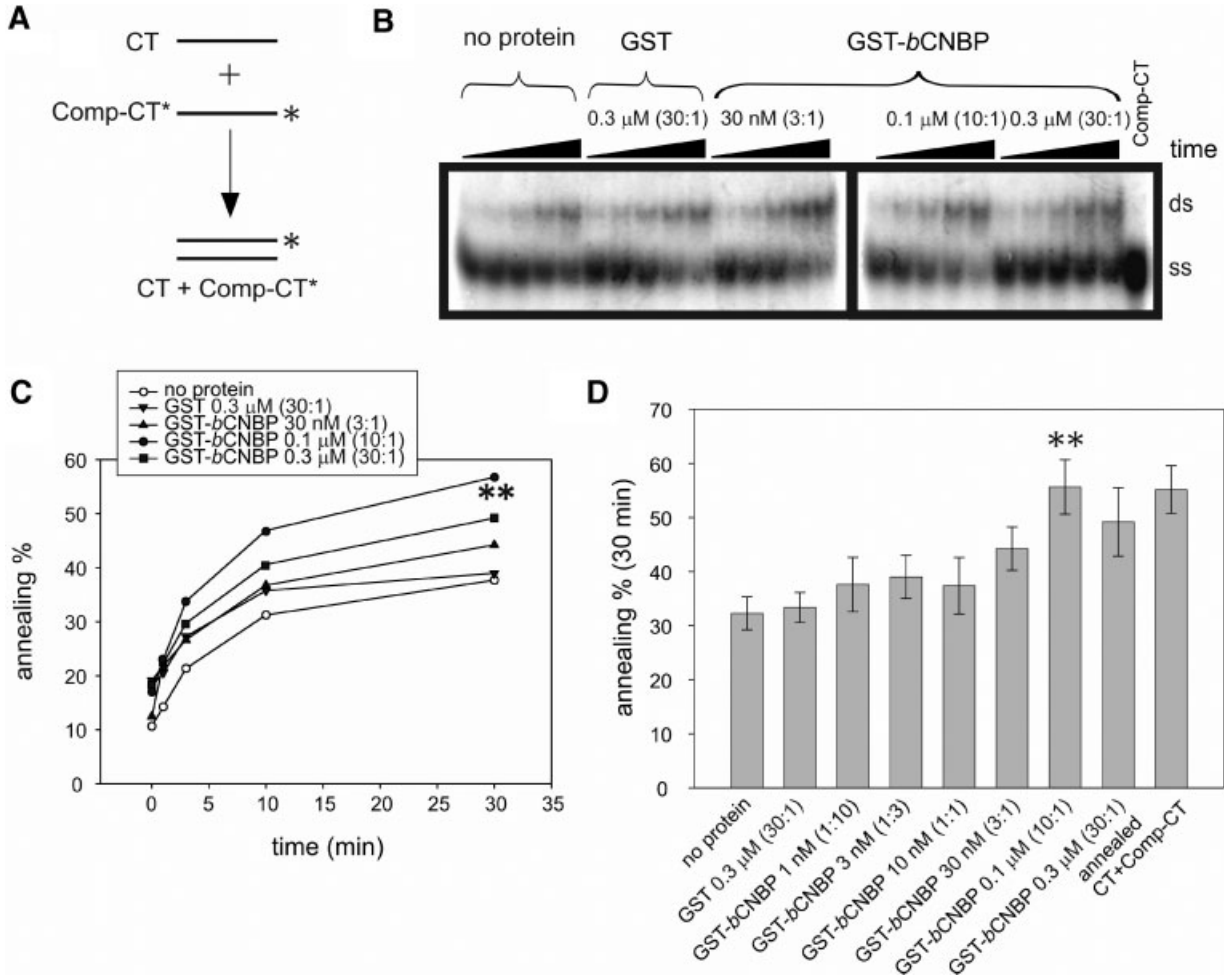

Fig. 9. *b*CNBP nucleic acid annealing promotion activity. **A**: Scheme of the annealing assay using labeled Comp-CT and unlabeled CT oligonucleotides. * Indicates the ^{32}P-labeled strand. **B**: Annealing assay performed without added protein or with increasing amounts (threefold serial increments from 0.03 to 0.3 μM) of GST-*b*CNBP. A reaction using 0.3 μM GST was carried out as a control. The last lane at the right corresponds to the labeled Comp-CT probe alone. For each protein concentration, reactions were sampled at 0, 1, 3, 10, and 30 min. Single-stranded (ss) and double-stranded (ds) probes are indicated at the right of each figure. The molar protein/probe ratio is indicated for each protein concentration. **C**: Plots of the annealing percentages (%) obtained in (B) versus time, represented as the average of three independent experiments. Annealing was analyzed for no protein (○), 0.3 μM GST (▼), and 30 nM (▲), 0.1 μM (●), and 0.3 μM (■) GST-*b*CNBP. **D**: Bar chart representing annealing percentage (%) obtained at 30 min (means ± SEM, n = 3). GST-*b*CNBP concentrations ranging from 1 nm to 0.3 μM were analyzed. Completely annealed reactions (annealed CT + Comp-CT) were used as control of the maximum expected annealing percentage. ** Indicates $P < 0.01$ in respect of the controls (One-way ANOVA, Tuckey test).

