A Polypyrimidine/Polypurine Tract within the *Hmga2* Minimal Promoter: A Common Feature of Many Growth-Related Genes†

Alessandra Rustighi, Michela A. Tessari, Fulvia Vascotto,‡ Riccardo Sgarra, Vincenzo Giancotti, and Guidalberto Manfioletti*

Dipartimento di Biochimica, Biofisica e Chimica delle Macromolecole, Università di Trieste, Italy

Received August 15, 2001; Revised Manuscript Received November 1, 2001

ABSTRACT: HMGA2 is an architectural nuclear factor which plays an important role in development and tumorigenesis, but mechanisms regulating its expression are largely unknown. The proximal promoters of the mouse and human genes coding for HMGA2 contain a conserved polypyrimidine/polypurine (ppyr/ppur) element which constitutes a multiple binding site for Sp1 and Sp3 transcription factors. In the present study we report that this region can adopt a single-stranded DNA conformation, as demonstrated in vitro by S1 nuclease sensitivity on supercoiled plasmids, indicative of an intramolecular triple-helical H-DNA structure. Moreover, we find that PTB (polypyrimidine tract binding protein), a member of the hnRNP family, binds the pyrimidine strand of *Hmga2* as well as similar ppyr/ppur elements of the c-Ki-*ras* (R·Y) and c-*myc* P1 promoters. Transfection experiments indicate that non-B-DNA conformers of the ppyr/ppur tract of the *Hmga2* promoter contribute to positive transcriptional activity. We propose a transcriptional mechanism, acting on the *Hmga2* non-B-DNA structure and functioning through interconversion between double-stranded and single-stranded DNA, that seems to be adopted by an increasing number of genes, mainly growth-related.

HMGA2 is a member of the HMGA[1] family of non-histone chromatin-associated proteins which function as general transcription factors (*1*). Three proteins of this family are known, two of them are HMGA1a and HMGA1b, together reported as HMGA1, because they originate by alternative splicing of the same gene and the third component is HMGA2, which arises from a different but related gene (*2*, *3*). All three HMGA proteins share a triplicate common DNA binding domain, called AT *hook*, that interacts with the minor groove of AT-rich DNA sequences (*1*).

HMGA1 proteins are essential structural components of inducible transcriptional multiprotein complexes and, as has been shown for the gene of human β-interferon, are responsible also for the correct timing of enhanceosome assembly and disruption (*4*, *5*). They function as structural factors in protein complexes that modulate the transcription of a growing number of genes such as tumor necrosis factor β, E-selectin, insulin receptor, and several others (*1*, *6*, and references therein). HMGA2 has been shown to have the same role as a transcriptional modulator (*7*). Given that HMGA proteins are responsible for the correct three-dimensional configuration of protein−DNA complexes, it is not surprising that they play a key role in important cellular processes, such as proliferation and differentiation, and recently also a link to apoptosis of leukemic cells has been reported (*8*).

HMGA proteins are expressed at high levels both in transformed cell lines and in tumors having different origins, while in normal adult cells they are not present or only barely detectable (*9*−*11*). This correlation is so consistent that elevated levels of HMGA proteins have been suggested as diagnostic markers of malignant transformation (*12*−*15*). Direct evidence that HMGA participate in the oncogenic process was first provided when the expression of antisense HMGA2 RNA was shown to prevent retrovirally induced neoplastic transformation of rat thyroid cells (*16*), and very recently the oncogenic properties have been extended to all three members of the HMGA family (*17*). HMGA2 dysregulation, as a result of specific chromosomal rearrangements, is also being identified in a variety of common benign mesenchymal tumors, such as lipomas and uterine leiomyomas, making the *HMGA2* gene probably one of the most frequently rearranged genes in human neoplasms (*18*). Transgenic mice expressing a rearranged form of HMGA2

† This work was supported by grants from Associazione Italiana per la Ricerca sul Cancro (AIRC), Milano, Italy, by Ministero della Ricerca Scientifica e Tecnologica, Roma (MURST), Italy (Grant 9806279300), by ASI (Grants ARS-99-47 and I/R/129/00), and by the Università di Trieste, Italy.

* Address correspondence to this author at the Università degli Studi di Trieste, Dipartimento di Biochimica, Biofisica e Chimica delle Macromolecole, via Giorgieri, 1-34127 Trieste, Italy. Phone: 39-040-6763675. Fax: 39-040-6763694. E-mail: manfiole@univ.trieste.it.

‡ Present address: International Centre for Genetic Engineering and Biotechnology, AREA Science Park, Padriciano 99, 34012 Trieste, Italy.

[1] Abbreviations: HMGA, high mobility group A; bp, base pairs; ppyr/ppur, polypyrimidine/polypurine; NSE, nuclease-sensitive element; EMSA, electrophoretic mobility shift assay; TBE, Tris−borate−EDTA buffer; EDTA, ethylenediaminetetraacetic acid; HNBF, single strand *Hmga2*-NSE 1-binding factor; PNK, polynucleotide kinase; RTase, reverse transcriptase; GST, glutathione *S*-transferase; hnRNP, heterogeneous nuclear ribonucleoprotein; PTB, polypyrimidine tract binding protein; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; SV40, Simian virus 40; CMV, citomegalovirus; Sp1, specificity factor 1; CTF/NF-1, CAAT-box binding transcription factor/nuclear factor 1; TFO, triple-forming oligonucleotide.

10.1021/bi011666o CCC: $22.00
Published on Web 12/28/2001

develop lipomas with high incidence, demonstrating a direct role of the protein in the development of these neoplasias (*19*, *20*).

The importance of *Hmga2* in proliferation and differentiation during embryogenesis, with special regard to the adipose tissue development, has been confirmed in vivo since genomic loss of *Hmga2* gives rise to the *pygmy* phenotype in mice (*21*), implicating this gene in the control of fat cell proliferation and the regulation of obesity (*22*).

To define the molecular mechanisms responsible for its regulation, we have previously isolated and characterized the genomic DNA of the mouse *Hmga2* gene (*3*) and identified its promoter region (*23*). One of the most striking findings obtained from the analysis of the *Hmga2* promoter sequence was the presence of a segment of 60 bp with an unusual sequence composition of one strand containing only pyrimidine residues and a complementary purine-rich strand (ppyr/ppur). This segment resides just a few base pairs upstream of the major transcription initiation site and is highly homologous to a region present in the epidermal growth factor receptor (EGF-R) promoter as well as in the 5′ flanking sequences of other growth-related genes including c-myc, insulin receptor (I-R), androgen receptor (AR), c-src, c-Ki-ras, transforming growth factor (TGF) β3, and platelet-derived growth factor (PDGF) A-chain, where it constitutes a functionally significant element (*24*−*31*). In particular, this element has been shown to be essential for gene activity, to interact with transcriptional regulators, to adopt non-B-DNA conformations, such as triple-helical H-DNA, and therefore to be sensitive to S1 nuclease in vitro.

In this report we identify two nuclease-sensitive elements, NSE 1 and NSE 2 in the *Hmga2* promoter region, of which NSE 1 is centered within the ppyr/ppur tract. A detailed analysis of NSE 1 by high-resolution mapping and EMSA suggests an intramolecular triple-helical H-DNA conformation of this tract. Moreover, we find that PTB (polypyrimidine tract binding protein) binds not only to the polypyrimidinic strand of the *Hmga2* promoter but also to other ppyr-containing sequences from the mouse c-Ki-*ras* and human c-*myc* P1 promoters, suggesting a common mechanism of regulation for these genes. Finally, transfection assays indicate a positive role of this non-B-DNA structure in transcription activity.

EXPERIMENTAL PROCEDURES

Recombinant Plasmid Constructs. BaP, 4P, HaP, and AaP, containing *Hmga2* regulatory regions cloned upstream of a luciferase reporter gene of the pGL2 basic vector (Promega), were previously described (*23*). Mutant constructs −63mut, −45mut, −38mut, and −24mut were constructed as follows: supercoiled plasmid BaP was treated with S1 nuclease and afterward was cleaved with *Spe*I (Amersham Pharmacia Biotech). The resulting products were separated on a 1% agarose gel, extracted, and purified, and their extremities were blunted with T4 DNA polymerase and religated. After transformation of *Escherichia coli* (DIH-101) cells the plasmid DNA of several colonies was sequenced, and clones differing in the deletion introduced inside the ppyr/ppur tract were selected. Sp1mut was produced by PCR-mediated site-directed mutagenesis, introducing an *Eco*RI endonuclease recognition site into HaP, from −40 to −35, and confirmed by sequencing.

Enzymatic Probing of Non-B-DNA Structures. Six micrograms of plasmid DNA was digested in a final volume of 100 μL of S1 incubation buffer with 5 units of S1 nuclease (Roche Molecular Biochemicals) for each microgram of DNA at 45 °C for 25 min. The DNA was extracted with phenol−chloroform to remove S1 nuclease and precipitated with ethanol. The S1-treated DNA was digested with the appropriate restriction enzyme, and after ethanol precipitation the digestions were analyzed by electrophoresis on 6% polyacrylamide gels. DNA was blotted onto a nylon membrane (Genescreen NEN, Du Pont), and the filters were hybridized with radiolabeled oligonucleotide AR17 (from +69 to +47) under standard conditions. Filters were exposed to an X-ray film (Hyperfilm MP, Amersham Pharmacia Biotech) using intensifying screens.

High-Resolution Mapping. Twelve micrograms of supercoiled plasmid BaP was treated with 12 units of S1 nuclease in 200 μL of S1 incubation buffer at 45 °C for 5 min. After phenol−chloroform extraction and ethanol precipitation the DNA was digested with *Hin*dIII. Half of the digestion was treated with 2.5 units of calf intestinal phosphatase (Amersham Pharmacia Biotech), phenol−chloroform extracted, ethanol precipitated, and labeled with 10 units of polynucleotide kinase (Epicenter Technologies) using [γ-^{32}P]ATP (3000 Ci/mmol, NEN Life Science Products) for 1 h at 37 °C. The other half-digestion was labeled at the 3′ ends using 20 units of M-MuLV reverse transcriptase (Roche Molecular Biochemicals) with [α-^{32}P]dCTP (3000 Ci/mmol), and then cold dATP and dGTP were added to fill in only the first three positions of the protruding extremities of the *Hin*dIII fragment, in 10 μL final volume at 42 °C for 1 h. Finally, both reactions were treated with *Spe*I and loaded on a 1% agarose gel to recover the ∼690 bp *Spe*I−*Hin*dIII fragment of interest. The same procedure without initial S1 nicking was performed as a control. A total of 5000 cpm of the four different *Spe*I−*Hin*dIII fragments was then separated on a 6% sequencing gel, dried, and autoradiographed. The software Quantity One from Bio-Rad was used for the quantitative analysis of the results.

