[CANCER RESEARCH 64, 5728–5735, August 15, 2004]

Identification of the Genes Up- and Down-Regulated by the High Mobility Group A1 (HMGA1) Proteins: Tissue Specificity of the HMGA1-Dependent Gene Regulation

Josefina Martinez Hoyos,[1] Monica Fedele,[1] Sabrina Battista,[1] Francesca Pentimalli,[1,2] Mogens Kruhoffer,[3] Claudio Arra,[4] Torben F. Orntoft,[3] Carlo Maria Croce,[2] and Alfredo Fusco[1]

[1]Dipartimento di Biologia e Patologia Cellulare e Molecolare e/o Istituto di Endocrinologia ed Oncologia Sperimentale del CNR, Facoltà di Medicina e Chirurgia di Napoli, Università degli Studi di Napoli "Federico II," Naples, Italy; [2]Kimmel Cancer Center, Jefferson Medical College, Philadelphia, Pennsylvania; [3]Department of Clinical Biochemistry, Aarhus University Hospital, Aarhus, Denmark; and [4]Istituto Dei Tumori Di Napoli "Fondazione Pascale," Naples, Italy.

ABSTRACT

High mobility group A (HMGA) proteins are chromatinic proteins that do not have transcriptional activity *per se*, however, by interacting with the transcription machinery, they regulate, negatively or positively, the expression of several genes. We searched for genes regulated by HMGA1 proteins using microarray analysis in embryonic stem (ES) cells bearing one or two disrupted *hmga1* alleles. We identified 87 transcripts increased and 163 transcripts decreased of at least 4-fold in *hmga1*−/− ES cells. For some of them, a HMGA1-dose dependency was observed, because an intermediate level was observed in the heterozygous ES cells. When the expression analysis of these genes was extended to embryonic fibroblasts and adult tissues such as heart, spleen, and liver from *hmga1*-knockout mice, contrasting results were obtained. In fact, aside some genes showing the same HMGA1 regulation observed in ES cells, there were some genes that did not modify their expression, and others showing a HMGA1-mediated regulation but in an opposite direction. These results clearly indicate that HMGA1-mediated gene regulation depends on the cellular context. Finally for a couple of analyzed HMGA1-regulated genes, electrophoretic mobility shift assay and chromatin immunoprecipitation revealed a direct binding of HMGA1 proteins to their promoters, suggesting a HMGA1-direct regulation of their expression.

INTRODUCTION

The high mobility group A (HMGA) protein family includes HMGA1a and HMGA1b, which are encoded by *hmga1* through alternative splicing (1) and the closely related HMGA2 protein (2). These proteins bind the minor groove of AT-rich DNA sequences. Their DNA-binding domain is located in the NH_2-terminal region of the protein and contains three short basic repeats, the so-called "AT-hooks." The mammalian HMGA proteins have long been known to play key roles in chromatin architecture and gene control by serving as generalized chromatin effectors, either enhancing or suppressing the ability of more usual transcriptional activators and repressors to act within the confines of chromatinized DNA (3, 4).

HMGA1 proteins seem to play their major physiological role during embryonic development (5). In fact, HMGA1 expression is very high during embryogenesis, whereas it is negligible in normal adult tissues. HMGA1 proteins has been found abundant in malignant neoplasias (6), where their expression appears critical for the acquisition of the neoplastic phenotype (7, 8).

To identify the differentiation pathways in which HMGA1 is involved and to assess the role of the HMGA1 proteins in development, we generated embryonic stem (ES) cells in which one or both *hmga1* alleles are disrupted. We reported recently that *hmga1*−/− ES cells generate less T-cell precursors than do wild-type ES cells after *in vitro*-specific differentiation. Indeed, they preferentially differentiate to B cells, probably consequent to decreased IL-2 expression and increased IL-6 expression, both of which are regulated directly by the HMGA1 proteins (9). Moreover, a lack of HMGA1 expression results in altered hemopoietic differentiation (*i.e.*, there is a reduction in the monocyte/macrophage population and an increase in megakaryocyte precursors, erythropoiesis, and globin gene expression). Re-expression of the *hmga1* gene in *hmga1*−/− ES cells restores the wild-type phenotype (9). These results indicate that drastic changes occur in the transcriptional activity of the *hmga1*−/− cells, and presumably they depend on the modification of the expression of HMGA1-regulated genes.

Using the powerful oligonucleotide microarray hybridization technique, we analyzed the expression profile of ES cells carrying two, one, and no *hmga1* functional alleles to identify the genes that are regulated, positively or negatively, by HMGA1. We screened an array in which 13,059 transcripts were represented, and we identified 87 transcripts that increased and 163 transcripts that decreased with a ≥4-fold change in hmga1−/− ES cells with respect to the wild-type ES cells. Semiquantitative and quantitative reverse transcription (RT)-PCR confirmed the differential expression between wild-type and *hmga1*-knockout ES cells. We obtained different results when we measured the expression of these genes in murine embryonic fibroblasts (MEF) and various adult tissues from *hmga1* knockout mice. The differential expression of some genes matched that found in ES cells, whereas the expression of other genes was either unchanged or opposite to that found in ES cells. Finally, electrophoretic mobility shift assay and chromatin immunoprecipitation experiments demonstrated that HMGA1 proteins bind to the promoters of some representative HMGA1-regulated genes, indicating a direct role of HMGA1 in the regulation of their transcription.

MATERIALS AND METHODS

Cell Culture. The generation and the culture of *hmga1*+/− and *hmga1*−/− ES cells are described elsewhere (9). MEF have been established from wild-type, *hmga1*+/− and *hmga1*−/− embryos 12.5 days post-coitum following standard procedures. They were grown in DMEM medium supplemented with 10% FCS, glutamine and non-essential aminoacids (Life Technologies, Inc.) in a 7% CO_2 atmosphere. Functioning rat thyroid line (FRTL)-5, FRTL-5-KiMSV and FRTL-5–HMGA1as-KiMSV cells, and their culture conditions are reported elsewhere (8).

RNA Extraction from Tissues and Cells. Tissues were snap-frozen in liquid nitrogen and stored at −80°C until use. Total RNAs were extracted from tissues and cell culture using TRI REAGENT (Molecular Research Center, Inc.) solution, according to the manufacturer's instructions. The integrity of the

Received 4/21/04; revised 6/17/04; accepted 6/23/04.

Grant support: from the Associazione Italiana Ricerca sul Cancro (AIRC), Progetto Strategico Oncologia Consiglio Nazionale delle Ricerche, the Ministero dell'Università e della Ricerca Scientifica e Tecnologica (MIUR), "Piani di Potenziamento della Rete Scientifica e Tecnologica" CLUSTER C-04, and the Programma Italia-USA sulla Terapia dei Tumori coordinated by Prof. Cesare Peschle, and Ministero della Salute."

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Requests for reprints: Alfredo Fusco, Dipartimento di Biologia e Patologia Cellulare e Molecolare e/o Instituto di Endocrinologia ed Oncologia Sperimentale del CNR, Facoltà di Medicina e Chirurgia di Napoli, Università degli Studi di Napoli "Federico II," via Pansini 5, 80131 Napoles, Italy. Phone: 39-081-746-3749 or 3056; fax: 39-081-746-3037; E-mail: afusco@napoli.com.

RNA was assessed by denaturing agarose gel electrophoresis (virtual presence of sharp 28S and 18S bands) and spectrophotometry.