activities (Fig. 9B,C) and no differences were observed with incubations longer than 30 min. GST did not cause changes in the annealing kinetics compared with controls, indicating that *b*CNBP itself was responsible for the observed effect. It is worth mentioning that the protein/probe ratio 10:1, which represents a molar nucleotide/protein ratio of ≅8:1, is similar to the optimal ratio obtained for the HIV-1 NC protein [Rein et al., 1998; Levin et al., 2005], and is in the same range as the determined for other nucleic acid chaperones like nucleolin [Hanakahi et al., 2000], FMRP [Gabus et al., 2004], and hnRNP A1 [Munroe and Dong, 1992].

Finally, we analyzed CNBP melting promotion activity taking into account that nucleic acid chaperones cause nucleic acid duplex denaturation [Martin and Bushman, 2001; Levin et al., 2005]. Melting assays were performed using CT-12 and Comp-CT-12 oligonucleotides (see Table I and Fig. 10A), and CT and either Comp-CT or Comp-CT-M2 oligonucleotides that form a perfect duplex or a mismatched duplex, respectively (see Table I and Fig. 10D$_1$,E$_1$).

Melting curves were performed with increasing amounts of GST-*b*CNBP at molar protein/probe ratios ranging from 1:1 to 1,000:1. The assayed temperatures were not higher than 60°C, since this temperature was experimentally determined as the GST-*b*CNBP thermal denaturation value (Fig. S7, Supplementary Material). For the CT-12/Comp-CT-12 duplex the effects

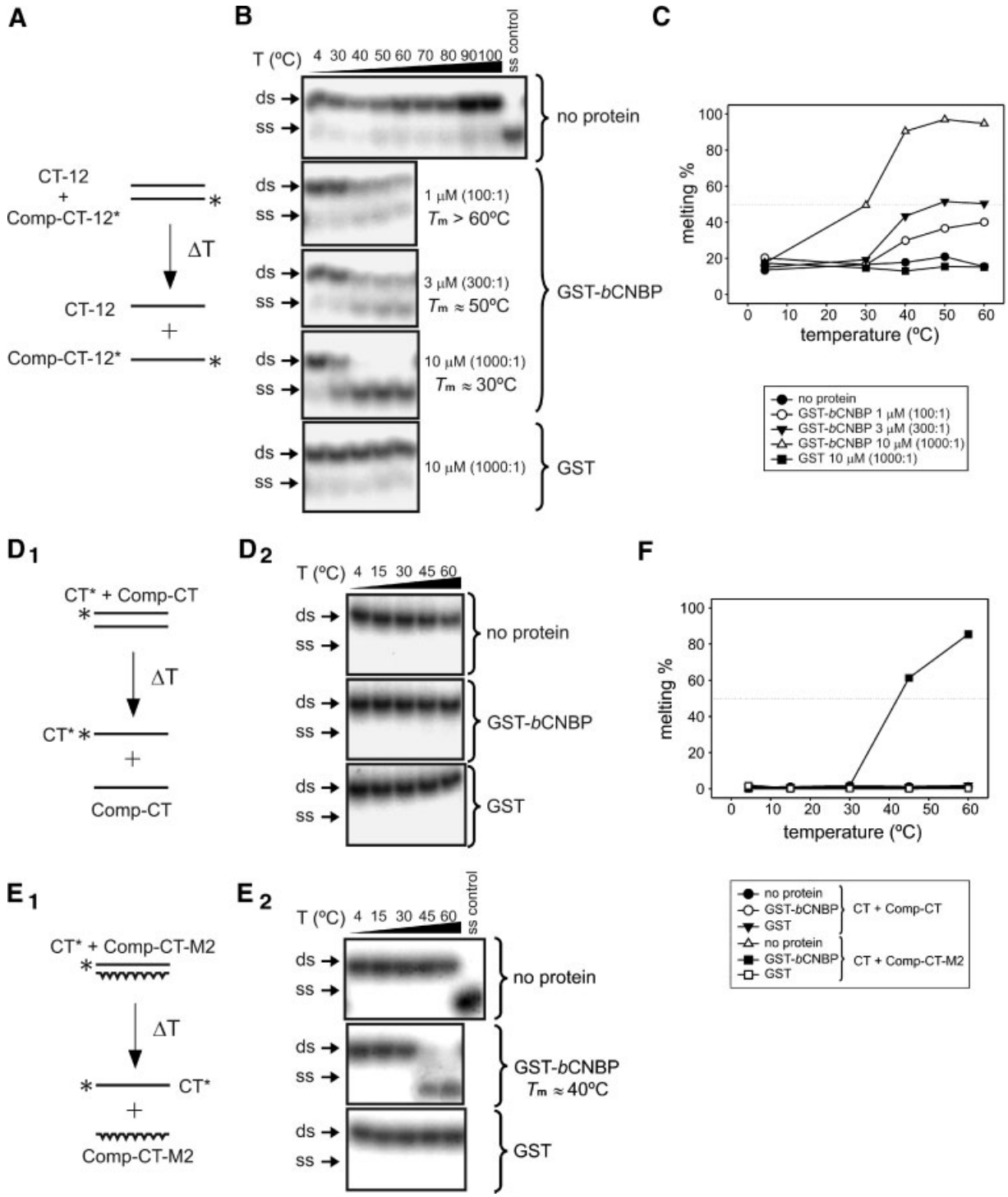

Fig. 10.