Electrophoretic Mobility Shift Assay of Triplex Formation. Oligonucleotides used in triplex-forming assays were as follows: T(A)s, 5′-GCCTCCTCCTTTCCTCCTCCTCCTC-3′ (from −84 to −62); T(A)as, 5′-GAGGAGGAGGAG-GAAAGGAGGAGGC-3′; T(BC)as, 5′-GGAGGAGGAA-GAGAGGAGGAGGA-3′ (from −60 to −38); GA3, 5′-TGGGGAGGGTGGGGAAGGTGGGGAGGAGAATT-3′.

For EMSA the labeled oligonucleotide T(A)ds had first to be PAGE-purified. Briefly, 15 pmol of both sense T(A)s and antisense T(A)as oligonucleotides was labeled with 20 pmol of [γ-^{32}P]ATP (3000 Ci/mmol) and annealed. The whole probe was separated on a 15% native PAGE, and duplex oligonucleotide T(A)ds was recovered from the gel by the crush-and-soak method. For triplex formation 100 fmol (5×10^5 cpm) of T(A)ds was mixed with increasing amounts of cold T(BC)as or GA3 in 50 mM Tris−acetate, pH 7.4, 50 mM NaCl, 10 mM $MgCl_2$, 2 mM spermidine, and 6% glycerol in 15 μL and incubated overnight at 37 °C. The samples were electrophoresed through a native 15% polyacrylamide gel in incubation buffer at 7 V/cm at 4 °C. The gels were then fixed, dried, and autoradiographed.

Protein Expression and Preparation of Nuclear Extracts. Glutathione *S*-transferase (GST), GST−hnRNP K, and

GST−PTB were purified by glutathione−agarose affinity chromatography (glutathione Sepharose 4B, Amersham Pharmacia Biotech) from extracts of *E. coli* (BL21 strain DE3) transformed with pGEX-2TK, pGEX-hnRNP K (*32*), and pGEX-KG-PTB (*33*). Fusion proteins were eluted with 10 mM glutathione and checked for purity, correct size, and concentration by SDS−PAGE.

Nuclear extracts from confluent NIH-3T3, FRTL5 Cl2 KiMol, PC Cl3 E1a + raf, and QT6 cells were prepared in a cold room as described (*23*), and protein concentration was determined by the Bradford microassay procedure.

Electrophoretic Mobility Shift Assays (EMSA). Fifty femtomoles of radiolabeled oligonucleotide (30000 cpm) was incubated with 1−4 μg of protein nuclear extract or 1−4 ng of recombinant protein in 20 μL reactions as described previously (*23*). For competition binding reactions the unlabeled competitor was included in the reaction at the indicated molar excess of the labeled probe. After incubation for 20 min at room temperature, samples were loaded onto a native 6% polyacrylamide gel in 45 mM Tris−borate buffer and 1 mM EDTA, pH 8.3 (0.5× TBE), electrophoresed at 4 °C and at 15 V/cm, dried, and autoradiographed. When used, anti-PTB antibodies (*34*), 1:300 dilution, were incubated 20 min at room temperature with the nuclear extracts prior to addition of the labeled probe. For Western blot analysis of EMSA, a band-shift assay was performed as usual but with 100 fmol of probe, and subsequently, proteins were transferred overnight by capillary to a nitrocellulose membrane in 1× TBE buffer containing 0.5 mM DTT. The membrane was then blocked and incubated with a 1:300 dilution of anti-PTB antibodies. After washings the membrane was incubated with a 1:5000 dilution of the horseradish peroxidase-conjugated secondary antibody, and immunocomplexes were visualized with the SuperSignal West Femto substrate (Pierce).

The following oligonucleotides were used in EMSA experiments (mutations are underlined): Cs, 5′-TCTCTTCCTCCTCCCCTCTCTCTTTTT-3′; Cas, 5′-AAAAAGAGAGAGGGGAGGAGGAAGAGA-3′; Cs-mut1, 5′-TCTCTTCCTCC<u>GAATTC</u>CTCTCTTTTT-3′; AP1, 5′-CGCTTGATGAGTCAGCCGGAA-3′; K-ras, 5′-GCTCCCTCCCTCCCTCCTTCCCTCCCTCCC-3′; 18Y, 5′-CTTTCTTCCCTTCCTTTC-3′; BCR7, 5′-AAAGGGTTTTCCGTGTTCCCTGCCCCCCTCCC-3′; CT3, 5′-AATTCTCCTCCCCACCTTCCCCACCCTCCCCA-3′; TC1, 5′-GGCTTCCTCCTTCCTCCTCCTCCCC-3′; AdMLP, 5′-GCGTTCGTCCTCACTCTCTTCCGCATCGTG; AdMLPmut, 5′-GCGTTCGTCCTCACTCT<u>GAG</u>CCGCATCGTG; α_1I3, 5′-CTCTGTCCTCCTGTCCTCC<u>TC</u>TGTCATCCTGTCCTCCTGTCCT.

In Situ UV Cross-Linking. EMSA with labeled Cs oligonucleotide was performed as above. The wet gel was autoradiographed at 4 °C, and the retarded band was excised from the gel. The protein−oligonucleotide complex was cross-linked within the gel slice by irradiation with an UV lamp (300 nm, 50 W) for 7.5 min at a 10 cm distance. The cross-linked complex was eluted by the crush-and-soak method in SDS−PAGE sample buffer overnight. Half-preparation was run in a denaturing 12% SDS−PAGE alongside protein molecular weight markers (Amersham Pharmacia Biotech) for 2 h at 20 V/cm. At the end the gel was stained with Coomassie blue to visualize the molecular weight markers, dried, and autoradiographed.

Southwestern Blot Analysis. Thirty micrograms of nuclear extracts was separated onto a 15% SDS−PAGE. The samples were then transferred to a nitrocellulose filter (Schleicher & Schuell) using the Trans-Blot SD semi-dry transfer cell (Bio-Rad). The membrane was then soaked in decreasing concentrations of guanidine hydrochloride and blocked for 1 h at 4 °C in blocking solution (5% skim milk powder, 0.2% BSA, and 0.1% Tween 20 in phosphate-buffered saline). Subsequently, the membrane was prehybridized for 1 h at 4 °C in binding buffer (50 mM Hepes, pH 7.2, 0.1 M KCl, 1 mM $MgCl_2$, 1 mM DTT, and 10% glycerol) and then hybridized overnight in the same conditions with single-stranded, end-labeled probe (10^5 cpm/mL). The blot was then rinsed twice in binding buffer for 15 min at 4 °C and autoradiographed.

Cell Cultures and Transient Transfections. FRTL5 Cl2 KiMol and PC Cl3 E1a + raf were cultured as described (*9*). QT6 and NIH-3T3 cells were grown in Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum at 37 °C in a humidified 5% CO_2 incubator. For transfections NIH-3T3 cells were plated at a density of 0.5 × 10^6 cells per 60 mm diameter culture dish and transfected by the calcium phosphate coprecipitation procedure as already described (*23*). Briefly, transfection precipitates, containing 5 μg of reporter plasmid and 0.4 μg of either SV40 or CMV β-galactosidase expression vector (Promega) to normalize for transfection efficiency, were applied to subconfluent cells and harvested respectively either 48 or 16 h after transfection. For cotransfection experiments the same amounts of reporter and normalization vectors were used together with 5 μg of Sp1 expression vector pEVR Sp1 or empty vector pRc/CMV. After transfection, cells were harvested, lysed with lysis buffer (Promega), and assayed as described (*23*).

RESULTS

Analysis of Hmga2 Promoter Truncation Mutants. We have previously reported that basal transcription of the *Hmga2* minimal promoter depends essentially on a ppyr/ppur tract of 60 bp, from −84 to −25, which constitutes a multiple binding site for Sp1 transcription factors, and a functional CTF/NF-1 binding site only few nucleotides downstream (*23*). To ascertain the importance for transcription of the whole ppyr/ppur sequence, we have tested in transfection experiments the luciferase activity of the minimal promoter (HaP) with some deletion mutants. Transient transfections were done with −63mut, −45mut, −38mut, and −24mut in HMGA2-expressing NIH-3T3 mouse fibroblasts (*35*), and results are reported in Figure 1. Removal of 20 bp from HaP (−63mut) already results in a 2-fold decrease of promoter activity. Further deletions of the ppyr/ppur tract (−45mut and −38mut) lead to a stepwise decrease in transcriptional activity, while a total deletion of the ppyr/ppur tract (−24mut) is accompanied by a very low transcriptional activity, comparable to that of the promoter-less construct AaP.

Similar results were obtained by transfecting the same constructs in another HMGA2-expressing cell line, PC Cl3 E1a + raf, derived from rat thyroid epithelial cells transformed with the oncogenes *E1A* and v-*raf* (*9*) (data not shown).

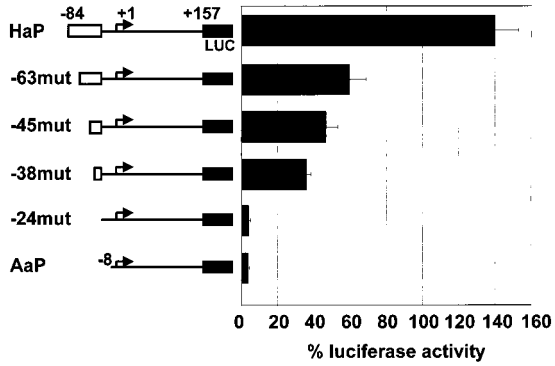

FIGURE 1: Promoter activity of truncation mutants of the minimal *Hmga2* promoter. NIH-3T3 cells were transiently transfected with the promoter construct HaP and a series of its deletion mutants, as represented schematically on the left part of the figure. Open boxes indicate the ppyr/ppur region, and numbering refers to the major transcription initiation site (*23*). Transfected cells were collected 48 h after transfection, and transcriptional activity of each clone is expressed as a percentage of luciferase activity measured for the early SV40 promoter (pGL2 promoter, Promega), which was arbitrarily set to 100%. Columns and bars represent the mean and the standard deviation, respectively, of at least three independent experiments. As additional control two separately obtained constructs, identical to −38mut and −24mut, were assayed in transfection experiments. A vector expressing β-galactosidase from an SV40 promoter was included to normalize for transfection efficiency.