Microarray Analysis. Microarray analysis was performed as described previously in detail.[5] Briefly, cRNA was prepared from 8 μg of total RNA, hybridized to MG-U74 Affymetrix oligonucleotide arrays (containing 13,059 murine transcripts), scanned, and analyzed according to Affymetrix (Santa Clara, CA) protocols. Scanned image files were visually inspected for artifacts and normalized by using GENECHIP 3.3 software (Affymetrix). Comparisons were made for each mutated sample *versus* wild-type sample, taking the wild-type sample as baseline by using GENECHIP 3.3. The fold-change values, indicating the relative change in the expression levels between mutated samples and the wild-type sample, were used to identify genes differentially expressed between these conditions.

Semiquantitative and Quantitative RT-PCR. RNAs were treated with DNaseI (Invitrogen) and reverse-transcribed using random exonucleotides as primers and MuLV-reverse transcriptase (Perkin-Elmer). To ensure that RNA samples were not contaminated with DNA, negative controls were obtained by performing PCR on samples that were not reversed-transcribed but otherwise identically processed.

The PCRs were performed with the same RNAs used for array analysis, and the primers sequences are available upon request. For semiquantitative PCR, reactions were optimized for the number of cycles to ensure product intensity within the linear phase of amplification. The PCR products were separated on a 2% agarose gel, stained with ethidium bromide, and scanned using a Typhoon 9200 scanner. Digitized data were analyzed using Imagequant (Molecular Dynamics).

Quantitative PCR was performed with SYBR Green PCR Master Mix (Applied Biosystems) as follows: 95°C 10 minutes and 40 cycles (95°C 15 seconds and 60°C 1 minute).

Protein Extraction, Western Blotting, and Antibodies. Tissues and cell culture were lysed in buffer 1% NP40, 1 mmol/L EDTA, 50 mmol/L Tris-HCl (pH 7.5), and 150 mmol/L NaCl, supplemented with complete protease inhibitors mixture (Roche Diagnostic Corp.). Total proteins were separated by SDS-PAGE and transferred to nitrocellulose membranes (Bio-Rad). Membranes were blocked with 5% nonfat milk and incubated with antibodies against Id3 and tubulin. All of them were purchased from Santa Cruz Biotechnology. Bound antibody was detected by the appropriate secondary antibody and revealed with an enhanced chemiluminescence system (Amersham-Pharmacia Biotech).

Electrophoretic Mobility-Shift Assay. DNA-binding assays with the purified proteins were performed as described previously (10). Five to 20 ng of recombinant protein were incubated in the presence of radiolabeled oligonucleotide. A 200-fold excess of specific unlabeled competitor oligonucleotide was added. The double-strand oligonucleotides used were Id3 spanning from base −632 to −615 of the murine Id3 promoter region, (5′-tgatttttttttttttt-tcaaatctg-3′; ref. 11) and p96 spanning from base −901 to −872 of the 5′ untranslated region of the murine p96 gene (5′-aagaaatatttgatatttttttcttttatcc-3′; Ref. 12).

The same oligonucleotides were also used in binding assays with total extract from wild-type and *hmga1*-knockout murine spleen tissues. Eight micrograms of extracts were incubated in 20 mmol/L HEPES (pH 7.6), 40 mmol/L KCl, 0.1 mmol/L EDTA, 0.5 mmol/L $MgCl_2$, 0.5 mmol/L DTT, and 0.1 mmol/L phenylmethylsulfonyl fluoride in a volume of 20 μl containing 1 μg of poly(dC-dG), 2 μg of BSA, and 10% glycerol, for 10 minutes at room temperature. Binding reactions were incubated for 10 minutes after the addition of 2.5 fmol of a ^{32}P-end-labeled oligonucleotide (specific activity, 8,000–20,000 cpm/fmol). For the antibody supershift analysis, the reactions were performed by preincubating extracts with 0.5 μg of antibody anti-HMGA1 (Santa Cruz Biotechnology) on ice for a minimum of 30 minutes.

The DNA-protein complexes were resolved on 6% non-denaturing acrylamide gels and visualized by exposure to autoradiographic films.

Chromatin Immunoprecipitation. Approximately 3×10^7 wild-type and *hmga1*-knockout MEF were grown on 75-cm^2 dishes and cross-linked by the addition of formaldehyde (to 1% final concentration) to attached cells. Cross-linking was allowed to proceed at room temperature for 5 minutes and was terminated with glycine (final concentration, 0.125 mol/L). Cells were collected and lysed in buffer containing 5 mmol/L PIPES (pH 8.0), 85 mmol/L KCl, 0.5% NP40, and protease inhibitors (1 mmol/L phenylmethylsulfonyl fluoride, 10 μg/ml aprotinin, 10 μg/ml leupeptin), on ice for 10 minutes. Nuclei were pelleted by centrifugation at 5,000 rpm for 5 minutes at 4°C and resuspended in buffer containing 50 mmol/L Tris-Cl (pH 8.1), 10 mmol/L EDTA, 1% SDS, the same protease inhibitors, and incubated on ice for 10 minutes. Chromatin was sonicated on ice to an average length of about 400 bp with a Branson sonicator model 250. Samples were centrifuged at 14,000 rpm for 10 minutes at 4°C. Chromatin was pre-cleared with protein G Sepharose (blocked previously with 1 mg/ml BSA) at 4°C for 2 hours. Pre-cleared chromatin of each sample was incubated with 2 μg of antibody anti-HMGA1 at 4°C overnight. An aliquot of wild-type sample was incubated also with anti-IgG antibody. Next, 60 μl of a 50% slurry of blocked protein G Sepharose was added, and immune complexes were recovered. The supernatants were saved as "input." Immunoprecipitates were washed twice with 2 mmol/L EDTA, 50 mmol/L Tris-Cl (pH 8.0) buffer and 4 times with 100 mmol/L Tris-Cl (pH 8.0), 500 mmol/L LiCl, 1% NP40, and 1% deoxycholic acid buffer. The antibody-bound chromatin was eluted from the beads with 200 μl of elution buffer (50 mmol/L NaHCO3, 1% SDS). Samples were incubated at 67°C for 5 hours in the presence of 10 μg RNase and NaCl to a final concentration of 0.3 mol/L to reverse formaldehyde cross-links. Samples were then precipitated with ethanol at −20°C overnight. Pellets were resuspended in 10 mmol/L Tris (pH 8)-1 mM EDTA and treated with proteinase K to a final concentration of 0.5 mg/ml at 45°C for 1 hour. DNA was extracted with phenol/chloroform/isoamyl alcohol, ethanol-precipitated, and resuspended in water. Input DNA and immunoprecipitated DNAs were analyzed by PCR for the presence of Id3 and p96 promoter sequences. PCR reactions were performed with AmpliTaq gold DNA polymerase (Perkin-Elmer). The primers used to amplify the sequence of the Id3 promoter were 5′-agggtttatgcagcaag-cac-3′ (forward) and 5′-atttgctgctcgtctgacct-3′ (reverse). The primers used to amplify the sequence of the p96 promoter were 5′-aactccagctgtgtcaagtt-3′ (forward) and 5′-gaaagaaagagaggggaaag-3′ (reverse). PCR products were resolved on a 2% agarose gel, stained with ethidium bromide, and scanned using a Typhoon 9200 scanner.