on melting temperature became evident at protein/probe ratios over 100:1. The maximal effect ($T_m \approx 30^\circ C$) was observed at a 1,000:1 protein/probe ratio (Fig. 10B,C). GST did not cause changes in the melting compared with controls, indicating that *b*CNBP itself was responsible for the melting effect. Protein/probe ratios of 100:1 and 1,000:1, which represent molar nucleotide/protein ratios between 1:5 and 1:50, are in the same range as the ones determined for other nucleic acid chaperones [Tsuchihashi and Brown, 1994]. To analyze the effect of GST-*b*CNBP on the melting of longer DNA duplexes, the assay was performed using CT/Comp-CT perfect duplex (Fig. $10D_1,D_2,F$). No evident effect of GST-*b*CNBP on the melting temperature was observed. However, when melting assay was performed using CT/Comp-CT-M2 mismatched duplex (Fig. $10E_1,E_2,F$) a noticeable effect of GST-*b*CNBP on the melting temperature was evident ($T_m \approx 40^\circ C$). GST did not cause changes in the melting profile compared with controls, indicating that the observed effect is caused by *b*CNBP. *b*CNBP promoted the melting of a mismatched duplex (CT/Comp-CT-M2) but failed to promote the melting of an stable 45-bp-long perfect duplex (CT/Comp-CT), suggesting that CNBP targets for melting activity are partial unpaired duplexes. In the case of CT-12/Comp-CT-12 duplex, as this is a shorter duplex than CT/Comp-CT and, consequently, less stable, it may experience transient partial denaturations that could be recognized by *b*CNBP allowing it to promote the duplex melting. On the other hand, EMSAs performed using partially mismatched duplexes (CT/Comp-CT-M2 and CT/Comp-CT-37) showed that GST-*b*CNBP shifted the labeled probes (Fig. S8, Supplementary Material). Thus, it seems that *b*CNBP is unable to bind to CT/Comp-CT perfect duplex (see Fig. 6C) while it actually binds partially mismatched duplexes, probably through recognition of unpaired nucleotides. Based on these data, it is tempting to speculate that CNBP binding to partially mismatched probes might be a consequence of the *b*CNBP-favored melting and subsequent binding to the labeled single-stranded molecule.

CNBP annealing and melting promotion activities may be the biochemical basis to explain the broad-spectrum functions and the diversity of processes wherein CNBP has been implicated, and may help to merge all the proposed actions of this protein over nucleic acids in a unique biochemical way of action: the nucleic acid chaperone activity.

DISCUSSION

The main goal of this work was to gain further insights into the CNBP single-stranded nucleic acid binding features and the identification and characterization of CNBP main biochemical activities. We aimed our work to find CNBP features that would explain how this protein may be involved in several gene expression regulation pathways.

Main Features of CNBP Binding to Single-Stranded Nucleic Acid Targets

The experiments on CNBP nucleic acid binding were mainly aimed to establish CNBP binding pattern behavior and biochemical requirements. Our data contribute to the knowledge of the binding stoichiometry and structural characterization of the bound single-stranded nucleic acids and, thus, begin to dissect the main features of putative CNBP cellular nucleic acid targets.

Fig. 10. *b*CNBP melting promotion activity. **A**: Scheme of the melting assays using a duplex formed by pre-annealed labeled Comp-CT-12 and unlabeled CT-12 oligonucleotides. * Indicates the ^{32}P-labeled strand. **B**: The melting assay represented in (A) was performed with no protein or increasing amounts (threefold serial increments from 1 to 10 μM) of GST-*b*CNBP. A control reaction was carried out adding 10 μM GST. The ''ss control'' corresponds to the single-stranded labeled Comp-CT-12 probe. The calculated melting temperatures (T_m) and the molar protein/probe ratio are indicated for each condition. Single-stranded (ss) and double-stranded (ds) probes are indicated at the left of each gel. **C**: Graph representing the melting curves (plot of the melting percentages vs. temperature) from (B) used for melting temperatures (T_m) calculations. Three independent experiments were performed for each condition. **D_1**: Scheme of the melting assay using a duplex formed by pre-annealed labeled Comp-CT and unlabeled CT oligonucleotides. **D_2**: The melting assay represented in (D_1) was performed with no protein or 10 μM (1,000:1 protein/probe ratio) GST-*b*CNBP or GST. **E_1**: Scheme of the melting assay using a duplex formed by pre-annealed labeled Comp-CT-M2 and unlabeled CT oligonucleotides. **E_2**: The melting assay represented in (E_1) was performed with no protein or 10 μM (1,000:1 protein/probe ratio) GST-*b*CNBP or GST. The ''ss control'' corresponds to the single-stranded labeled CT probe. **F**: Graph representing the melting curves (plot of the melting percentages vs. temperature) from (D_1) and (E_1) used for melting temperatures (T_m) calculations. Three independent experiments were performed for each condition.

