

# The second AT-hook of the architectural transcription factor HMGA2 is determinant for nuclear localization and function

Giacomo Cattaruzzi<sup>1</sup>, Sandro Altamura<sup>2</sup>, Michela A. Tessari<sup>2</sup>,  
Alessandra Rustighi<sup>2,3</sup>, Vincenzo Giancotti<sup>2</sup>, Carlo Pucillo<sup>1</sup> and Guidalberto Manfioletti<sup>2,\*</sup>

<sup>1</sup>Dipartimento di Scienze e Tecnologie Biomediche, University of Udine, P.le Kolbe, 4 - 33100 Udine, Italy

<sup>2</sup>Dipartimento di Biochimica, Biofisica e Chimica delle Macromolecole, University of Trieste, via Giorgieri, 1 - 34127 Trieste, Italy and <sup>3</sup>Laboratorio Nazionale Consorzio Interuniversitario Biotecnologie, AREA Science Park, 34012 Trieste, Italy

Received September 27, 2006; Revised December 1, 2006; Accepted December 4, 2006

## ABSTRACT

**High Mobility Group A (HMGA) is a family of architectural nuclear factors which play an important role in neoplastic transformation. HMGA proteins are multifunctional factors that associate both with DNA and nuclear proteins that have been involved in several nuclear processes including transcription. HMGA localization is exclusively nuclear but, to date, the mechanism of nuclear import for these proteins remains unknown. Here, we report the identification and characterization of a nuclear localization signal (NLS) for HMGA2, a member of the HMGA family. The NLS overlaps with the second of the three AT-hooks, the DNA-binding domains characteristic for this group of proteins. The functionality of this NLS was demonstrated by its ability to target a heterologous  $\beta$ -galactosidase/green fluorescent protein fusion protein to the nucleus. Mutations to alanine of basic residues within the second AT-hook resulted in inhibition of HMGA2 nuclear localization and impairment of its function in activating the cyclin A promoter. In addition, HMGA2 was shown to directly interact with the nuclear import receptor importin- $\alpha$ 2 via the second AT-hook. HMGA proteins are overexpressed and rearranged in a variety of tumors; our findings can thus help elucidating their role in neoplastic transformation.**

## INTRODUCTION

The High Mobility Group A (HMGA1a, HMGA1b and HMGA2) are non-histone chromatin proteins which participate in a variety of cellular processes, including gene transcription, integration of retroviruses into chromatin, induction of neoplastic transformation and promotion of metastatic progression (1).

HMGA1a and HMGA1b, together referred to as HMGA1, result from alternative splicing of the same gene and differ for an internal region of 11 amino acids (2), while the highly homologous HMGA2 derives from a different but related gene (3).

The expression level of HMGA is maximal during embryonic development but is low in fully differentiated cells (1). Consistent with a role in proliferation and differentiation, homozygous mutations in the *Hmga2* gene result in the *pygmy* or 'mini-mouse' phenotype (4), while haploinsufficiency of the *Hmga1* gene causes cardiac hypertrophy and myelo-lymphoproliferative disorders in mice (5). In addition, lack of HMGA1 causes insulin resistance and diabetes in humans and mice (6).

In several human tumor cells, the HMGA1 protein level is high and correlates with the increasing degree of malignancy or metastatic potential, suggesting its use as a molecular marker for malignancy (1). HMGA2 expression is more restricted and was reported to localize to the invasive front of carcinomas, where cells exhibit epithelial–mesenchymal transition (7). Indeed, it has been recently reported that transforming growth factor- $\beta$  employs HMGA2 to elicit epithelial–mesenchymal transition (8). Of additional interest, rearrangements of the *HMGA2* gene resulting in loss of the acidic C-terminal tail, have been frequently detected in benign human tumors, mostly of

\*To whom correspondence should be addressed. Tel: +39 040 5583675; Fax: +39 040 5583694; Email: manfiol@units.it  
Present address:

Michela A. Tessari BioFocus DPI, Archimedesweg 4, P.O. Box 2048, 2301CA Leiden, The Netherlands

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

mesenchymal origin (9). Direct evidence for a role played by HMGA proteins in neoplastic transformation is given by overexpression of HMGA proteins, which was shown to promote a transformed phenotype in different cell lines, and by transgenic mice overexpressing HMGA which develop tumors (1,10,11). Recently, a novel role for HMGA in cellular senescence has been reported, adding more complexity to the functions played by this group of proteins (12).

HMGA proteins consist of about 100 amino acid residues and are characterized by the presence of three DNA-binding domains called 'AT-hooks', which preferentially bind to the narrow minor groove of AT-rich DNA sequences. Because of their intrinsic flexibility, HMGA proteins participate in specific protein-DNA and protein-protein interactions inducing both structural changes in chromatin and formation of stereospecific high-order complexes called 'enhanceosomes' on the promoter/enhancer DNA sequences, regulating expression of a large number of mammalian genes (1,13). One of the best-studied mechanisms of gene regulation in which HMGA1 has been involved in is that of the interferon- $\beta$  gene (IFN- $\beta$ ) (14). HMGA2 instead has been reported to associate with the transcriptional repressor p120<sup>E4F</sup> and to activate the cyclin A gene (15) and more recently to interact with pRB inducing E2F1 activity playing a crucial role in pituitary tumorigenesis (16).