RESULTS

Gene Expression Profile Analysis. RNAs extracted from wild-type, hmga1+/− and hmga1−/− ES cells were hybridized to MG-U74 Affymetrix oligonucleotide arrays containing 13,059 transcripts. The expression profile of the heterozygous and homozygous ES cells was compared with that of the wild-type ES cells that were used as a common reference. The number of transcripts increased or decreased in the heterozygous and the homozygous mutant versus wild-type sample is shown in the Fig. 1. Of the 13,059 transcripts represented on the array, 1,300 had a 2- to 3-, 313 had a 3- to 4-, 227 had a 4- to 10-,

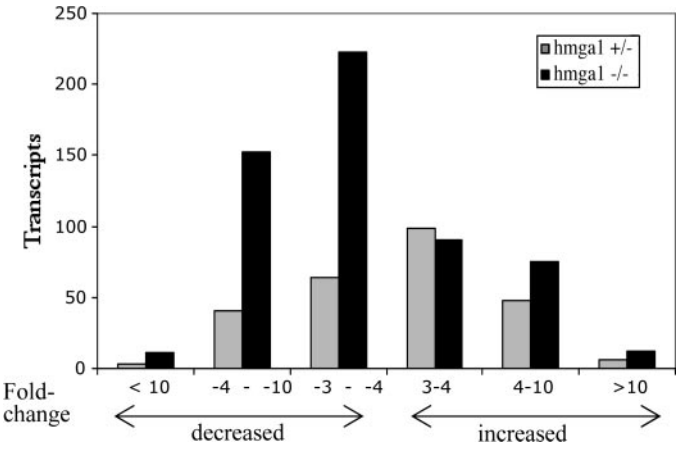

Fig. 1. Gene expression profile in ES cells carrying one or two *hmga1* disrupted alleles. The expression profile of the heterozygous and the homozygous *hmga1*-knockout ES cells was compared with that of the wild-type ES cells that were used as a common reference. Transcripts are grouped according to fold change.

[5] http://www.cancergenetics.med.ohio-state.edu/microarray.

Table 1 *The genes differentially expressed with a ≥4-fold change in heterozygous (+/−) and homozygous (−/−) hmga1-knockout ES cells were grouped in families*

Description	GenBank accession no.	FC +/−	FC −/−
A. Signal transduction			
Mus musculus parathyroid hormone/parathyroid hormone related-peptide receptor (PTHR) gene, exons 10–15, complete cds.	L34611	1.1	18.4
M. musculus mRNA for ryudocan core protein, complete cds.	D89571	3.9	10.6
M. musculus mitogen-responsive M_r 96,000 phosphoprotein p96 mRNA, alternatively spliced p67 mRNA, and alternatively spliced p93 mRNA, complete cds.	U18869	1.7	9.3
M. musculus mRNA for MAP kinase-activated protein kinase 2	BC063064	10	4.7
Mouse Wnt-6 mRNA, complete cds.	M89800	1.1	4.1
M. musculus JIP-1 (JIP-1) mRNA, complete cds.	AF003115	−1.7	−4
Murine macrophage gene, encoding bmk (B cell/myeloid kinase).	J03023	−1.4	−4.1
M. musculus TGF-β-inducible protein (TSC-36) mRNA, complete cds.	M91380	1.3	−4.2
M. musculus mRNA for retinoic acid receptor-α.	X57528	1.1	−4.3
Mouse growth factor-inducible protein (pip92) mRNA, complete cds.	M59821	−2.5	−4.4
Mouse oxytocin-neurophysin I gene, complete cds.	M88355	1.7	−4.5
M. musculus-patched mRNA, complete cds	U46155	−1.2	−4.6
Mouse G protein α subunit (GNA-15) mRNA, complete cds.	M80632	−9.4	−4.7
Mouse mRNA for NBL4, complete cds.	D28818	−2.1	−5.1
M. musculus receptor protein tyrosine phosphatase-λ (ptp-λ) mRNA, complete cds.	U55057	1.3	−5.4
M. musculus protein-serine/threonine kinase (pim-2) mRNA, complete cds.	L41495	−1.7	−5.6
Mouse (clone M1) GTPase (Ran) mRNA, complete cds	L32751	2.6	−5.7
M. musculus ShcC mRNA, complete cds.	U46854	1.3	−6.3
M. musculus syk mRNA for protein-tyrosine kinase.	Z49877	−1.3	−6.3
M. musculus c-Src kinase (Csk) mRNA, complete cds.	U05247	−1.6	−8.6
B. Transcription factors			
M. musculus Thing1 mRNA, complete cds.	U21226	3.8	14
House mouse; Musculus domesticus adult testis mRNA for truncated form of Sox17, complete cds.	D49473	1.3	13.7
M. musculus melanocyte-specific gene 1 (msg1) mRNA, complete cds.	U65091	1.4	9.7
M. musculus mRNA for fos-related antigen-2	X83971	3.1	4
M. musculus Sox4 (Sox4) mRNA, partial cds.	ET62444	−1.4	−4.1
Mouse c-fos oncogene.	V00727	−12	−4.2
Mouse c-myc gene exon 3.	L00039	−4.3	−4.3
M. musculus AP-2.2 gene.	X94694	−1.6	−4.4
Mouse mRNA for Zfp-57, complete cds.	D21850	−2.5	−4.6
M. musculus Kruppel-like factor LKLF mRNA, complete cds.	U25096	−1.4	−4.8
M. musculus transcription factor TFEB mRNA, partial cds.	U36393	−2.2	−4.9
M. musculus transcription factor NF-YC subunit mRNA, complete cds	U62297	−2.6	−5.2
M. musculus DNA binding protein NFI-X (NfiX) mRNA, partial cds	U57636	−1.9	−6.2
M. musculus mRNA for TIF1 β protein	X99644	−5.8	−6.2
M. musculus transcription factor junB (junB) gene, 5 region and complete cds.	U20735	−3.5	−6.4
M. musculus Oct-6 mRNA for octamer binding protein.	X57482	−8.1	−6.5
Mouse helix-loop-helix protein (Id related) mRNA, complete cds.	M60523	−5.3	−7.4
Mouse growth factor-induced protein (zif/268) mRNA, complete cds.	M22326	−8.2	−24.9
C. Cell proliferation			
M. musculus chop-10 mRNA.	X67083	15.1	16.9
Mouse D-type cyclin (CYL2) mRNA, complete cds.	M83749	−2.5	9
M. musculus gly96 mRNA.	X67644	4.4	8
M. musculus mRNA for zyxin.	Y07711	1.3	5.3
D. Extracellular-matrix and cellular-structure			
M. musculus laminin A chain mRNA, complete cds.	J04064	1.2	14
M. musculus α-2 type IV collagen mRNA, complete cds.	J04695	1.4	9.8
M. musculus α-1 type IV collagen (Col4a-1) mRNA, complete cds.	J04694	1.4	9.1
Mouse α-B2-crystallin gene, complete cds	M73741	1.4	7.7
Mouse laminin B1 mRNA, complete cds.	M15525	1.3	6.3
Mouse mRNA for cysteine-rich glycoprotein SPARC	NM_009242	1.2	5.5
Mouse gelsolin gene, complete cds.	J04953	1.5	4.5
M. musculus mRNA for hair keratin, mHb6.	X99143	3.5	4.1
Mouse α-B crystallin mRNA.	M63170	1	4
Mouse COL1A2 mRNA for pro-α-2(I) collagen.	X58251	−3.8	−4.1
M. musculus ρB gene.	X99963	−7.4	−4.2
M. musculus neurofibromatosis 2 (Nf-2) gene mRNA, complete cds.	L27090	1.3	−4.5
M. musculus SM22 α mRNA, complete cds.	L41154	−1.5	−4.5
Mouse COL1A2 mRNA for pro-α-2(I) collagen	NM_007743	−3.9	−4.7
M. musculus α 1 type I collagen gene, partial cds and 3 flanking region	U50767	−1.9	−5.6
M. musculus mRNA for myosin I	X97650	−4.6	−6.3
Mouse tau microtubule binding protein mRNA, complete cds.	M18776	−1.3	−6.6
M. musculus FVB/N collagen pro-α-1 type I chain mRNA, complete cds.	U08020	−2.1	−7.5
Mouse mRNA for vascular smooth muscle α-actin.	X13297	−2.1	−8.6
E. Metabolism, transport and secretion			
Mouse mRNA for mastocytoma proteoglycan core protein, serglycin.	X16133	2.6	40.9
M. musculus cathepsin H prepropeptide (ctsH) mRNA, complete cds.	U06119	1.5	17.3
M. musculus calcium binding protein D-9k mRNA, complete cds.	AF028071	2.3	9.8
Mouse serine protease inhibitor homologue (J6) mRNA, complete cds	J05609	1.7	8.7
M. musculus heparan sulfate D-glucosaminyl 3-*O*-sulfotransferase-1 precursor (3OST1) mRNA, alternatively spliced, complete cds.	AF019385	−1.4	6.5
Mouse mRNA for preproinsulin-like growth factor IA.	X04480	−1.5	6.2
M. musculus preprodipeptidyl peptidase I mRNA, complete cds.	U89269	1.2	5.9
M. musculus mRNA for cytoplasmic dynein heavy chain (partial, ID mdhc10).	ET63396	1.6	5.5
M. musculus steroid cytochrome p450 7-α hydroxylase mRNA, complete cds.	L06463	−1.1	−4
M. musculus very-long-chain acyl-CoA synthetase (VLCS) mRNA, complete cds.	AF033031	−2	−4
M. musculus (clone E31.1 in pGEM7Zf(+)) N-acetylglucosaminyltransferase I mRNA, complete cds.	L07037	2.5	−4.1
Mouse mRNA for inward rectifier K+ channel, complete cds.	D50581	1	−4.1
M. musculus preprocortistatin (Cort) mRNA, complete cds.	AF013253	2.7	−4.3
M. musculus extracellular superoxide dismutase (SOD3) mRNA, complete cds.	U38261	1.2	−4.6