*b*CNBP from embryonic developmental stages previous to MBT was able to bind the single-stranded nucleic acid probes, RNA or DNA, containing the sequence from the 5′ UTR from *X. laevis* L4 rp-mRNA. On the contrary, *b*CNBP from developmental stages posterior to MBT, did not bind the RNA probe (Fig. 1A) and bound the ssDNA probe with lower affinity (Fig. S4). These results suggest that beyond MBT embryos contain either lower amounts or less available *b*CNBP. CNBP changes its sub-cellular localization from cytoplasm to nuclei beyond MBT [Armas et al., 2001; Armas et al., 2004]. It would be possible that the signal causing the CNBP sub-cellular localization change may also be affecting its nucleic acid binding behavior. The MBT determines the beginning of zygotic transcription [Newport and Kirschner, 1982] and, consequently, embryonic CNBP may be engaged in binding to newly synthesized RNAs and/or to transcriptional control elements of developmentally regulated genes. Therefore, beyond MBT CNBP may interact with nuclear targets that block its binding site, making it inaccessible to the external probes used in EMSAs.

Recombinant CNBP bound RNA as well as ssDNA probes, but with some differences according to the nature of the nucleic acid molecule. Previous reports proposed the existence of monomer-dimer CNBP forms binding to nucleic acid targets [Pellizzoni et al., 1998; Armas et al., 2004]. Our results from UV cross-linking, EMSA and competition EMSA showed that dimer binds to nucleic acids target as a pre-formed species, suggesting the existence of non-interconvertible monomeric and dimeric CNBP forms co-existing in solution. Competition EMSA also revealed differences in the stoichiometry of complexes formed between GST-*b*CNBP and ssDNA or RNA probes. GST-*b*CNBP bound ssDNA as monomer and dimer while RNA was only bound by GST-*b*CNBP dimer. It was proposed that the L4 rp-mRNA 5′ UTR contains two independent CNBP binding sites that rely on its secondary structure [Pellizzoni et al., 1998]. Since the predicted secondary structures of ssDNA and RNA probes are identical, the difference in their stabilities may cause the exposure or hiding of alternative binding sites determining the difference in complexes stoichiometries. Indeed, the RNA probe displays a more stable secondary structure and, thus, may expose CNBP dimer binding sites. The lower stability of ssDNA probe secondary structure may enable it to explore alternative conformations allowing it to adopt another conformation able to be bound by the CNBP monomer. In a similar fashion, the Comp-CT probe secondary structure stability may allow it to be bound by the monomeric GST-*b*CNBP while little interaction may occur with the dimeric CNBP form.

CNBP Preferred Targets Contain G-Rich Single-Stranded Stretches

The main structural features of putative CNBP nucleic acid targets were analyzed by means of sets of artificial mutant probes systematically designed from the L4-5′ UTR and Comp-CT CNBP targets. The analysis was performed taking into account the predicted secondary structures and other atypical structures possibly displayed by each of the assayed probes.

From the analysis of DNA-L4-UTR and Comp-CT probes and their derivatives it was evident a direct correlation between the CNBP binding affinity and the amount of unpaired guanosines present in the probes. On the other hand, the assays performed with homopolymeric probes suggest that the G-rich single-stranded stretches need a sequence and/or structural constrain to be bound by CNBP. These findings agree with data from the literature which show that other CNBP single-stranded nucleic acid target sequences, for example, the sterol regulatory element (SRE, GTG(G/T)GGTG) [Rajavashisth et al., 1989], the distal suppressor sequence from the human β-myosin heavy chain (MHC, GTGGTCGTG) [Flink and Morkin, 1995a], the JC virus early promoter-enhancer (JCV(E)) tandem pentanucleotide repeat element (AGGGAAGGGA) [Liu et al., 1998], the macrophage colony-stimulating factor (M-CSF or CSF-1) promoter region upstream the phorbol ester-responsive element (CTCAGTGGGCCTCTGGGGTGTA) [Konicek et al., 1998], and the 5′ region of rat *c-Myc* pre-mRNA intron 1 (AGUUGGGGUGGA) [Yasuda et al., 1995] are mainly composed by guanosine-rich stretches. The CNBP preference for guanosine nucleotides is coherent when CNBP structural motifs are considered. Structural studies of the HIV-1 NC protein revealed that each of the CCHC zinc knuckles bind to one exposed guanosine base located in a single-stranded G-rich loop. HIV-1 NC protein binds its target in a mutually induced fashion, in

which NC protein pulls unstacked guanosines into hydrophobic clefts on the surface of the zinc knuckles [De Guzman et al., 1998; Amarasinghe et al., 2000]. On the other hand, the RGG box motif interacts with nucleic acids in a sequence-unspecific manner by electrostatic interactions between the positively charged Arg residues and the negatively charged nucleic acids phosphate backbone, disrupting the nucleic acids secondary structure by unstacking adjacent bases through intercalation of the aromatic residues present in the RGG motif [Raman et al., 2001]. However, the concept of the RGG box as a non-specific binding motif has changed since the RGG box was demonstrated to act as a sequence-specific binding domain for the fragile X mental retardation protein (FMRP), that recognizes guanosine-rich sequences that fold into G-quadruplex [Darnell et al., 2004].