Taken together, these results show the importance of the integrity of the ppyr/ppur region for a positive regulation of the transcriptional activity of the *Hmga2* promoter.

Identification of Two S1 Nuclease-Sensitive Sites, NSE 1 and NSE 2, in the Hmga2 Promoter. In addition to the importance of the ppyr/ppur tract as a multiple binding site for Sp1 transcription factors, such regions of unusual DNA sequence composition are known to confer the ability to adopt non-B-DNA conformations, such as slipped helices, H-DNA, or others which generate extrusions of single-stranded DNA (*36*, *37*). Studies about the role of these non-B-DNA structures in gene regulation indicate that they may be important regulatory elements and are often sensitive to the treatment with single strand specific nucleases, such as nucleases S1 or P1, provided that there is conformational/torsional stress (*36*, *37*). Therefore, to verify the presence of such structures also in the *Hmga2* gene, supercoiled plasmids containing regulatory regions of the *Hmga2* gene, previously used to characterize the promoter region (*23*), were treated with S1 nuclease (Figure 2A). Constructs 4P and HaP, containing respectively 280 and 84 bp of the 5′ flanking region of the *Hmga2* promoter upstream of a luciferase reporter gene, as well as a deletion construct, AaP, in which the ppyr/ppur region has been deleted, were treated with S1 nuclease, digested with *Xba*I, which cuts only once inside the luciferase coding sequence, and separated on a 6% native polyacrylamide gel (Figure 2B). Plasmid 4P produced two fragments of 280 and 520 bp, HaP one fragment of 280 bp, and AaP no fragments at all. The generation of the 280 bp fragment is consistent with S1 cutting in the ppyr/ppur region, which was therefore indicated as NSE 1 (nuclease-sensitive element 1). Consistently, this sequence is present only in constructs 4P and HaP but not in AaP (Figure 2A). In addition, also a weaker band of 520 bp was generated only in the digestion of construct 4P. The presence of this band can be explained by assuming S1

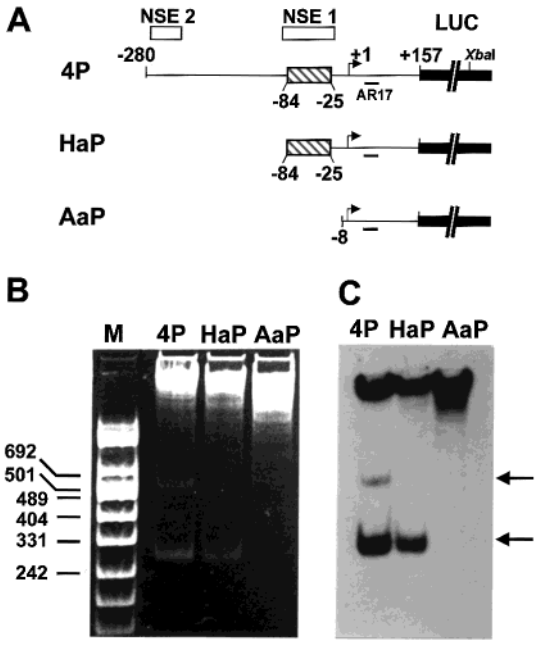

FIGURE 2: S1 nuclease sensitivity of the *Hmga2* promoter. (A) Schematic representation of the three plasmids used in the S1 nuclease sensitivity experiment. The numbers refer to the major transcription initiation site (+1), indicated by a bent arrow. The position of AR17 is denoted by a black bar. The luciferase reporter gene (LUC) is indicated by a black box, while the pyrimidine-rich element (ppyr/ppur) is given by a hatched one. Two open boxes indicate the position of the two mapped S1-sensitive elements NSE 1 and NSE 2. The *Xba*I restriction site within the luciferase gene is indicated. (B) Ethidium bromide stained 6% polyacrylamide gel. Lane M is DNA size markers (MBI, Fermentas), whose principal weights (bp) are indicated on the left. Lanes 4P, HaP, and AaP show the result of treatment with S1 nuclease and subsequent restriction with *Xba*I of the supercoiled plasmids 4P, HaP, and AaP as a negative control. (C) Southern blot of the gel of panel B, hybridized with radiolabeled oligonucleotide AR17 (indicated in panel A). The two arrows on the right indicate the position of the fragments released by the above-mentioned treatment.

cutting in a region at the 5′ end of the promoter sequence in clone 4P (NSE 2), which contains a tract quite rich in pyrimidine bases. The preference of S1 nuclease for NSE 1 is even more evident by comparing the intensity of the two bands recognized by oligonucleotide AR17, derived from a region common to all three plasmids, in the Southern blot of Figure 2C. No bands were observed when supercoiled plasmids were linearized with *Hin*dIII, cutting in the polylinker of the cloning vector, prior to S1 nuclease treatment (data not shown). Thus, formation of the single-stranded region is dependent on supercoiling.

Since NSE 1 maps in the ppyr/ppur region, which has a functional activity in the *Hmga2* promoter, we decided to further characterize this element.

High-Resolution Mapping of NSE 1. In an effort to precisely map NSE 1, we next determined which strand and which nucleotides were cleaved by S1 nuclease. For this purpose supercoiled plasmid BaP (*23*), containing the *Hmga2* promoter region from −495 to +157, was partially digested with S1 nuclease to generate a population of nicked but not linearized DNAs (Figure 3A). The plasmid was then treated with *Hin*dIII, which cuts only once at the end of the polylinker of the cloning vector. To visualize the S1-created nicks separately on the coding (pyr) and on the noncoding strands (pur), one half of the digestion products was labeled

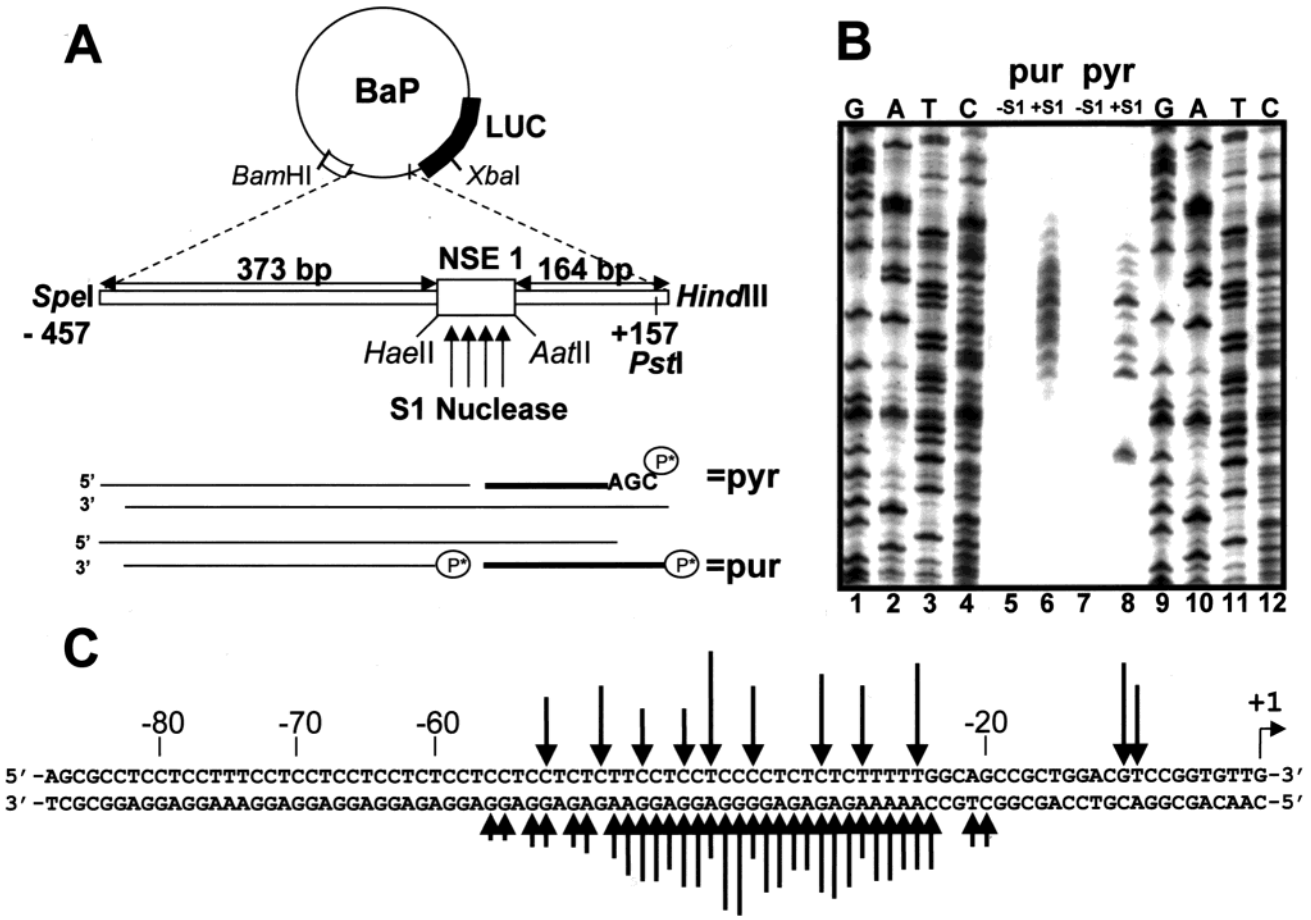

FIGURE 3: High-resolution mapping of the S1 nuclease-sensitive element NSE 1 within the *Hmga2* ppyr/ppur element. (A) Schematic representation of the *Hmga2* promoter containing plasmid BaP. The black box represents the luciferase coding region of pGL2 basic vector (Promega). The open box indicates the nuclease-sensitive element 1. The restriction sites *Spe*I (−450) and *Hin*dIII (pGL2 basic polylinker) used in the fine mapping experiment are shown, and the expected fragment lengths are indicated. The lower part of panel A shows the fragments resulting from the experiment, and bold bars indicate schematically the S1-freed, labeled strands which are visible in the figure of panel B. (B) Lanes 1−4 and 9−12 are sequencing reactions used as molecular size markers. Lanes 5−8 show the fragments of the S1 fine mapping experiment resolved on the same 6% sequencing gel as the sequencing reactions. Equal amounts of cpm have been loaded in each lane. Lanes 5 and 6 are the fragments deriving from the labeling of the purinic strands with PNK. Lanes 7 and 8 show the RTase-labeled pyrimidine fragments. Lanes 5 and 7 are the S1 mock-treated controls. (C) Schematic interpretation of the experiment of panel B. The sequence of *Hmga2* from −88 to +1 is shown. Arrows above and below the DNA sequence represent S1 cleavage sites on the upper (pyr) and on the lower (pur) strand, respectively. The length of the arrows corresponds to the intensity of the bands as quantified by the Quantity One software.