Table 1 *Continued*

Description	GenBank accession no.	FC +/−	FC −/−
Mouse mRNA for a preprothyrotropin-releasing hormone.	X59387	−1.1	−4.7
M. musculus steroid sulfatase (Sts) mRNA, complete cds.	U37545	−2.3	−4.7
M. musculus mRNA for dihydropyrimidinase related protein 4, complete cds.	AB006715	5.7	−4.9
M. musculus putative chloride channel protein CLC6 (Clc6), exon 23 and complete cds.	AF030106	−1	−4.9
Mouse placental alkaline phospatase mRNA, complete cds.	J02980	−1.1	−4.9
M. musculus Balb/c cytochrome *c* oxidase subunit VIaH mRNA, complete cds.	U08439	3.1	−5
Mouse metallothionein-III gene, complete cds.	M93310	1.4	−5.3
M. musculus carboxypeptidase E (Cpe) mRNA, complete cds.	U23184	−3.5	−5.8
M. musculus hormone-sensitive lipase mRNA, complete cds.	U08188	−1.7	−8
M. musculus ADP-ribosylation factor-like protein 4 mRNA, complete cds.	U76546	−2.4	−8.7
F. Growth factors and related proteins			
M. musculus lefty gene.	AJ000083	−1.1	−4.3
M. musculus mRNA for insulin-like growth factor binding protein-3	X81581	−3.2	−6.4
M. musculus mRNA for insulin-like growth factor binding protein-4.	X81582	−1.4	−7.3
M. musculus acid labile subunit (ALS) gene, complete cds.	U66900	2.4	−7.4
Mouse Cyr61 mRNA, complete cds.	M32490	−4.2	−8.4
M. musculus follistatin-like protein (mac25) mRNA, complete cds.	L75822	−2.6	−11
G. Immune functions			
Mouse gene for M_r 47,000 heat shock protein(HSP47), exon 6.	D12907	1.1	5.7
M. musculus MHC class I B(2)-microglobulin gene (W4 allele), partial cds	M84366	1.2	5
Mouse lymphocyte differentiation antigen (Ly-6.2) mRNA, complete cds.	M18184	1.7	4.9
M. musculus anti-digoxin immunoglobulin heavy chain variable region precursor mRNA, partial cds.	ET62206	2	4.7
Mouse CD19 gene, exons 6–15.	M62553	3.3	−4.1
M. musculus (clone B6) myeloid secondary granule protein mRNA.	L37297	−2	−4.3
M. musculus putative TNF-resistance related protein mRNA, complete cds.	U90926	−1.5	−6.5
Mouse mRNA for poliovirus receptor homolog protein soluble form, complete cds.	D26107	−3.7	−12.2
H. Other functions			
Mouse surfeit locus surfeit 3 protein gene	M14689	3.3	4.5
Mouse ERA-1–993 mRNA, complete cds, and alternate ERA-1–339 mRNA, complete cds.	M22115	3.3	4.1
M. musculus imprinted in placenta and liver (Ipl) gene, complete cds.	AF002708	−1.7	−4.1
M. musculus H19 mRNA.	X58196	2.6	−5.9

Abbreviations: cds., coding sequence; FC, fold change.