We considered the possibility that the preferred CNBP targets were nucleic acid folded into G-quadruplex bearing in mind that CNBP preferred targets should contain unpaired G-stretches interspersed with other nucleotides. In our assays, no significant differences were detected when comparing *b*CNBP binding behavior to probes potentially able to form the G-quadruplex and probes impaired to form them. Instead, the analysis of these probes allowed us to reinforce the hypothesis that the main feature of CNBP putative targets are the number of unpaired guanosines constrained in their spatial position by the secondary structure.

Taken together, here we propose that CNBP preferred targets are single-stranded nucleic acids, either RNA or DNA, containing unpaired guanosine-rich stretches. In a similar fashion than retroviral NC protein, CNBP binding to its targets may involve a co-folding of the protein and the single-stranded nucleic acid during the interaction, which may cause the entrance of guanosine nucleobases to the pockets formed by the CCHC zinc knuckles and the RGG box. CNBP binding sequence specificity may be favored by constrained spatial positions of guanosine nucleobases governed by the target secondary structure, which may not necessarily imply the G-quadruplex folding.

CNBP May Function as Nucleic Acid Chaperone

Nucleic acid chaperones are nucleic acid binding proteins that catalyze the rearrangement and folding of nucleic acids into conformations that have the maximal number of base pairs and form the thermodynamically most stable three-dimensional structure needed for biological function [Rein et al., 1998; Tompa and Csermely, 2004; Ivanyi-Nagy et al., 2005; Levin et al., 2005]. These chaperones assist folding by randomly disrupting the misformed bonds via repeated cycles of binding and release, allowing the substrate to resume search in the conformational space towards the global energy minimum. As locally unpaired strands of the substrates are kept at a close range by the bound chaperone, this proximity limits the subsequent conformational search and speeds up the folding process. All these distinct elements of rapid and versatile binding, local unfolding, and proximal positioning of segments of the misfolded substrate by the flexible/disordered region(s) of chaperones were assembled into an ATP-independent mechanistic model of chaperone action called the "entropy transfer model" [Tompa and Csermely, 2004]. The disordered structure predicted from CNBP amino acid sequence, in addition to the annealing and melting promotion activities, strongly suggests that CNBP functions as a nucleic acid chaperone. Internal mismatches of preformed mismatched duplex substrates may be recognized and bound by CNBP as local single-stranded regions, thereby exerting the helix destabilization (melting promotion) activity and favoring the disassembly of incorrect structures. On the other hand, CNBP annealing promotion activity may prevent nucleic acid complete denaturation by neutralizing the negative charge of the phosphodiester backbone, reducing electrostatic repulsion between the strands and allowing them to perform the conformational search to reach a more favorable structure.

The nucleic acid chaperone activity has been thoroughly analyzed for retroviral NC proteins. The HIV-1 NC chaperone activity is based on two related activities, hybridization and helix-destabilization [Levin et al., 2005]. The first one is sequence-unspecific and a consequence of the basic amino acids in the backbone of HIV-1 NC protein, which provide positive charges that act as counter ions of nucleic acids, binding to the phosphate backbone and allowing their attraction and annealing. The second activity, that is, helix-destabilization, is a relatively weak activity that partially destabilizes the secondary

structure of nucleic acids in order to provide single-stranded regions that can nucleate new duplex formation. Helix-destabilization activity depends on zinc knuckles due to their preference for binding single-stranded (unpaired) nucleotides. These activities are experimentally separable and the relative importance of each one depends strongly on the particular assayed targets. Besides, the RGG box motifs from other nucleic acid chaperones have been directly related with their annealing promotion and duplex destabilization activities [Munroe and Dong, 1992; Hanakahi et al., 2000; Gabus et al., 2004]. Therefore, the CNBP RGG box may act as a basic backbone functioning in a sequence-unspecific way by remodeling the nucleic acids secondary structure in order to yield a better binding target. On the other hand, the CNBP zinc knuckles may contribute in a sequence-specific fashion to the helix-destabilization activity by binding to single-stranded unpaired nucleobases. Results reported here suggest that the CNBP biological function may depend on its nucleic acid chaperone activity, which may consist in the binding and refolding of nucleic acids secondary structures through melting and annealing promotion activities, to render the most stable conformation.