by a fill-in reaction with reverse transcriptase and the other half at the 5′ ends with polynucleotide kinase. Subsequently, the DNA of both halves was digested with *Spe*I, cutting at −457, to obtain only one labeled extremity (Figure 3A), and the partially nicked *Spe*I−*Hin*dIII fragments were resolved on a sequencing polyacrylamide gel (Figure 3B). The results show a collection of labeled DNA fragments that are present on both strands of the S1-digested DNA (Figure 3B, lanes 6 and 8) but not in the S1-mock-treated controls (Figure 3B, lanes 5 and 7). The sites at which S1 nuclease nicked each strand of the DNA are indicated in Figure 3C. The major nicking sites appear to be within −52 and −25 on both strands, although some faint bands were detected also downstream and upstream on the noncoding strand. It is interesting to note that while the nicking on the noncoding strand occurred symmetrically (mirror symmetry around nucleotides −38/−39), that on the coding strand shows an asymmetrical distribution, having two distant nicking sites at −10 and −9.

A detailed inspection of the sequence composition of the ppyr/ppur tract reveals the presence of several different direct and/or mirror repeats. Such symmetries, especially within a ppyr/ppur context, make the formation of several different triple helices of this tract very likely to occur, which could account for the single strandedness of NSE 1 (*37*).

The ppyr/ppur Element of the Hmga2 Promoter Forms Triple-Helical DNA in Vitro. To strengthen the hypothesis of an intramolecular triplex conformation of the NSE 1 in the supercoiled plasmid, we performed a band-shift experiment of intermolecular triplex formation on the basis of the high-resolution mapping results. Therefore, the labeled duplex oligonucleotide T(A)ds, derived from the −84 to −62 region of the ppyr/ppur tract, was used as a double-stranded target and incubated with increasing amounts of cold oligonucleotide T(BC)as (from −60 to −38) as a TFO. This experiment should mimic the situation of a purinic H-DNA form of NSE 1, as shown in the model of Figure 4A, which is more likely to occur at physiological pH (*38*). The reactions were then separated on native 10% polyacrylamide gel in order to discern shifts of the labeled oligonucleotides obtained by triplex formation (Figure 4B). Indeed, oligonucleotide T(BC)as, although having 2 mismatches over 25

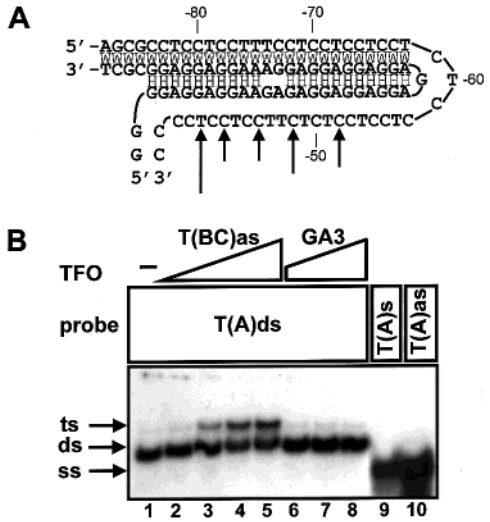

FIGURE 4: Triplex-forming potential of the *Hmga2* NSE 1. (A) Schematic drawing of one of the possible non-B-DNA structures of the *Hmga2* NSE 1. W indicates Watson−Crick interactions, and H stands for Hoogsten bondings. The numbering refers to the major transcription initiation point of the *Hmga2* gene, arrows show the position of S1 nuclease-created nicks. (B) Band-shift analysis of intermolecular triplex formation with double-stranded probe T(A)-ds, from −84 to −62 of the *Hmga2* promoter. Lane 1, free probe. Lanes 2−5: labeled T(A)ds was incubated with increasing amounts of TFO T(BC)as, derived from −60 to −38 of *Hmga2*, in 1:1, 1:10, 1:100, and 1:1000 molar ratios. Lanes 6−8: labeled T(A)ds was incubated with increasing amounts of a GA-rich oligonucleotide derived from the ppyr/ppur tract close to the human c-*myc* P1 promoter (*25*) with 1:10, 1:100, and 1:1000 molar ratios. Lanes 9 and 10 show the migration of single-stranded, labeled oligonucleotides T(A)s and T(A)as, respectively. The arrows on the left indicate the position of single-stranded (ss), duplex (ds), and triplex (ts) DNA.

pairings, with regard to oligonucleotide T(A)ds (Figure 4A), assuming formation of antiparallel reverse Hoogsten bondings, was able to shift the probe with a 10-fold molar excess (Figure 4B, lane 3). The double-stranded target is not shifted completely at 1000-fold concentration, probably because of autoassociation of the G-rich TFO, as already reported (*39*). On the contrary, a GA-rich oligonucleotide, derived from the ppyr/ppur element of the human c-*myc* P1 promoter (*32*), was not able to retard the migration of labeled T(A)ds even at a 1000-fold molar excess (Figure 4B, lanes 6−8), demonstrating that the interaction between T(A)ds and T(BC)as is a high affinity, specific interaction.

Sequence-Specific Nuclear Factors Bind to Single-Stranded DNA in the ppyr/ppur Tract of the Hmga2 Minimal Promoter. Since it appears from fine mapping results that the promoter region present in a single-stranded form involves nucleotides from −52 to −25, band-shift assays were carried out using as probes either the sense (Cs) or the antisense strand (Cas) of this region to search for potential single strand binding factors that could be involved in *Hmga2* gene regulation. When the pyrimidine strand was labeled (Cs) and incubated with nuclear extracts prepared from several HMGA2-expressing cell lines such as NIH-3T3, PC Cl 3 E1A + raf, and QT6 (*9*, *23*, *35*), a specific predominant complex, called "A", was detected for each extract (Figure 5, lanes 2−4). This complex has the same mobility and similar intensity in all nuclear extracts tested and was efficiently competed by a 50-fold molar excess of the self-cold probe (lane 7). Competition with a 50-fold molar excess of tRNA or single-stranded poly(dI·dC) (lanes 8 and 9) was ineffective as well as the use of double-stranded oligonucleotide (Cds) derived from the same region (lane 10), indicating specific interaction between Cs and a single strand specific binding factor(s). The presence of two complexes with lower mobility and weaker intensity can also be observed (lanes 2−4, 6, 12), but their formation is not reproducible using different preparations of nuclear extract from the same cell line. Using the purine strand (Cas) as a probe and incubating with same nuclear extracts, some specific but very weak complexes were detected (data not shown). Since complex A was the predominant single strand binding activity detected, using both sense and antisense oligonucleotides, we focused our attention on this complex, while the nature of the others was not explored further. Therefore, we performed several competition band-shift experiments using oligonucleotide Cs as probe with nuclear extracts from either mouse NIH-3T3 or another HMGA2-expressing rat thyroid cell line, FRTL5 Cl2 KiMol (*9*), and different single-stranded oligonucleotides as competitors, obtaining essentially the same results for both cell types (Figure 5, lanes 11−33). We used oligonucleotide 18Y, consisting of only pyrimidine nucleotides without any marked repetition and oligonucleotide BCR7, derived from the promoter sequence of the human BCR gene (*40*) and characterized by some stretches of irregular pyrimidine nucleotide arrays. 18Y slightly competes at a 100-fold molar excess (lanes 15 and 16), while BCR7 is completely ineffective even at a 100-fold molar excess (lanes 17 and 18). As a control an unrelated single-stranded oligonucleotide with the consensus for the transcription factor AP1 was included (lane 19). These experiments demonstrate therefore that there is a sequence-specific interaction between the identified nuclear protein(s) and the single-stranded polypyrimidine sequence Cs of the *Hmga2* promoter and that the factor(s) does (do) not simply recognize single-stranded DNA or pyrimidine stretches, but rather either a specific structure or a specific pyrimidine-rich sequence or a combination of both.