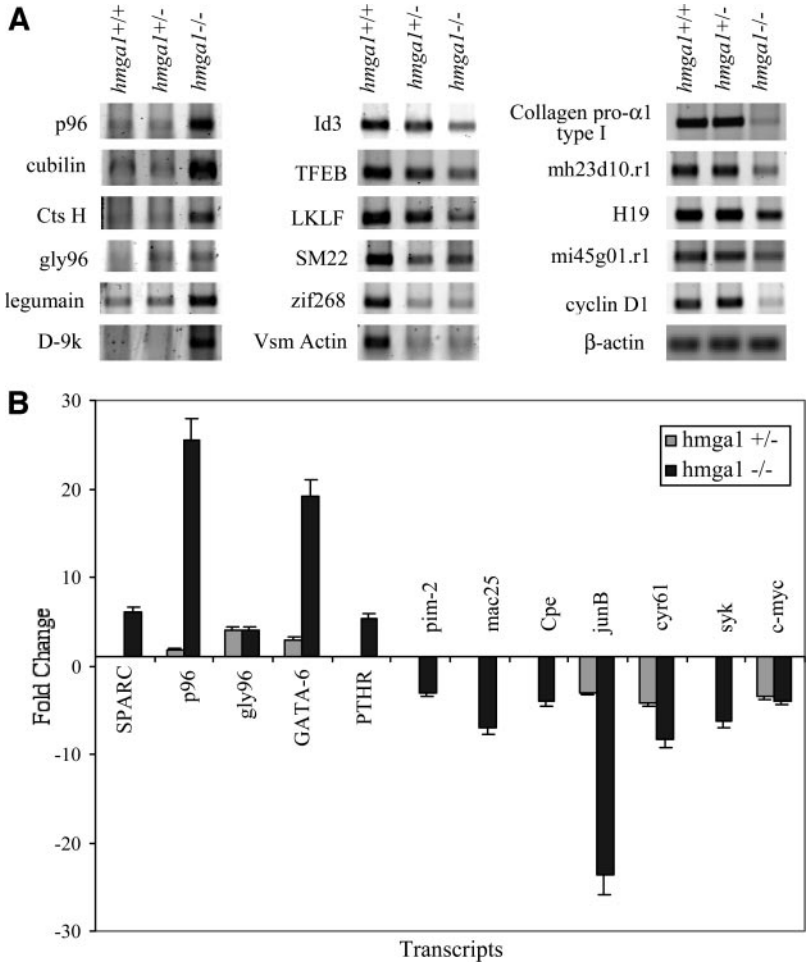

Fig. 2. Validation of microarray data by semiquantitative and quantitative RT-PCR. We confirmed some of the microarray data by semiquantitative (*A*) or quantitative (*B*) RT-PCR. Amplification of the *β*-actin gene has been evaluated as a control of the RNA amount used. *Cpe*, carboxypeptidase E.

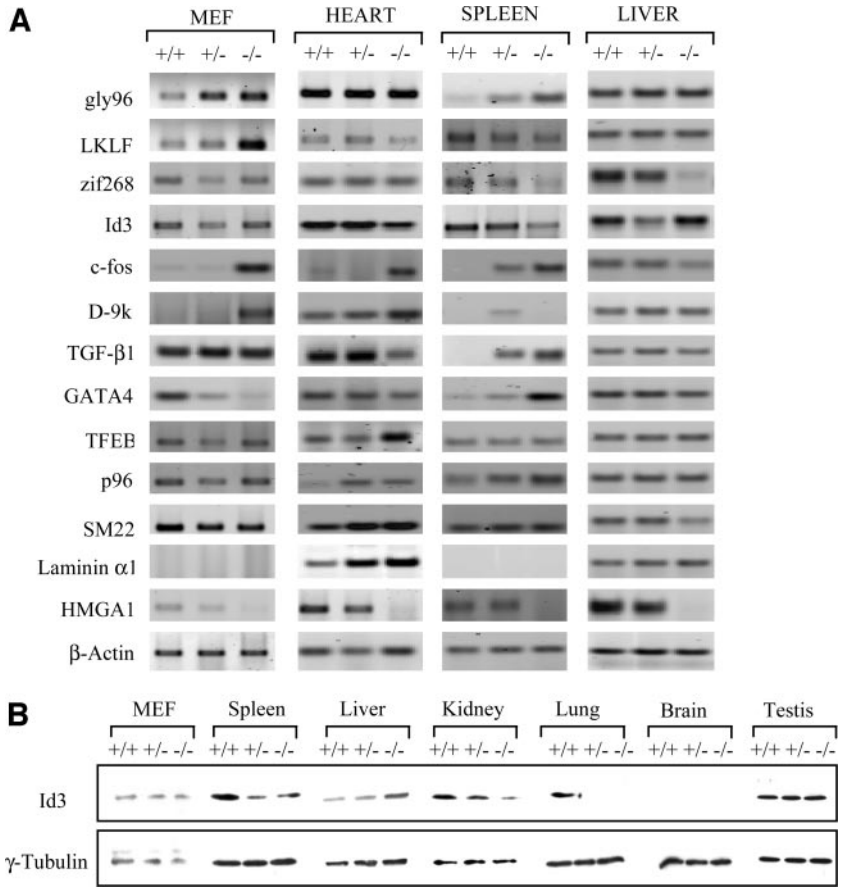

Fig. 3. Gene expression in *hmga1*-knockout cells and tissues. *A*, gene expression analysis by semiquantitative RT-PCR in *hmga1*-knockout MEFs and tissues. β-Actin has been used as a housekeeping gene to normalize the RNA amount used. *B*, protein expression analysis by Western blot in *hmga1*-knockout MEFs and tissues. As a control for equal protein loading, the blotted proteins were incubated with γ-tubulin-specific antibodies. The sources of proteins are indicated. *Cpe*, carboxypeptidase E.

and 23 had a >10-fold change. We examined the 250 transcripts (1.9%) that had a ≥4-fold-change in the homozygous mutant versus the wild-type sample. Among these 250 transcripts, 87 were increased and 163 were decreased, including 103 known genes (37 increased and 66 decreased), 118 expression sequence tags (ESTs) (40 increased and 78 decreased), and 29 unknown genes (10 increased and 19 decreased). As a control of microarray analysis, we verified that the HMGA1 was not expressed in hmga1−/− ES cells. The genes with ≥4-fold change in hmga1−/− ES cells were grouped according to their function: (*a*) signal-transduction pathways, (*b*) transcription factors, (*c*) cell proliferation, (*d*) extracellular-matrix and cellular-structure proteins, (*e*) metabolic pathways, transport and secretion, (*f*) growth factors and related proteins, (*g*) genes with immune functions, and (*h*) other genes. The relative fold changes in these genes, grouped as described above, are shown in Table 1. It is noteworthy that among the HMGA1-regulated genes, we found Id3 (13), lefty (14), and Wnt-6 (15) that are important in embryonic development and some oncogenes such as c-myc, junB, pim-2, and c-fos.

Validation of Microarray Analysis. To validate the results obtained by microarray analysis we evaluated the expression of 50 transcripts by semiquantitative RT-PCR in the wild-type, hmga1+/− and hmga1−/− ES cells. For all of them, RT-PCR analysis confirmed the differential expression associated with the expression of the HMGA1 proteins. Some representative RT-PCR analysis are shown in Fig. 2*A*. The expression of some genes (*i.e.*, TFEB, LKLF, and Id3) was hmga1 dose-dependent. In fact, the changes in hmga1+/− ES cells were intermediate between those found in wild-type and hmga1−/− ES cells. Conversely, the expression of other genes (*i.e.*, cubilin, p96, D-9K, legumain, and collagen) was not modified in hmga1+/− ES cells in comparison to the wild-type ES cells. We also analyzed 12 genes with quantitative RT-PCR. The results coincide with those of the microarray analysis (Fig. 2*B*). The primers used for semiquantitative and quantitative PCR are available upon request.

Some, but Not All, of the Genes Differentially Expressed in Wild-Type and *hmga1*-Knockout ES Cells Depend on HMGA1 Expression in Other Cells and Tissues. We next verified whether the genes differentially expressed in hm*ga1*-knockout ES cells showed a differential expression also in embryonic fibroblasts isolated from *hmga1*-knockout mice. A semiquantitative RT-PCR analysis showed that the differential expression of some genes (*i.e.*, D-9k, gly 96, and LKLF) in *hmga1*−/− fibroblasts matched that found in *hmga1*-knockout ES cells, whereas the expression of other genes did not (Fig. 3*A*). We next evaluated the expression of the HMGA1-regulated genes in adult heart, liver, and spleen tissue from *hmga1*+/− and *hmga1*−/− mice.