It is noticeable that the CNBP nucleic acid chaperone activity is compatible with its possible cellular functions. The model wherein CNBP participates in rp-mRNA translational control involves 5′ UTR equilibrium between open and closed conformations. A closed CNBP-bound form would result in translation inhibition while an open La-bound form would, instead, allow translation [Pellizzoni et al., 1998]. Cell signals might influence the affinity of CNBP for the 5′ UTR, and its differential binding may lead, either alone or together with additional proteins, to diverse effects on the translation of rp-mRNAs. CNBP may control the translational process affecting diverse cellular processes dependent on RNA structure through the nucleic acid chaperone activity, as it was well documented for other RNA chaperones [Weeks, 1997; Rein et al., 1998; Tompa and Csermely, 2004; Ivanyi-Nagy et al., 2005; Levin et al., 2005].

Single-stranded nucleic acid binding proteins, together with double-stranded transcriptional factors, have been associated with the transcriptional modulation of a variety of genes. This is the case for CNBP in regards of the *c-Myc* promoter model. Certainly, it was reported that CNBP binds to the purine-rich strand of the CT element together with the hnRNP K, and possibly additional factors, determining the formation of an open protein–DNA complex [Tomonaga et al., 1998]. The CT promoter element is also bound by dsDNA binding factors, such as Sp1 [DesJardins and Hay, 1993]. Consequently, the CT element conformation modulation might either allow or block the action of transcriptional factors, allowing it to confer different properties upon nearby promoters [Tomonaga et al., 1998]. Here we have shown that CNBP binds to Comp-CT single-stranded sequence and promotes the annealing of the CT and Comp-CT strands as well as the melting of partially unpaired duplexes. The CNBP nucleic acid chaperone activity may promote or restrict the interaction of some other conventional transcription factors with the *c-Myc* promoter by controlling the formation of single-stranded loops. Based on these findings, we hypothesize that CNBP controls *c-Myc* transcriptional rate trough its annealing and melting promotion activities, which would remodel chromatin structure and influence *trans* factors binding to the CT element. This hypothesis directly links alterations produced by CNBP in DNA conformation and topology with potential changes in *c-Myc* protooncogene transcription efficiency. It is worth mentioning that some other multifunctional proteins have been reported to function through a similar biochemical mechanism. For example, Purα, a single-stranded DNA and RNA binding protein, has helix-destabilizing properties and activates the platelet-derived growth factor A (PDGF-A) gene transcription via interactions with a promoter G-rich single-stranded region [Zhang et al., 2005].

The purine-rich strand of the CT element has also been related to the formation of alternative structures such as G-quadruplex critical for the *c-Myc* transcriptional silencing mechanism [Simonsson et al., 1998; Siddiqui-Jain et al., 2002; Seenisamy et al., 2004; Yang and Hurley, 2006]. In this context, it would be tempting to speculate that CNBP binding to the purine-rich strand of the *c-Myc* promoter CT-element may affect the stability not only of duplex but also of G-quadruplex formation. Indeed, FMRP [Zanotti et al., 2006] and HIV-1 NC protein [Lyonnais et al., 2003; Kankia et al., 2005] are both nucleic acid chaperones that share

structural motifs with CNBP and interact with G-quadruplex changing their folding stabilities. However, this hypothesis needs to be experimentally proved.

In light of the results presented here, the CNBP nucleic acid chaperone activity is a novel biochemical function that broadens the field of study about CNBP biological function. This finding is a significant contribution for settling the molecular basis of the diverse ways in which CNBP may control gene expression. Further research is being performed to completely elucidate the relationship between CNBP nucleic acid chaperone activity and its structural and/or regulative features. In addition, in vivo experiments are being performed to more precisely correlate the CNBP biochemical and cellular activities with its biological functions.

ACKNOWLEDGMENTS

N. B. Calcaterra is a Staff Member and P. Armas is a Fellow of the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina. We acknowledge Dr. P. Pierandrei-Amaldi for the *X. laevis* L4 ribosomal protein cDNA clone (GenBank Accession Number X05216).

REFERENCES

Abe Y, Chen W, Huang W, Nishino M, Li YP. 2006. CNBP regulates forebrain formation at organogenesis stage in chick embryos. Dev Biol 259:116–127.

Amarasinghe GK, De Guzman RN, Turner RB, Chancellor KJ, Wu ZR, Summers MF. 2000. NMR structure of the HIV-1 nucleocapsid protein bound to stem-loop SL2 of the psi-RNA packaging signal. Implications for genome recognition. J Mol Biol 301:491–511.

Armas P, Cabada MO, Calcaterra NB. 2001. Primary structure and developmental expression of Bufo arenarum cellular nucleic acid-binding protein: Changes in subcellular localization during early embryogenesis. Dev Growth Differ 43:13–23.

Armas P, Cachero S, Lombardo VA, Weiner A, Allende ML, Calcaterra NB. 2004. Zebrafish cellular nucleic acid-binding protein: Gene structure and developmental behaviour. Gene 337:151–161.

Burd CG, Dreyfuss G. 1994. Conserved structures and diversity of functions of RNA-binding proteins. Science 265:615–621.

Calcaterra NB, Palatnik JF, Bustos DM, Arranz SE, Cabada MO. 1999. Identification of mRNA-binding proteins during development: Characterization of Bufo arenarum cellular nucleic acid binding protein. Dev Growth Differ 41:183–191.