We next evaluated if other natural ppyr/ppur-containing promoters, known to be regulated through ppyr/ppur regulatory elements, share the interaction with the same complex-forming factor(s) as Cs. Therefore, three pyrimidine-rich single-stranded oligonucleotides, TC-1, K-ras, and CT3, were included as competitors (Figure 5, lanes 20−33). Oligonucleotide TC1 represents the ppyr element of the human c-*Src* promoter and has been chosen because of sequence homology with Cs and because it has been shown to be a specific binding site for the novel pyrimidine binding protein Spy (*28*). Oligonucleotide K-ras derives from the pyrimidine strand of the murine c-Ki-*ras* R·Y promoter site and contains six repeats of the tetranucleotide TCCC (*29*). Finally, also oligonucleotide CT3, containing several repeats of the so-called CT element, which derives from the P1 promoter of human c-*myc* gene, was included (*32*). While oligonucleotide K-ras is able to efficiently compete for complex A formation already at a 20-fold molar excess (lanes 28−30) showing an affinity similar to Cs (lanes 22−24), CT3 (lanes 31−33) has an intermediate affinity, and TC1 weakly competes for the formation of complex A (lanes 25−27). It appears therefore that the factor in complex A, which we call now HNBF (single strand *Hmga2-N*SE1-*b*inding *f*actor), in addition to the NSE 1 of *Hmga2* binds to c-Ki-ras R·Y with

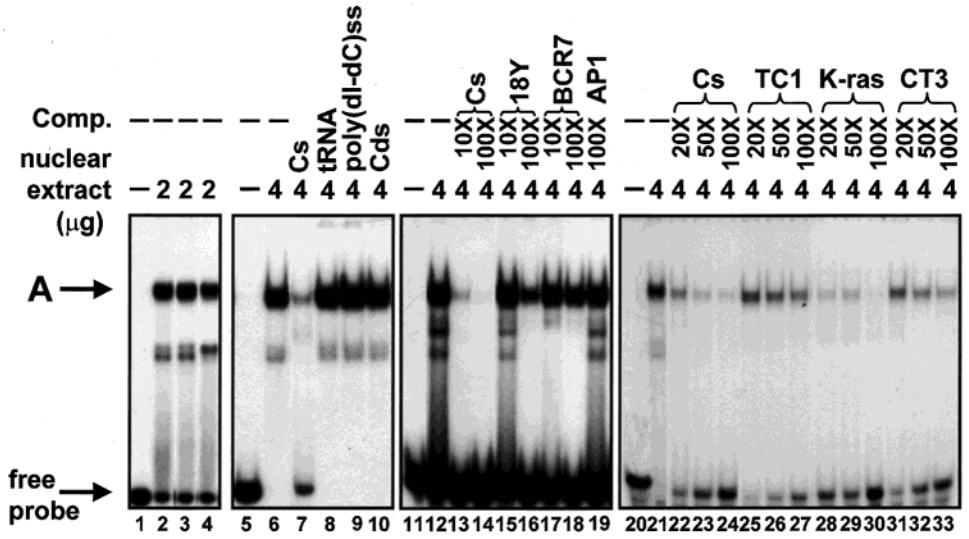

FIGURE 5: Sequence-specific, single strand binding factors interact with the NSE 1 of *Hmga2*. Band-shift experiments using the single-stranded oligonucleotide Cs (from −51 to −25) as a probe. Lane 1: free probe. Lanes 2−4: Cs was incubated with 2 μg of nuclear extract from NIH-3T3, PC Cl3 E1a + raf, and QT6 cells, respectively. Lanes 5−33: competition band-shift experiments with labeled probe Cs, using 4 μg of nuclear extract from NIH-3T3 cells, are shown. Lanes 5, 11, and 20: free probe. Competitions are with the self-Cs single-stranded oligonucleotide, lane 7, with nonspecific single-stranded competitors, lanes 8 and 9, and with double-stranded oligo Cds, lane 10, at 100-fold molar excess. Lanes 13 and 14: competition with the self-Cs. Lanes 15−19: competition with ppyr containing (18Y and BCR 7) and noncontaining (AP1) single-stranded oligonucleotides at the indicated molar excess. Lanes 22−33: competition with single-stranded ppyr oligonucleotides derived from specific promoter elements: Cs = mouse *Hmga2*, TC1 = human c-*src*, K-ras = mouse c-Ki-*ras*, and CT3 = human c-*myc* P1 promoters, at the indicated molar excess.

similar affinity and, also, albeit with much lower affinity to the CT element of c-*myc*.

Molecular Characterization of HNBF. As a first approach to unravel the identity of HNBF, we thought to determine its composition and molecular weight by in situ UV cross-linking. UV-irradiated complex A from an EMSA experiment with probe Cs with nuclear extracts from NIH-3T3 cells was eluted, and the DNA cross-linked proteins were analyzed on a SDS−PAGE. A radioactive doublet was observed with a migration distance of 73 and 78 kDa, as inferred from the calibration curve obtained from the migration in the same gel of protein molecular weight standards, that corresponds to proteins of 63 and 68 kDa, taking into account the molecular weight of the DNA probe (Figure 6A, lane 1). Without UV irradiation, no retarded complex was observed, showing that native interactions were effectively disrupted in denaturing SDS gels (Figure 6A, lane 2). A possible explanation for the doublet observed for the HNBF−DNA complex could simply be the presence of different subunits or, alternatively, the presence of differently migrating HNBF−DNA populations. The presence of a doublet in UV cross-linking experiments has in fact already been reported for polypyrimidine tract binding protein (PTB), a member of the heterogeneous nuclear ribonucleoprotein (hnRNP) family, and explained assuming that each band corresponds to the binding of the DNA strand to different DNA binding domains of PTB (*41*). To investigate these possibilities and to have a more precise evaluation of the molecular weight, we have performed a Southwestern blot analysis. Several crude nuclear extracts already tested for the presence of HNBF were therefore assayed with radiolabeled Cs probe (Figure 6B). One band, with a migration distance indicative of an apparent molecular mass of 62 kDa, is recognized by probe Cs in all extracts tested (Figure 6B, lanes 1−3). From these experiments it thus appears that HNBF is not constituted by different subunits and has an apparent molecular mass of 62 kDa.

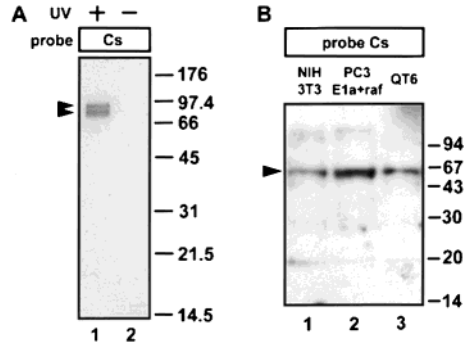

FIGURE 6: Molecular characterization of HNBF. (A) SDS−PAGE analysis of in situ UV cross-linking of complex A from EMSA with NIH-3T3 nuclear extract. Protein size markers (kDa) are indicated on the right of the figure; arrowheads show the position of the protein−DNA complexes. Lane 1: UV cross-linked product of complex A formed by probe Cs. Lane 2: the same as in lane 1 but without UV cross-linking. (B) Southwestern analysis with probe Cs. 30 μg of the indicated nuclear extracts was separated on a 15% SDS−PAGE and transferred to a nitrocellulose membrane. The membrane was then probed with radiolabeled Cs. The arrowhead indicates the recognized band. Protein size markers (kDa) are indicated on the right of the figure.

HNBF Corresponds to Polypyrimidine Tract Binding (PTB) Protein. From a combined literature search looking for proteins which are able to bind single-stranded polypyrimidine DNA and having the estimated molecular mass, we restricted our attention to two members of the hnRNP family, PTB, also known as hnRNP I, and hnRNP K. Both in fact have a molecular mass very similar to that we estimated for HNBF, and both specifically bind polypyrimidine single-stranded DNA tracts (*41*, *42*). We therefore tested whether these proteins were able to bind Cs oligonucleotide from the