Some genes, such as Id3 and p96, showed the same expression trend as in ES cells, being down-regulated and up-regulated by HMGA1, respectively, also in heart and spleen (Fig. 3*A*). Some other genes only changed in one type of tissue. For example, gly96 and LKLF expression was different only in spleen, whereas TFEB and Laminin α1 were different only in heart. Interestingly, the regulation of some genes, such as TFEB and Laminin α1, in adult tissues was opposite to that found in ES cells. In fact, they were decreased in ES cells but increased in heart (Fig. 3*A*). These results suggest that HMGA1 function depends on the cellular context.

To verify that changes in RNA levels were associated with changes at protein levels, we analyzed by Western blot the expression of Id3 in MEF and tissues from wild-type and *hmga1*-knockout mice. As shown in the Fig. 3*B* for the Id3 protein, protein levels paralleled RNA levels and were characteristic in each tissue.

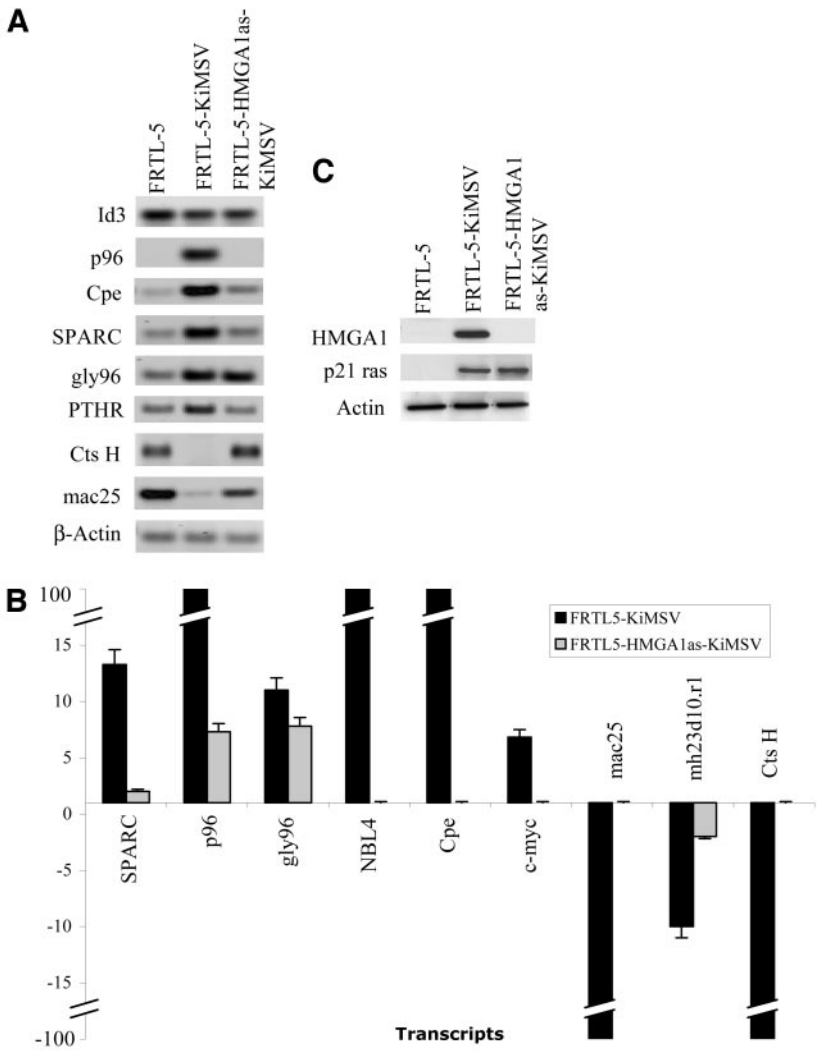

Fig. 4. Gene expression in a transformed cell system-overexpressing HMGA1. The expression of some genes has been studied by semiquantitative (*A*) or quantitative (*B*) RT-PCR in a transformed cell system overexpressing HMGA1; β-actin has been used as a housekeeping gene to normalize the RNA amount used; (*C*) Western blot analysis of v-*ras*-Ki and HMGA1 proteins. Sources of RNA or proteins are FRTL-5, rat thyroid epithelial cell line; FRTL-5-KiMSV, FRTL-5 infected with the Kirsten murine sarcoma virus; FRTL-5-HMGA1as-KiMSV, FRTL-5 transfected with a construct carrying HMGA1 mRNA in an antisense orientation, and then infected with the Kirsten murine sarcoma virus.

Analysis of the HMGA1-Dependent Genes in a Transformed Cell System. We demonstrated previously that HMGA1 overexpression is a necessary event in cell transformation. In fact, when HMGA1 expression was blocked by transfecting rat thyroid cells (FRTL-5) with an antisense *hmga1* cDNA construct and infected with the Kirsten murine sarcoma virus (KiMSV) carrying the v-*ras-Ki* oncogene, they (FRTL-5-HMGA1as-KiMSV) did not acquire the typical markers of neoplastic transformation (ability to grow in soft agar and induce tumors after injection into athymic mice), although the differentiation markers (*i.e.*, thyrotropin-dependency, ability to trap iodide, thyroglobulin synthesis, and secretion) were lost. Conversely, the neoplastic markers were shown by the untransfected rat thyroid cells infected with the same murine retrovirus (FRTL-5-KiMSV; ref. 8). Therefore, we analyzed, by semiquantitative and quantitative RT-PCR, the expression of some *hmga1*-dependent genes in FRTL-5, FRTL-5-KiMSV, and FRTL-5-HMGA1as-KiMSV cells. The experiments revealed two sets of genes. Some genes showed the same regulation observed in the ES cells, *i.e.*, carboxypeptidase E (Cpe) that decreased in *hmga1*-knockout ES cells and increased in the neoplastic cells or cathepsin H that increased in *hmga1*-knockout ES cells and decreased in the neoplastic cells compared with the wild-type controls. Other genes were regulated in an opposite direction compared with ES cells (*i.e.*, p96 and mac25), which demonstrated an increased and a decreased level, respectively, in both *hmga1*-knockout ES and the FRTL5 KiMSV cells compared with the respective controls, although the HMGA1 proteins were expressed only in the latter cells. Some representative results are shown in Fig. 4, *A* and *B*. In Fig. 4*C* we show the expression of the proteins v-*ras*-Ki and HMGA1 in the normal and neoplastic thyroid cells.

HMGA1 Proteins Bind to Id3 and p96 Promoters. We next evaluated whether the differential gene expression was a direct effect of the presence of HMGA1. We examined the Id3 and p96 genes because they were modified at RNA level in different cells and tissues and because their promoter regions contain AT-rich sequences that are a preferential-binding site for the HMGA proteins.