Carey J. 1991. Gel retardation. Methods Enzymol 208:103–117.

Chen W, Liang Y, Deng W, Shimizu K, Ashique AM, Li E, Li YP. 2003. The zinc-finger protein CNBP is required for forebrain formation in the mouse. Development 130: 1367–1379.

Darnell JC, Warren ST, Darnell RB. 2004. The fragile X mental retardation protein, FMRP, recognizes G-quartets. Ment. Retard Dev Disabil Res Rev 10:49–52.

De Guzman RN, Wu ZR, Stalling CC, Pappalardo L, Borer PN, Summers MF. 1998. Structure of the HIV-1 nucleocapsid protein bound to the SL3 psi-RNA recognition element. Science 279:384–388.

DesJardins E, Hay N. 1993. Repeated CT elements bound by zinc finger proteins control the absolute and relative activities of the two principal human c-myc promoters. Mol Cell Biol 13:5710–5724.

Fabini G, Rutjes SA, Zimmermann C, Pruijn GJ, Steiner G. 2000. Analysis of the molecular composition of Ro ribonucleoprotein complexes. Identification of novel Y RNA-binding proteins. Eur J Biochem 267:2778–2789.

Flink IL, Morkin E. 1995a. Alternatively processed isoforms of cellular nucleic acid-binding protein interact with a suppressor region of the human beta-myosin heavy chain gene. J Biol Chem 270:6959–6965.

Flink IL, Morkin E. 1995b. Organization of the gene encoding cellular nucleic acid-binding protein. Gene 163:279–282.

Gabus C, Mazroui R, Tremblay S, Khandjian EW, Darlix JL. 2004. The fragile X mental retardation protein has nucleic acid chaperone properties. Nucleic Acids Res 32:2129–2137.

Hanakahi LA, Bu Z, Maizels N. 2000. The C-terminal domain of nucleolin accelerates nucleic acid annealing. Biochemistry 39:15493–15499.

Huppert JL, Balasubramanian S. 2005. Prevalence of quadruplexes in the human genome. Nucleic Acids Res 33:2908–2916.

Ivanyi-Nagy R, Davidovic L, Khandjian EW, Darlix JL. 2005. Disordered RNA chaperone proteins: From functions to disease. Cell Mol Life Sci 62:1409–1417.

Kankia BI, Barany G, Musier-Forsyth K. 2005. Unfolding of DNA quadruplexes induced by HIV-1 nucleocapsid protein. Nucleic Acids Res 33:4395–4403.

Konicek BW, Xia X, Rajavashisth T, Harrington MA. 1998. Regulation of mouse colony-stimulating factor-1 gene promoter activity by AP1 and cellular nucleic acid-binding protein. DNA Cell Biol 17:799–809.

Krecic AM, Swanson MS. 1999. hnRNP complexes: Composition, structure, and function. Curr Opin Cell Biol 11: 363–371.

Levin JG, Guo J, Rouzina I, Musier-Forsyth K. 2005. Nucleic acid chaperone activity of HIV-1 nucleocapsid protein: Critical role in reverse transcription and molecular mechanism. Prog Nucleic Acid Res Mol Biol 80:217–286.

Liu JX, Gui JF. 2005. Expression pattern and developmental behaviour of cellular nucleic acid-binding protein (CNBP) during folliculogenesis and oogenesis in fish. Gene 356:181–192.

Liu J, Levens D. 2006. Making myc. Curr Top Microbiol Immunol 302:1–32.

Liu M, Kumar KU, Pater MM, Pater A. 1998. Identification and characterization of a JC virus pentanucleotide repeat element binding protein: Cellular nucleic acid binding protein. Virus Res 58:73–82.

Lombardo VA, Armas P, Weiner AM, Calcaterra NB. 2007. In vitro embryonic developmental phosphorylation of the cellular nucleic acid binding protein by cAMP-dependent protein kinase, and its relevance for biochemical activities. FEBS J 274:485–497.

Lyonnais S, Gorelick RJ, Mergny JL, Le Cam E, Mirambeau G. 2003. G-quartets direct assembly of HIV-1 nucleocapsid protein along single-stranded DNA. Nucleic Acids Res 31:5754–5763.

Maraia RJ, Intine RV. 2002. La protein and its associated small nuclear and nucleolar precursor RNAs. Gene Expr 10:41–57.

Martin SL, Bushman FD. 2001. Nucleic acid chaperone activity of the ORF1 protein from the mouse LINE-1 retrotransposon. Mol Cell Biol 21:467–475.

Mathews DH, Sabina J, Zuker M, Turner DH. 1999. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. J Mol Biol 288:911–940.

McGrath CF, Buckman JS, Gagliardi TD, Bosche WJ, Coren LV, Gorelick RJ. 2003. Human cellular nucleic acid-binding protein Zn2+ fingers support replication of human immunodeficiency virus type 1 when they are substituted in the nucleocapsid protein. J Virol 77:8524–8531.