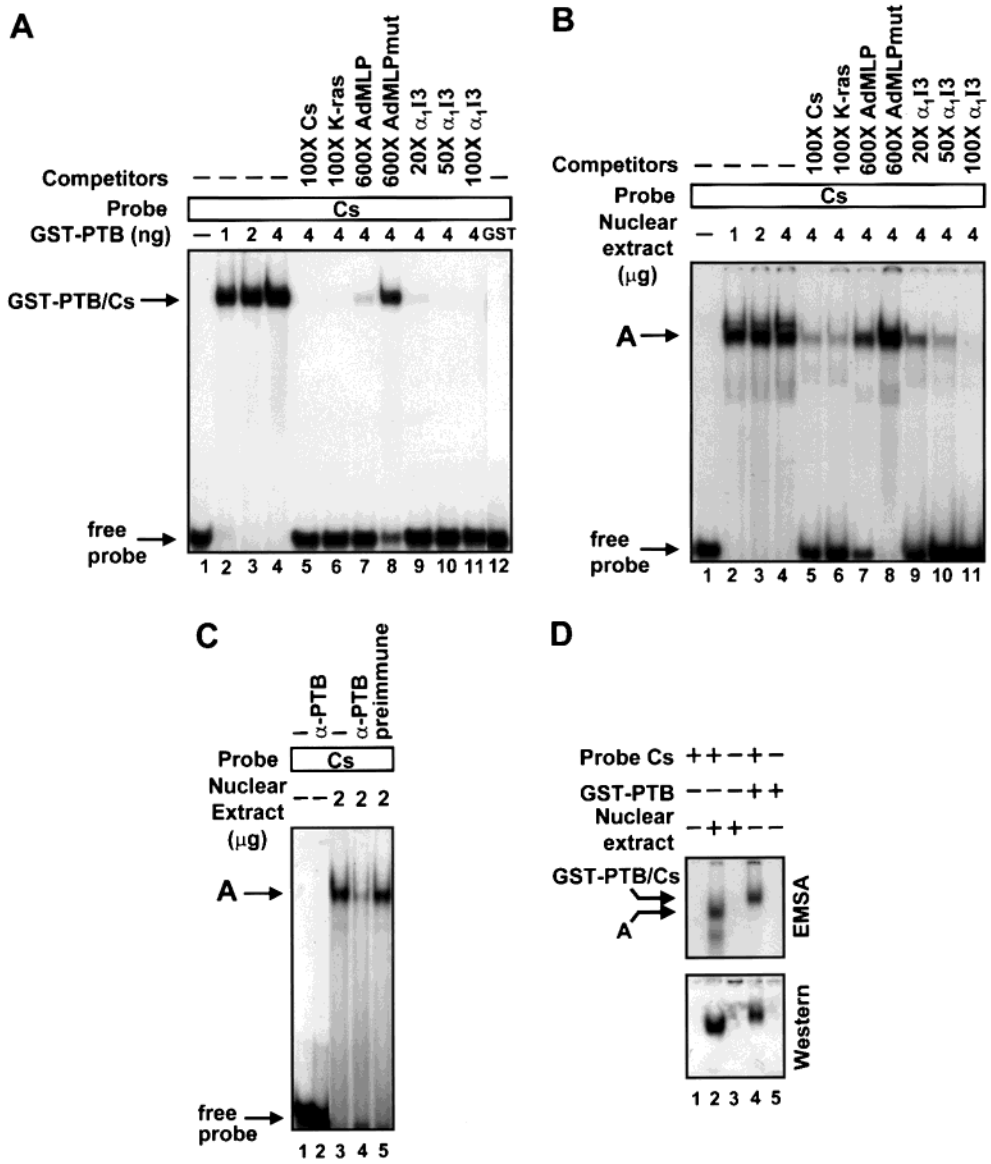

FIGURE 7: Identification of HNBF as PTB. (A) Band-shift analysis with single-stranded labeled probe Cs with increasing amounts of purified recombinant PTB fused to GST (lanes 2−4), as indicated. Lane 1: free probe. Lanes 5−11 show competition band-shifts with different oligonucleotides at the indicated molar excess. In lane 12, 4 μg of GST protein was used as a negative control. (B) Band-shift analysis with single-stranded labeled probe Cs with increasing amounts of NIH-3T3 nuclear extract (lanes 2−4), as indicated. Lane 1: free probe. Lanes 5−11 show competition band-shifts with different oligonucleotides at the indicated molar excess. (C) Band-shift analysis with probe Cs, nuclear extract from NIH-3T3 cells, and anti-PTB antibodies. Lane 1: free probe. 2 μg of nuclear extracts was incubated with the probe in lanes 3−5. Anti-PTB-specific antibodies were added in lanes 2 and 4 while preimmune serum was added in lane 5. (D) Western blot of EMSA. EMSA was performed with the Cs probe, NIH-3T3 nuclear extracts, and recombinant PTB (upper panel, 3 h exposure). Western blot analysis with anti-PTB antibodies of the same gel blotted to a nitrocellulose membrane (lower panel, 6 s exposure). Lane 1: 100 fmol of free probe. 16 μg of NIH-3T3 nuclear extract (lanes 2 and 3) and 80 ng of GST-PTB (lanes 4 and 5) were used.

Hmga2 gene. In the experiment shown in Figure 7A, increasing amounts of purified recombinant PTB fused to GST were incubated with labeled oligonucleotide Cs. A complex is produced (lanes 2−4) which can be effectively competed by 100-fold molar excess of the self-cold probe Cs and oligo K-ras (lanes 5 and 6) while recombinant GST alone is not able to shift the probe (lane 12). Oligo AdMLP from the adenovirus major late promoter (*41*) and oligo α_1I3 from the proteinase α_1-inhibitor 3 variant I gene (*33*), which are low and high affinity binding sites, respectively, for PTB, were included as competitors. Oligo AdMLP is able to specifically compete for complex formation (lane 7) even though at higher molar excess, while the mutated oligo AdMLPmut, harboring a PTB-specific mutation, is not (lane 8). Oligo α_1I3 competes very efficiently already at a 20-fold molar excess (lane 9). The same competitors were used in an EMSA with nuclear extract obtaining the same result (Figure 7B), demonstrating therefore that HNBF and PTB have the same DNA binding specificity and affinity. Recombinant hnRNP K fused to GST was also used in a similar experiment, but no binding was detected with oligo Cs (data not shown).

The identity of HNBF was further investigated using specific antibodies against PTB (*34*). Preincubation of nuclear extracts with PTB-specific antibodies resulted in abrogation of complex A formation (Figure 7C, lane 4). The result is specific since the antiserum was not able to produce any effect in the absence of nuclear extract (lane 2) and the pre-

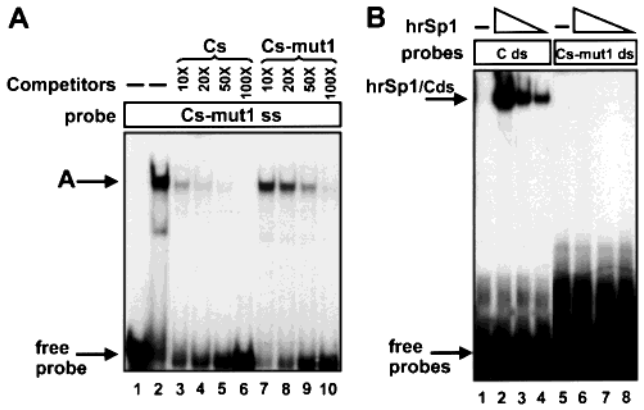

FIGURE 8: Characterization of the binding activities of single- and double-stranded oligonucleotide Cs-mut1. (A) Single-stranded oligonucleotide Cs-mut1, which harbors a 6 base substitution with respect to Cs, is used as a probe in competition band-shift experiments with NIH-3T3 nuclear extract. Lane 1: free probe. Lanes 2−10: labeled Cs-mut1 with 4 μg of nuclear extract. The arrow on the left indicates the position of the complex. Lanes 3−10 show cross-competition with Cs and Cs-mut1 as competitors at the indicated molar excess. (B) Band-shift experiment with double-stranded oligonucleotides Cds, lanes 1−4, and Cs-mut1 ds, lanes 5−8, with decreasing amounts (4, 1, 0.5 footprint forming units), lanes 2−4 and 6−8, of human recombinant Sp1 (Promega). Lanes 1 and 5: free probes.

immune serum did not affect complex A formation (lane 5). To further confirm the identity of the protein, a band-shift was carried out, transferred by capillarity to a nitrocellulose membrane and subjected to Western blot analysis with PTB-specific antibodies. As can be seen in Figure 7D, no band was detected with anti-PTB antibodies in the absence of the DNA probe (lane 3), whereas in the presence of Cs a band was obtained (lane 2) in the same position of that obtained in EMSA (compare the upper with the lower panel of Figure 7D). As a control, the recombinant GST−PTB protein was recognized only in the presence of probe Cs (lane 4) while it is not in its absence (lane 5). From the reported experiments it appears therefore that HNBF is either identical to or highly related to PTB.

Promoter Activity Correlates with Non-B-DNA Structure. To test whether the non-B-DNA structure, we evidenced in vitro, could have an effect in transcription, implying therefore a role for DNA topology in *Hmga2* promoter activity, transfection experiments were performed using both supercoiled and linearized reporter vectors. In addition to HaP and AaP, two mutant reporter constructs, −45mut and Sp1mut, have been transfected. −45mut has only 21 bp of the ppyr/ppur tract of the *Hmga2* promoter left, has lost the S1 nuclease sensitivity (data not shown), and still retains the ability to bind Sp1 transcription factors, since it contains the sequence from −40 to −35, which is the strongest binding site in the ppyr/ppur tract (*23*; data not shown). Sp1mut was created in order to abolish binding of Sp1 to this site but maintaining the ability to bind PTB and the sensitivity to S1. This construct was obtained by replacing the sequence 5′-TCCCCT-3′ (from −40 to −35) with 5′-GAATTC-3′. From the EMSA experiments reported in Figure 8, it can be clearly inferred that single-stranded oligonucleotide Cs-mut1, carrying this mutation, is still able to bind PTB (Figure 8A, lane 2) albeit with reduced affinity, as determined by the cross-competitions with Cs-mut1 and Cs (Figure 8A, lanes 3−6 and 7−10), whereas in its double-

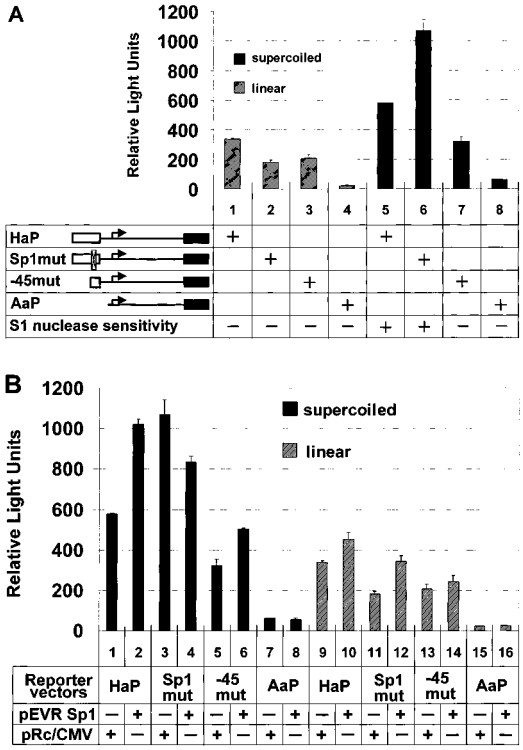

FIGURE 9: Promoter activity correlates with the ability to form non-B-DNA structures. Transfection experiments of *Hmga2* promoter reporter constructs are shown. The transfected cells were harvested 16 h after transfection and assayed for luciferase activity. The mean values of at least three independent experiments are shown. Error bars represent the range of standard errors from three different experiments. A vector expressing β-galactosidase from a CMV promoter was included as internal control. (A) NIH-3T3 cells were transiently transfected with luciferase reporter plasmids either linearized with *Sal*I, lanes 1−4, or supercoiled, lanes 5−8. Reporters are schematically represented in the lower part of the figure along with S1 nuclease sensitivity; the open box represents the ppyr/ppur region and the black box the luciferase reporter gene. (B) NIH-3T3 cells were transiently cotransfected with the indicated *Hmga2* reporter plasmids, in the supercoiled, lanes 1−8, and in the linearized form, lanes 9−16, along with either 5 μg of Sp1 expression vector (pEVR Sp1) or the empty vector (pRc/CMV), as indicated in the lower part of the figure.

stranded form does not any more interact with Sp1 (Figure 8B, lanes 6−8). The resulting construct Sp1mut therefore still retains the whole ppyr/ppur tract and the sensitivity to S1 nuclease (data not shown) but has lost the −40 to −35 Sp1 consensus binding site.

Equal amounts of the described reporter vectors were transfected, either in their supercoiled or as *Sal*I linearized forms, and cells were harvested 16 h after addition of the precipitates because the transfected, supercoiled DNA becomes progressively relaxed during the first day post-transfection, and therefore, collecting cells after 48 h, the effect of supercoiling on the transcription of the reporter cannot be established (*43*).