To investigate whether the HMGA1 proteins were able to bind the AT-rich promoter regions of both Id3 and p96, we performed an electrophoretic mobility shift assay using oligonucleotides spanning nucleotides −632 to −615 of the murine Id3 promoter region and −901 to −872 of the 5′ untranslated region of the murine p96 gene. As shown in Fig. 5*A*, a recombinant HMGA1 protein was able to bind directly to these regions. Binding specificity was demonstrated by competition experiments showing loss of binding with the addition of 200-fold molar excess of a specific unlabeled oligonucleotide. Subsequently, we performed binding assays with total extract from wild-type and *hmga1*-knockout murine spleens. Two specific complexes with mobility corresponding to the HMGA1 proteins (isoforms A1a and A1b) were present in extracts from wild-type and heterozygous (data not shown) spleens, whereas they were absent in extracts from homozygous *hmga1*-knockout mice (Fig. 5*B*). These complexes were

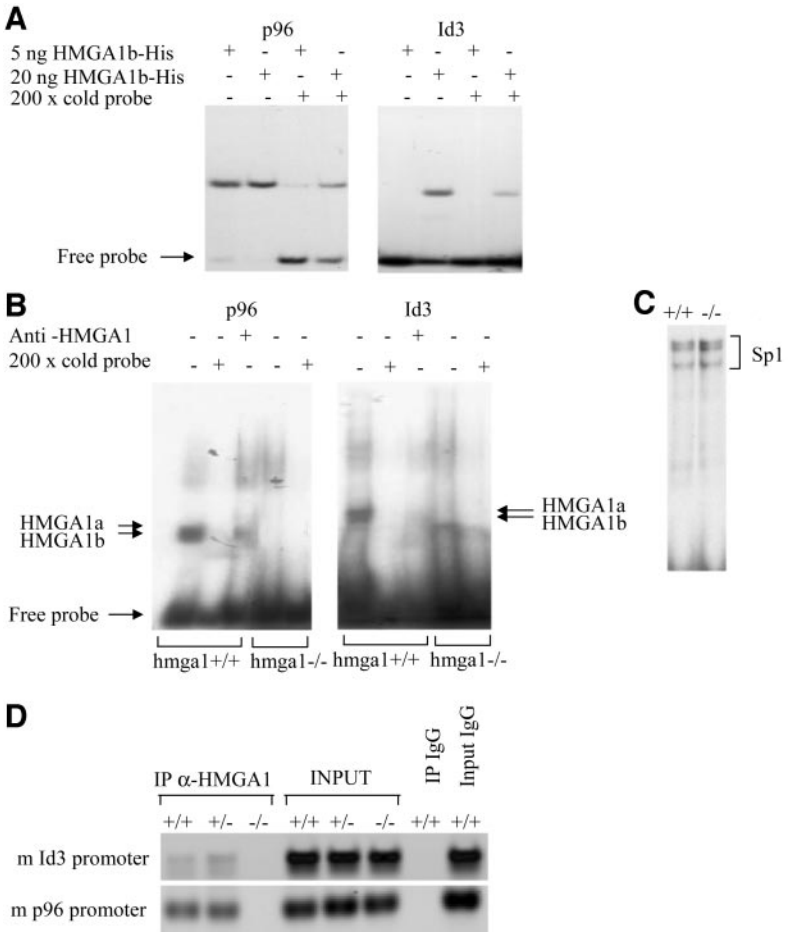

Fig. 5. HMGA1 proteins bind murine Id3 and p96 promoters *in vitro* and *in vivo*. *A*, electrophoretic mobility shift assay performed with radiolabeled oligonucleotides spanning from −209 to −169 bp of the murine Id3 promoter, spanning from −901 to −872 of the 5′ untranslated region of the murine p96 gene, and incubated with increasing amounts of recombinant HMGA1 as indicated. To assess the specificity of the binding, a 200× excess of unlabeled oligonucleotides were incubated as specific competitors. *B*, electrophoretic mobility shift assay performed with the same oligonucleotides as in *panel A*, incubated with total extracts from wild-type and hmga1-knockout murine spleens. In total extracts from wild-type (hmga1+/+) spleens there were two specific complexes, absent in total extracts from homozygous (hmga1−/−) spleen. Anti-HMGA1 antibody was used as a specific competitor. *C*, a control gel shift for Sp1 was performed to normalize the spleen extracts used. *D*, chromatin immune precipitation assay was performed on wild-type and hmga1-knockout MEFs. The recovered DNA was used as a template for PCR reactions with primers that amplify the murine Id3 and p96 promoters.

specifically displaced by the incubation with an antibody directed against the HMGA1 proteins, demonstrating that these complexes do consist of the HMGA1 proteins (Fig. 5*B*). A control gel shift for Sp1 was performed to normalize the spleen extracts used (*Panel C*).

To verify that HMGA1 proteins bind to Id3 and p96 promoters *in vivo*, we performed experiments of chromatin immunoprecipitation in MEF from wild-type and *hmga1*-knockout mice. Chromatin prepared as described under Materials and Methods was immunoprecipitated with anti-HMGA1 or normal rabbit IgG antibodies. The results shown in Fig. 5*D* demonstrate that HMGA1 proteins bind to these promoters. In fact, the Id3 and p96 promoter regions were amplified from the DNA recovered with anti-HMGA1 antibody in wild-type and *hmga1*+/− but not in *hmga1*−/− MEFs. Moreover, no amplification was observed in samples immunoprecipitated with normal rabbit IgG.

DISCUSSION

We have analyzed the expression profile of ES cells carrying two, one, and no *hmga1* functional allele by screening an "Affymetrix microarray" to identify genes that are regulated, positively or negatively, by the HMGA1 proteins. We found 87 transcripts increased and 163 decreased with a ≥4-fold change in *hmga1*−/− ES cells. The validity of these assays was confirmed by the absence of HMGA1 expression in the ES knockout cells. Semiquantitative and quantitative RT-PCR confirmed that all of these genes were differentially expressed in wild-type and *hmga1*-knockout ES cells. Several genes displayed *hmga1* dose-dependency, the phenotype of heterozygous cells was intermediate between those of wild-type and homozygous knockout cells. Thus for some genes, the level of *hmga1* expression may be critical for appropriate gene expression. In this case both alleles seem to be necessary to regulate the expression of these genes. For some other genes, the dependency on the *hmga1* expression levels was even more pronounced because the gene expression level in heterozygous ES cells was very close to that observed in homozygous cells. This type of regulation by *hmga1* expression levels may explain the appearance of pathologies, such as cardiac hypertrophy and B cell lymphomas, in mice heterozygous for *hmga1* gene disruption.[6] Several other genes showed the same expression level in wild-type and heterozygous ES cells. In this case, one *hmga1* allele is sufficient to regulate gene expression.

The genes regulated by HMGA1 in ES cells were also analyzed in MEF and in liver, spleen, and heart from wild-type, *hmga1*+/− and *hmga1*−/− mice. Different results were obtained in comparison to those observed in ES cells. In fact, the expression of some genes was either not modified by *hmga1* gene expression, or their regulation occurred in an opposite direction. It is noteworthy that the HMGA1 regulation of several genes was cell- and tissue-specific. It is known that by interacting with partner proteins, the HMGA1 proteins are able to enhance or suppress the effect of more "traditional" transcriptional activators and repressors. The fact that partner proteins are critical for HMGA1 activity may account for the cell- and tissue-specific regulation exerted by the HMGA1 proteins.

The same occurred when the HMGA1-regulated genes in ES were investigated in a cell system constituted by normal rat thyroid cells (FRTL-5) that do not express the HMGA1 proteins, the same cells

[6] M. Fedele, V. Fidanza, S. Battista, A. Fusco, manuscript in preparation.

malignantly transformed by the KiMSV (FRTL-5-KiMSV) that express high-HMGA1 levels and FRTL-5-KiMSV cells in which the synthesis of the HMGA1 protein was blocked by an antisense construct (FRTL-5-HMGA1as-KiMSV). These experiments revealed two sets of genes, those showing the same kind of regulation observed in ES and those genes showing regulation that occurred in an opposite direction.