Michelotti EF, Tomonaga T, Krutzsch H, Levens D. 1995. Cellular nucleic acid binding protein regulates the CT element of the human c-myc protooncogene. J Biol Chem 270:9494–9499.

Munroe SH, Dong XF. 1992. Heterogeneous nuclear ribonucleoprotein A1 catalyzes RNA. RNA annealing. Proc Natl Acad Sci USA 89:895–899.

Newport J, Kirschner M. 1982. A major developmental transition in early Xenopus embryos: II. Control of the onset of transcription. Cell 30:687–696.

Pellizzoni L, Lotti F, Maras B, Pierandrei-Amaldi P. 1997. Cellular nucleic acid binding protein binds a conserved region of the 5′ UTR of Xenopus laevis ribosomal protein mRNAs. J Mol Biol 267:264–275.

Pellizzoni L, Lotti F, Rutjes SA, Pierandrei-Amaldi P. 1998. Involvement of the Xenopus laevis Ro60 autoantigen in the alternative interaction of La and CNBP proteins with the 5′UTR of L4 ribosomal protein mRNA. J Mol Biol 281:593–608.

Peng K, Vucetic S, Radivojac P, Brown CJ, Dunker AK, Obradovic Z. 2005. Optimizing long intrinsic disorder predictors with protein evolutionary information. J Bioinform Comput Biol 3:35–60.

Rajavashisth TB, Taylor AK, Andalibi A, Svenson KL, Lusis AJ. 1989. Identification of a zinc finger protein that binds to the sterol regulatory element. Science 245:640–643.

Raman B, Guarnaccia C, Nadassy K, Zakhariev S, Pintar A, Zanuttin F, Frigyes D, Acatrinei C, Vindigni A, Pongor G, Pongor S. 2001. N(omega)-arginine dimethylation modulates the interaction between a Gly/Arg-rich peptide from human nucleolin and nucleic acids. Nucleic Acids Res 29:3377–3384.

Rein A, Henderson LE, Levin JG. 1998. Nucleic-acid-chaperone activity of retroviral nucleocapsid proteins: Significance for viral replication. Trends Biochem Sci 23: 297–301.

Ruble DM, Foster DN. 1998. Molecular cloning and characterization of a highly conserved chicken cellular nucleic acid binding protein cDNA. Gene 218:95–101.

Seenisamy J, Rezler EM, Powell TJ, Tye D, Gokhale V, Joshi CS, Siddiqui-Jain A, Hurley LH. 2004. The dynamic character of the G-quadruplex element in the c-MYC promoter and modification by TMPy P4. J Am Chem Soc 126:8702–8709.

Shimizu K, Chen W, Ashique AM, Moroi R, Li YP. 2003. Molecular cloning, developmental expression, promoter analysis and functional characterization of the mouse CNBP gene. Gene 307:51–62.

Siddiqui-Jain A, Grand CL, Bearss DJ, Hurley LH. 2002. Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. Proc Natl Acad Sci USA 99:11593–11598.

Siebenlist U, Hennighausen L, Battey J, Leder P. 1984. Chromatin structure and protein binding in the putative regulatory region of the c-myc gene in Burkitt lymphoma. Cell 37:381–391.

Simonsson T, Pecinka P, Kubista M. 1998. DNA tetraplex formation in the control region of c-myc. Nucleic Acids Res 26:1167–1172.

Tomonaga T, Michelotti GA, Libutti D, Uy A, Sauer B, Levens D. 1998. Unrestraining genetic processes with a protein-DNA hinge. Mol Cell 1:759–764.

Tompa P, Csermely P. 2004. The role of structural disorder in the function of RNA and protein chaperones. FASEB J 18:1169–1175.

Tsuchihashi Z, Brown PO. 1994. DNA strand exchange and selective DNA annealing promoted by the human immunodeficiency virus type 1 nucleocapsid protein. J Virol 68:5863–5870.

Weeks KM. 1997. Protein-facilitated RNA folding. Curr Opin Struct Biol 7:336–342.

Weiner AMJ, Allende ML, Becker TS, Calcaterra NB. 2007. CNBP mediates neural crest cell expansion by controlling cell proliferation and cell survival during rostral head development. J Cell Biochem, doi 10.1002/jcb.21380.

Yang D, Hurley LH. 2006. Structure of the biologically relevant G-quadruplex in the c-MYC promoter. Nucleosides Nucleotides Nucleic Acids 25:951–968.

Yasuda J, Mashiyama S, Makino R, Ohyama S, Sekiya T, Hayashi K. 1995. Cloning and characterization of rat cellular nucleic acid binding protein (CNBP) cDNA. DNA Res 2:45–49.

Zanotti KJ, Lackey PE, Evans GL, Mihailescu MR. 2006. Thermodynamics of the fragile X mental retardation protein RGG box interactions with G quartet forming RNA. Biochemistry 45:8319–8330.

Zhang Q, Pedigo N, Shenoy S, Khalili K, Kaetzel DM. 2005. Puralpha activates PDGF-A gene transcription via interactions with a G-rich, single-stranded region of the promoter. Gene 348:25–32.

Zuker M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res 31:3406–3415.