From the results reported in Figure 9A the activity of the linearized reporter constructs (lanes 1−4) is directly related to the integrity of the ppyr/ppur tract. In fact, both mutants, Sp1mut and −45mut (lanes 2 and 3), have a lower luciferase activity than the wild-type HaP (lane 1). Comparing the reporter activities of transfections with supercoiled plasmids (lanes 5−8) with those of the linearized counterpart (lanes 1−4), there is a 2-fold increase of transcriptional activity

for the supercoiled constructs. This was expected since it is known that gene expression, starting from transfected DNA, depends on DNA topology and supercoiling is known to increase this expression (*44*), while the efficiency of transfection for both supercoiled and linear DNA is the same (*45*). Apart from these considerations, it is surprising that the activity of Sp1mut is higher than HaP (Figure 9A, lanes 6 and 5), suggesting therefore a positive contribution to the reporter activity of non-B-DNA structures of the *Hmga2* promoter. On the contrary, −45mut, which has lost the ability to form non-B-DNA structures, shows, as in the transfections with the linearized reporter, a level of activation comprised between HaP and AaP (lane 7). This result may be explained assuming that in the supercoiled plasmids HaP and Sp1mut there is an equilibrium between B-DNA and single strand extruding conformers, equilibrium which depends on the local concentration of B-DNA binding factors, such as Sp1, and single strand interacting proteins. In Sp1mut this equilibrium is expected to be shifted toward single-stranded DNA, because the mutation impairs binding of Sp1 to its consensus binding site; the contribution of the single strand regulatory element is therefore more evident than in HaP.

If it holds true that single-stranded *cis* elements of Sp1mut are responsible for the observed transcriptional increase, compared to the wild-type promoter, then shifting the balance toward the B-DNA structure should revert the situation. Therefore, we decided to cotransfect an Sp1 expression vector, with the idea that overexpression of this factor, in addition to endogenous Sp1, should lead to a higher promoter occupancy of the whole ppyr/ppur tract by Sp1 proteins and to a shift toward B-DNA even in Sp1mut. Cotransfection of the Sp1 expression vector with supercoiled reporters determines an increase in HaP (Figure 9B, lanes 1 and 2) and −45mut (lanes 5 and 6) and not on AaP (lanes 7 and 8). This increase is modest, probably due to the already high levels of endogenous Sp1 protein, for all the reporters except for AaP where no binding sites for Sp1 are present. These results are consistent with our previous data using a *Drosophila* cell line which is devoid of endogenous Sp1 (*23*). In agreement with our expectations Sp1mut behaves differently, being repressed by Sp1 overexpression (lanes 3 and 4). This result demonstrates therefore that Sp1, by binding sequences adjacent to the NSE 1, acts as a clamp, holding together the double-stranded DNA and shifting therefore the preexisting equilibrium toward the double-stranded DNA. The results obtained cotransfecting Sp1 along with the linearized reporter vectors (Figure 8B, lanes 9−16) are in agreement with this explanation since all reporters, even Sp1mut, are activated by Sp1.

DISCUSSION

In this report we demonstrate that the ppyr/ppur tract present in the mouse *Hmga2* proximal promoter region is able to adopt non-B-DNA conformations which lead to the formation of a single-stranded DNA tract that is specifically recognized by the single strand polypyrimidine binding protein PTB and that has a positive effect on transcription in transfected cells.

Results presented in our previous work (*23*) demonstrated that in the *Hmga2* gene this ppyr/ppur tract is a multiple binding site for Sp1 transcription factors, whose recruitment seems to be essential in TATA-less promoters for the stabilization of the TFIID complex at the transcription initiation site (*46*). The existence of ppyr/ppur stretches in promoter proximal locations in a number of TATA-less mammalian genes has been reported. In addition to *Hmga2* these include EGF-R, c-myc, malic enzyme, I-R, AR, c-src, c-Ki-ras, TGFβ3, PDGF A-chain, and several other genes, mainly growth-related (*24*−*31*, *47*). It has now become clear that such regions are able to adopt unorthodox DNA structures such as H-DNA and slipped helices, and here we show that also the ppyr/ppur tract present in the *Hmga2* promoter is able to adopt such structures under torsional stress. To explain digestions on both the pyrimidinic and purinic strands, we have to assume that more than one such structure coexists in a conformational equilibrium and attack on both strands by S1 nuclease during the state of transition can occur (*48*, *49*). The presence of several direct and mirror repeats in the sequence suggests by itself the potential for generating such structures (*37*); however, the capability of adopting an H-DNA conformation is strongly suggested by our in vitro experiments using a TFO which is able to form an intermolecular triplex DNA.

Studies that examined the role of these structures in gene regulation indicate that they may be important regulatory elements that enhance the transcription of some genes (*43*, *50*, *51*) and act as repressors of others (*52*, *53*). The differences in the effect of the non-B-DNA structures on promoter activity may depend on cell lineages, the position of the element in relation to other elements such as the TATA box, or other double strand specific transcription factor binding sites and, overall, on the nature of nuclear factors recognizing the non-B-DNA structure itself. The case of *Hmga2* is particularly similar to that of the c-*myc* gene; in fact, the CT element of c-*myc* is a binding site for Sp1 when the DNA is double stranded and for hnRNP K when single stranded. Clearly, the binding of single strand- and double strand-requiring proteins is mutually exclusive. hnRNP K activates transcription through single-stranded *cis* elements in vivo, and therefore the transition between single and double strands provides a regulatory device that allows a single DNA segment to bind different factors and confer alternative properties to regulatory sequences (*43*).

We have demonstrated that the ppyr/ppur element of *Hmga2*, when single stranded, is a binding site for PTB which is an abundant and highly conserved factor. PTB was originally identified as a protein interacting with polypyrimidine tracts present in pre-mRNAs and proposed to be involved in different aspects of RNA metabolism, particularly in the regulation of several alternatively spliced genes but also in the process of internal ribosome entry (*54*). PTB has also been shown to interact with single strand pyrimidine-rich DNA sequences of the rat tyrosine aminotransferase liver enhancer, the human transferrin promoter, and the adenovirus major late promoter and in the 5′ region of the rat proteinase α_1-inhibitor 3 variant I gene (*33*, *41*, *55*, *56*). These elements play a positive role in transcription, but the function of PTB in this context has still to be elucidated; in fact, there is no evidence that these elements may exist in vitro or in vivo as single-stranded DNA. In this paper we show that an element which is important in the transcription of the *Hmga2* gene may adopt, at least in vitro, a conformation which involves single-stranded DNA able to bind PTB with high affinity,

implicating therefore for the first time more directly this factor in transcription. Cross-talk between transcriptional and posttranscriptional processes has been suggested (*57*), and an increasing number of proteins appear in fact to be multifunctional, participating in transcriptional and posttranscriptional events. TFIII is required for both the transcription of 5S rRNA genes and the packaging of 5S rRNA, the tumor suppressor gene WT1 first identified as a transcription factor has been involved in splicing, and finally hnRNP K, a pre-mRNA binding protein, appears to be a transcription factor (*57*), only to mention some of the most important cases. It is possible therefore that PTB could play a similar role.

One of the crucial steps in transcription initiation is melting of double-stranded DNA in the promoter region, and conditions that promote melting not only stimulate transcription initiation but also reduce or eliminate the requirements for several general transcription factors (*58*−*60*). Therefore, on the basis of the obtained results we propose the following working model for the *Hmga2* proximal promoter. The stress of supercoiling, which is present in chromatin structure (*61*), should cause the ppyr/ppur tract to melt and adopt one of the several possible non-B-DNA conformations. Binding of PTB to the single-stranded top strand may be sufficient to maintain the open conformation, which in turn may facilitate entry of RNA polymerase into the template. Parameters that locally alter the melting temperature such as matrix attachment, topoisomerase activity, and nucleosome placement and removal might dramatically influence the relative input of both single strand and double strand specific factors. The role of PTB in this context could be simply "architectural" stabilizing the non-B-DNA structure or it could act as a transcription factor increasing the transcription level, as is the case of hnRNP K. The DNA region that binds PTB contains also a binding site for Sp1; therefore, according to this view the final outcome on transcription would depend on the balance between Sp1 and PTB activity that could be influenced, apart from the topology of DNA, by different tissue- or cell-specific regulatory factors, stimuli that could alter the relative level, the DNA affinity, and transcription activity of the two factors. Indeed, in vivo the presence of nucleosomes may severely interfere with the formation and functional significance of the non-B-DNA structures we found in vitro; therefore, chromatinized DNA templates are needed to be considered as substrates for the evaluation of a possible function of non-B-DNA structure on *Hmga2* gene transcription.

Since the time they were first discovered, the expression of HMGA proteins has been correlated with neoplastic transformation being overexpressed in several malignant tumors. Conversely, the expression of these proteins in adult tissues is negligible and is essentially restricted to embryonic development (*21*, *62*). The suppression of HMGA2 synthesis by a vector expressing an antisense construct prevented the neoplastic transformation induced by v-*mos* and v-*ras*, demonstrating the success of such an approach and at the same time suggesting this strategy as a potential therapy for human malignant neoplasias in which HMGA gene overexpression is a general event (*16*). Indeed, very recently this potentiality has been successfully exploited, generating recombinant adenovirus carrying the HMGA1 sequence in the reverse orientation (*63*). Another possibility to suppress HMGA2 protein synthesis would be to target its gene using oligonucleotides that form triple helices at the ppyr/ppur element. This method has in fact been successfully used to restrict the expression of several genes by inhibiting transcription factor binding or by interfering with formation of the initiation complex or blocking elongation, so that oligonucleotides in the last years have become attractive as potential therapeutic agents (*64*). The ability of a TFO to target the double-stranded ppyr/ppur tract that we demonstrate suggests the possibility of using such an approach to block the expression of the protein and therefore expand the therapeutic possibilities for a variety of tumors overexpressing HMGA2.

ACKNOWLEDGMENT

We are grateful to Dr. D. Levens, NCI, Bethesda, MD, for kindly providing hnRNP K expression plasmid, to Dr. M. Schweizer, Department of Biological Sciences, Heriot-Watt University, Edinburgh, U.K., for pGEX-KG-PTB plasmid, to Dr. B. Majello, University Federico II of Naples, Naples, Italy, for Sp1 expression vectors, and to Dr. P. Sharp, MIT Center for Cancer Reserach, Cambridge MA, for providing antiserum against PTB. Acknowledgments are due to Dr. K. M. Vasquez, MD Anderson Cancer Center, University of Texas, Texas, for critical reading of the manuscript and to Dr. M. Ferletta, Uppsala University, Uppsala, Sweden, for technical help in the initial part of the work.