The differential gene expression depending on the HMGA1 presence could depend on an indirect effect of the HMGA1 proteins in the sense that HMGA1 might induce some proteins that may interfere with the expression of some genes. To exclude this possibility and demonstrate a direct effect of HMGA1 on the regulation of some genes expressed differentially in *hmga1*-knockout cells, we performed electrophoretic mobility shift assay and chromatin immunoprecipitation experiments. We demonstrated the binding of the HMGA1 proteins to the promoters of Id3 and p96. We note of particular interest the finding that p96 and Id3 are regulated by the HMGA1 proteins because they are believed to have a critical role in the process of carcinogenesis. In fact, although no putative alterations on Id genes have been identified in primary human tumors to date to certify Ids as true cellular proto-oncogenes, Id proteins that are basic helix-loop-helix transcription factors have been implicated in regulating a variety of cellular processes (*i.e.*, cellular growth, senescence, differentiation, apoptosis, and angiogenesis) that regulate tumorigenesis (16). In particular, Id3 has been frequently found increased in human neoplasias (16). Equally, p96, a mitogen-responsive phosphoprotein cloned from a mouse macrophage cell expression library, is consistently down-regulated in mouse mammary carcinogenesis and in human ovarian carcinomas as compared with normal surface epithelium (17, 18). It is likely that Id3 up-regulation and p96 down-regulation in human neoplasias depend also on the HMGA1 overexpression, a feature of most of the human-malignant neoplasias (19).

When we analyzed the expression of some genes in MEF and adult tissues taken from HMGA2 knockout mice, we found that several genes do not appear to be regulated by HMGA2 (data not shown). This result could depend on the different action of these two members of the same HMGA protein family and confirms that although HMGA1 and HMGA2 have a similar structure and expression profile (high during embryogenesis and neoplastic tissue), they exert different functions. This is consistent with a body of evidence indicating that the two proteins exert different function: (*a*) the BRCA1 promoter is regulated negatively by HMGA1 but not by HMGA2 (20); (*b*) HMGA2 is critical for adipocytic cell growth (21, 22), whereas HMGA1 has negative effect on the growth of the preadipocytic cells 3T3 L1 (23); and (*c*) the phenotype of the *hmga1*- and *hmga2*-knockout mice is divergent: *i.e.*, a reduction in size and fat tissue of *hmga2*-null mice and in cardiac hypertrophy and B-cell lymphomas of *hmga1*-null mice.[7]

In conclusion, this study indicates that HMGA1 proteins are involved in the regulation of several genes. For some genes, such as Id3 and p96, we demonstrate that the regulation is direct. The positive or negative regulation appears to be tissue-specific because it likely depends on the multiprotein complex in which HMGA1 proteins are inserted.

[7] M. Fedele, V. Fidanza, S. Battista, A. Fusco, manuscript submitted for publication.

ACKNOWLEDGMENTS

We thank the Associazione Partenopea per le Ricerche Oncologiche (APRO) for its support and Jean Ann Gilder (Scientific Communication) for editing the text.

REFERENCES

1. Johnson KR, Lehn DA, Reeves R. Alternative processing of mRNAs encoding mammalian chromosomal high-mobility-group proteins HMG-I and HMG-Y. Mol Cell Biol 1989;9:2114–23.
2. Manfioletti G, Giancotti V, Bandiera A, et al. cDNA cloning of the HMGI-C phosphoprotein, a nuclear protein associated with neoplastic and undifferentiated phenotypes. Nucleic Acids Res 1991;19:6793–7.
3. Reeves R, Nissen MS. The AT DNA binding domain of mammalian high mobility group I chromosomal protein. A novel peptide motif for recognizing DNA structure. J Biol Chem 1990;265:8576–82.
4. Thanos D, Maniatis T. Virus induction of human IFN β gene expression requires the assembly of an enhanceosome. Cell 1995;83:1091–100.
5. Chiappetta G, Avvantaggiato V, Visconti R, et al. High level expression of HMGI(Y) gene during embryonic development. Oncogene 1996;13:2439–46.
6. Reeves R. Molecular biology of HMGA proteins: hubs of nuclear function. Gene 2001;277:63–81.
7. Scala S, Portella G, Fedele M, Chiappetta G, Fusco A. Adenovirus-mediated suppression of HMGI(Y) protein synthesis as potential therapy of human malignant neoplasias. Proc Natl Acad Sci USA 2000;97:4256–61.
8. Berlingieri MT, Pierantoni GM, Giancotti V, Santoro M, Fusco A. Thyroid cell transformation requires the expression of the HMGA1 proteins. Oncogene 2002;21:2971–80.
9. Battista S, Pentimalli F, Baldassarre G, et al. Loss of Hmga1 gene function affects embryonic stem cell lympho-hematopoietic differentiation. FASEB J 2003;17:1496–8.
10. Baldassarre G, Fedele M, Battista S, et al. Onset of natural killer cell lymphomas in transgenic mice carrying a truncated HMGI-C gene by the chronic stimulation of the IL-2 and IL-15 pathway. Proc Natl Acad Sci USA 2001;98:7970–5.
11. Yeh K, Lim RW. Genomic organization and promoter analysis of the murine Id3 gene. Gene 2000;254:163–71.
12. Sheng Z, Smith ER, et al. Chromosomal location of murine disabled-2 gene and structural comparison with its human ortholog. Gene 2001;268:31–49.
13. Yokota Y. Id and development. Oncogene 2001;20:8290–8.
14. Hamada H, Meno C, Watanabe D, Saijoh Y. Establishment of vertebrate left-right asymmetry. Nat Rev Genet 2002;3:103–13.
15. Cadigan KM, Nusse R. Wnt signaling: a common theme in animal development. Genes Dev 1997;11:3286–305.
16. Sikder HA, Devlin MK, Dunlap S, Ryu B, Alani RM. Id proteins in cell growth and tumorigenesis. Cancer Cell 2003;3:525–30.
17. Schwahn DJ, Medina D. p96, a MAPK-related protein, is consistently down-regulated during mouse mammary carcinogenesis. Oncogene 1998;17:1173–8.
18. Mok SC, Chan WY, Wong KK, et al. DOC-2, a candidate tumor suppressor gene in human epithelial ovarian cancer. Oncogene 1998;16:2381–7.
19. Fedele M, Battista S, Manfioletti G, Croce CM, Giancotti V, Fusco A. Role of the high mobility group A proteins in human lipomas. Carcinogenesis (Lond) 2001;22:1583–91.
20. Baldassarre G, Battista S, Belletti B, et al. Negative regulation of BRCA1 gene expression by HMGA1 proteins accounts for the reduced BRCA1 protein levels in sporadic breast carcinoma. Mol Cell Biol 2003;23:2225–38.
21. Zhou X, Benson KF, Ashar HR, Chada K. Mutation responsible for the mouse pygmy phenotype in the developmentally regulated factor HMGI-C. Nature (Lond) 1995;376:771–4.
22. Anand A, Chada K. In vivo modulation of HMGI-C reduces obesity. Nat Genet 2000;24:377–80.
23. Melillo RM, Pierantoni GM, Scala S, et al. Critical role of the HMGI(Y) proteins in adipocytic cell growth and differentiation. Mol Cell Biol 2001;21:2485–95.