REFERENCES

1. Reeves, R., and Beckerbauer, L. (2001) *Biochim. Biophys. Acta 1519*, 13−29.
2. Manfioletti, G., Giancotti, V., Bandiera, A., Buratti, E., Sautière, P., Cary, P., Crane-Robinson, C., Coles, B., and Goodwin, G. H. (1991) *Nucleic Acids Res. 19*, 6793−6797.
3. Manfioletti, G., Rustighi, A., Mantovani, F., Goodwin, G. H., and Giancotti, V. (1995) *Gene 167*, 249−253.
4. Thanos, D., and Maniatis, T. (1993) *Cell 74*, 887−898.
5. Munshi, N., Agalioti, T., Lomvardas, S., Merika, M., Chen, G., and Thanos, D. (2001) *Science 293*, 1133−1136.
6. Brunetti, A., Manfioletti, G., Chiefari, E., Goldfine, I. D., and Foti, D. (2001) *FASEB J. 15*, 492−500.
7. Mantovani, F., Covaceuszach, S., Rustighi, A., Sgarra, R., Heath, C., Goodwin, G. H., and Manfioletti, G. (1998) *Nucleic Acids Res. 26*, 1433−1439.
8. Diana, F., Sgarra, R., Manfioletti, G., Rustighi, A., Poletto, D., Sciortino, M. T., Mastino, A., and Giancotti, V. (2001) *J. Biol. Chem. 276*, 11354−11361.
9. Giancotti, V., Pani, B., D'Andrea, P., Berlingieri, M. T., Di Fiore, P. P., Fusco, A., Vecchio, G., Philip, R., Crane-Robinson, C., Nicolas, R. H., Wright, C. A., and Goodwin, G. H. (1987) *EMBO J. 6*, 1981−1987.
10. Giancotti, V., Buratti, E., Perissin, L., Zorzet, S., Balmain, A., Portella, G., Fusco, A., and Goodwin, G. H. (1989) *Exp. Cell Res. 184*, 538−545.
11. Giancotti, V., Bandiera, A., Buratti, E., Fusco, A., Marzari, R., Coles, B., and Goodwin, G. H. (1991) *Eur. J. Biochem. 198*, 216−211.
12. Chiappetta, G., Bandiera, A., Berlingieri, M. T., Visconti, R., Manfioletti, G., Battista, S., Martinez-Tello, F. J., Santoro, M., Giancotti, V., and Fusco, A. (1995) *Oncogene 10*, 1307−1313.
13. Fedele, M., Bandiera, A., Chiappetta, G., Battista, S., Viglietto, G., Manfioletti, G., Casamassimi, A., Santoro, M., Giancotti, V., and Fusco, A. (1996) *Cancer Res. 56*, 1896−1901.
14. Bandiera, A., Bonifacio, D., Manfioletti, G., Mantovani, F., Rustighi, A., Zanconati, F., Fusco, A., Di Bonito, L., and Giancotti, V. (1998) *Cancer Res. 58*, 426−431.

15. Abe, N., Watanabe, T., Masaki, T., Mori, T., Sugiyama, M., Uchimura, H., Fujioka, Y., Chiappetta, G., Fusco, A., and Atomi, Y. (2000) *Cancer Res. 60*, 3117−3122.
16. Berlingieri, M. T., Manfioletti, G., Santoro, M., Bandiera, A., Visconti, R., Giancotti, V., and Fusco, A. (1995) *Mol. Cell. Biol. 15*, 1545−1553.
17. Wood, L. J., Maher, J. F., Bunton, T. E., and Resar, L. M. S. (2000) *Cancer Res. 60*, 4256−4261.
18. Tallini, G., and Dal Cin, P. (1999) *Adv. Anat. Pathol. 6*, 237−236.
19. Battista, S., Fidanza, V., Fedele, M., Klein-Szanto, A. J. P., Outwater, E., Brunner, H., Santoro, M., Croce, C. M., and Fusco, A. (1999) *Cancer Res. 59*, 4793−4797.
20. Arlotta, P., Tai, A. K. F., Manfioletti, G., Cliffort, C., Jay, G., and Ono, S. J. (2000) *J. Biol. Chem. 275*, 14394−14400.
21. Zhou, X., Benson, K. F., Ashar, H. R., and Chada, K. (1995) *Nature 376*, 771−774.
22. Anand, A., and Chada, K. (2000) *Nat. Genet. 24*, 377−380.
23. Rustighi, A., Mantovani, F., Fusco, A., Giancotti, V., and Manfioletti, G. (1999) *Biochem. Biophys. Res. Commun. 265*, 439−447.
24. Johnson, A. C., Jinno, Y., and Merlino, G. T. (1988) *Mol. Cell. Biol. 8*, 4174−4184.
25. Kinniburgh, A. J. (1989) *Nucleic Acids Res. 17*, 7771−7778.
26. Tewari, D. S., Cook, D. M., and Taub, R. (1989) *J. Biol. Chem. 264*, 16238−16245.
27. Chen, S., Supakar, P. C., Vellanoweth, R. L., Song, C. S., Chatterjee, B., and Roy, A. K. (1997) *Mol. Endocrinol. 11*, 3−15.
28. Ritchie, S., Boyd, F. M., Wong, J., and Bonham, K. (2000) *J. Biol. Chem. 275*, 847−854.
29. Pestov, D. G., Dayn, A., Siyanova, E. Y., George, D. L., and Mirkin, S. M. (1991) *Nucleic Acids Res. 19*, 6527−6532.
30. Lafyatis, R., Denhez, F., Williams, T., Sporn, M., and Roberts, A. (1991) *Nucleic Acids Res. 19*, 6419−6425.
31. Wang, Z., Lin, X., Qiu, Q., and Deuel, T. F. (1992) *J. Biol. Chem. 267*, 17022−17031.
32. Tomonaga, T., and Levens, D. (1995) *J. Biol. Chem. 270*, 4875−4881.
33. Sickinger, S., and Schweizer, M. (1999) *Biol. Chem. 380*, 1217−1223.
34. Gil, A., Sharp, P. A., Jamison, S. F., and Garcia-Blanco, M. A. (1991) *Genes Dev. 5*, 1224−1236.
35. Lanahan, A., Williams, J. B., Sanders, L. K., and Nathans, D. (1992) *Mol. Cell. Biol. 12*, 3919−3929.
36. Larsen, A., and Weintraub, H. (1982) *Cell 29*, 609−622.
37. Mirkin, S. M., Lyamichev, V. I., Drushlyak, K. N., Dobrynin, V. N., Filippov, S. A., and Frank-Kamenetskii, M. D. (1987) *Nature 330*, 495−497.
38. Beal, P. A., and Dervan, P. B. (1991) *Science 251*, 1360−1363.
39. Henderson, E., Hardin, C. C., Walk, S. K., Tinoco, I., Jr., and Blackburn, E. H. (1987) *Cell 51*, 899−908.
40. Zhu, Q. S., Heisterkamp, N., and Groffen, J. (1990) *Nucleic Acids Res. 18*, 7119−7125.
41. Brunel, F., Zakin, M. M., Buc, H., and Buckle, M. (1996) *Nucleic Acids Res. 24*, 1608−1615.
42. Takimoto, M., Tomonaga, T., Matunis, M., Avigan, M., Krutzsch, H., Dreyfuss, G., and Levens, D. (1993) *J. Biol. Chem. 268*, 18249−18258.
43. Tomonaga, T., and Levens, D. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 5830−5835.
44. Weintraub, H., Cheng, P. F., and Conrad, K. (1986) *Cell 46*, 115−122.
45. Hirt, B. (1967) *J. Mol. Biol. 26*, 365−369.
46. Pugh, B. F., and Tjian, R. (1991) *Genes Dev. 5*, 1935−1945.
47. Xu, G., and Goodridge, A. G. (1996) *J. Biol. Chem. 271*, 16008−16019.
48. Wells, R. D., Collier, D. A., Hanvey, J. C., Shimizu, M., and Wohlrab, F. (1988) *FASEB J. 2*, 2939−2949.
49. Frank-Kamenetskii, M. D., and Mirkin, S. M. (1995) *Annu. Rev. Biochem. 64*, 65−95.
50. Santra, M., Danielson, K. G., and Iozzo, R. V. (1994) *J. Biol. Chem. 269*, 579−587.
51. Ko, J. L., and Loh, H. H. (2001) *J. Biol. Chem. 276*, 788−795.
52. Grossmann, M. E., and Tindall, D. J. (1995) *J. Biol. Chem. 270*, 10968−10975.
53. Bacolla, A., Ulrich, M. J., Larson, J. E., Ley, T. J., and Wells, R. D. (1995) *J. Biol. Chem. 270*, 24556−24563.
54. Valcarcel, J., and Gebauer, F. (1997) *Curr. Biol. 7*, R705−R708.
55. Jansen-Durr, P., Boshart, M., Lupp, B., Bosserhoff, A., Frank, R. W., and Schutz, G. (1992) *Nucleic Acids Res. 20*, 1243−1249.
56. Brunel, F., Alzari, P. M., Ferrara, P., and Zakin, M. M. (1991) *Nucleic Acids Res. 19*, 5237−5245.
57. Ladomery, M. (1997) *Bioassays 19*, 903−909.
58. Pan, G., and Greenblatt, J. (1994) *J. Biol. Chem. 269*, 30101−30104.
59. Holstege, F. C. P., Tantin, D., Carey, M., van der Vliet, P. C., and Timmers, H. T. M. (1995) *EMBO J. 14*, 810−819.
60. Parvin, J. D., and Sharp, P. A. (1993) *Cell 73*, 533−540.
61. Champoux, J. J. (2001) *Annu. Rev. Biochem. 70*, 369−413.
62. Chiappetta, G., Avantaggiato, V., Visconti, R., Fedele, M., Battista, S., Trapasso, F., Merciai, B. M., Fidanza, V., Giancotti, V., Santoro, M., Simeone, A., and Fusco, A. (1996) *Oncogene 5*, 2439−2446.
63. Scala, S., Portella, G., Fedele, M., Chiappetta, G., and Fusco, A. (2000) *Proc. Natl. Acad. Sci. U.S.A. 97*, 4256−4261.
64. Vasquez, K. M., and Wilson, J. H. (1998) *Trends Biochem. Sci. 23*, 4−9.

BI011666O