Even though HMGA are considered important nodes in the chromatin protein network, the precise mechanisms of nuclear import for these proteins remain unknown. To determine the mechanism of nuclear import for HMGA proteins, we used a deletion mutagenesis approach on HMGA2 and identified a NLS corresponding to the second AT-hook of the protein. Expression of this sequence in fusion with a heterologous  $\beta$ -galactosidase ( $\beta$ -gal) and green fluorescent protein (GFP) efficiently directs nuclear targeting, thereby demonstrating functionality of this NLS. Point mutations of the basic residues within the NLS, in the context of full length protein, strongly inhibit HMGA2 nuclear localization and the ability to activate the cyclin A, a known HMGA2 target gene (15). Moreover, since the region identified is identical in all HMGA proteins, these results can be expected to be extended to the entire HMGA family. Our data provide the first reported mechanism of nuclear import for a HMGA protein and might help in finding strategies able to interfere with the nuclear localization and function of these proteins.

## MATERIALS AND METHODS

### Cells, tissue culture and drug treatments

NIH-3T3 (Mouse fibroblast), CHO (Chinese hamster ovary), HeLa (human cervical carcinoma) and 293 (human embryonic kidney) cell lines were grown in Dulbecco's modified Eagle's medium supplemented with fetal calf serum (10%), L-glutamine (4 mM), penicillin (100 units/ml) and streptomycin (100  $\mu$ g/ml) at 37°C in 5% CO<sub>2</sub> atmosphere. Ratjadone C (Alexis Corporation,

Lausen, Switzerland) treatment was performed for 6 h at a final concentration of 5 ng/ml. Ratjadone C is an analogue of the nuclear export inhibitor leptomycin B.

### Plasmid constructs

Plasmids expressing the wild-type and C-terminal deletion mutants of the human HMGA2 protein in fusion with  $\beta$ -gal and GFP, were generated via PCR using pcDNA3 HMGA2 (15) as template and the primers (carrying the restriction site, underlined, indicated in brackets) summarized in Table 1. Plasmids pHM829 HMGA2, pHM829 HMGA2 1-83, pHM829 HMGA2 1-73 and pHM829 HMGA2 1-54 were created using as upstream primer the A2\_109up oligonucleotide and, as downstream, the primers A2\_109dw, A2\_83dw, A2\_73dw and A2\_54dw respectively. Plasmid pHM829 HMGA2 1-43 was obtained with A2\_43up and A2\_43dw primers. PCR products were digested with the appropriate restriction enzymes and ligated into the pHM829 plasmid (17), a kind gift of Dr. T. Stamminger.

Plasmid pEGFP HMGA2 was generated by PCR with pcDNA3 HMGA2 as a template, and the primers A2EGFP\_up and A2EGFP\_dw (see Table 1). Primers were designed to incorporate a 5' *Hind* III site and a 3' *Bam* HI site (underlined) into the PCR product, for subsequent restriction and ligation into the pEGFP-N1 plasmid (Clontech Laboratories, Palo Alto, CA).

Plasmid pcDNA3HA TrxA2-45/53 was created by cloning into the *Rsr* II site of pcDNA3HATrx plasmid (15), carrying the *E. coli* thioredoxin ORF, the A2\_45/53s and A2\_45/53as oligonucleotides (see Table 1), which were phosphorylated and annealed; plasmid pcDNA3HA TrxA2-45/63 was created in the same way using the A2\_45/63s and A2\_45/63as oligonucleotides. Plasmids pHM829 TRX, pHM829 TRXA2-45/63 and pHM829 TRXA2-45/53 were created via PCR using as template the pcDNA3HATrx, pcDNA3HATrxA2-45/63 and pcDNA3HATrxA2-45/53 plasmids respectively, and the TRXup and TRXdw primers summarized in Table 1. Plasmid pGFP-HDAC4 was a kind gift from Dr. C. Brancolini, plasmid pGEX3X-Pendulin (importin- $\alpha$ 2) was supplied by Dr B. Birshstein, plasmid pGEX4T1 importin- $\beta$  was supplied by Dr I. Mattaj. Plasmid pGL2CycAwt was previously described (15).

### Site-directed mutagenesis

Plasmids pEGFP-HMGA2mS2 (K46A/R47A), pEGFP-HMGA2mS3 (K46A/R47A/R49A), pEGFP-HMGA2mS4 (K46A/R47A/R49A/R51A), pEGFP-HMGA2mS5 (K46A/R47A/R49A/R51A/K53A) and pEGFP-HMGA2mT6 (K74A/R75A/R77A/R79A/R81A/K82A) were obtained using the QuikChange site-directed mutagenesis kit (Stratagene, La Jolla, CA). The mutagenesis reaction was performed following the manufacturer's protocol using as template the pEGFP HMGA2 plasmid for pEGFP-HMGA2mS2, the pEGFP-HMGA2mS2 for pEGFP-HMGA2mS3, the pEGFP-HMGA2mS3 for pEGFP-HMGA2mS4, the pEGFP-HMGA2mS4 for pEGFP-HMGA2mS5.

**Table 1.** Name and sequence of oligonucleotides used in this work

Oligo name	Oligo sequence
A2EGFPup	5'-GAT TAC GCT <u>AAG CTT</u> ATG AGC GCA-3' ( <i>Hind</i> III)
A2EGFPdw	5'-GAC CGG TGG <u>GGA TCC</u> GTC CTC TTC G-3' ( <i>Bam</i> HI)
A2_109up	5'-GAT TAC GCT <u>GCT AGC</u> ATG AGC GCA-3' ( <i>Nhe</i> I)
A2_109dw	5'-GAC CGG TGG <u>TCT AGA</u> GTC CTC TTC G-3' ( <i>Xba</i> I)
A2_83dw	5'-AAC AAC TTG <u>TCT AGA</u> CCA TTT CCT AG-3' ( <i>Xba</i> I)
A2_73dw	5'-GCC TCT TGG <u>TCT AGA</u> TTC TCC AGT GG-3' ( <i>Xba</i> I)
A2_54dw	5'-GGA CTC TTG TTT <u>CTA GAG</u> CCT TTG GGT C-3' ( <i>Xba</i> I)
A2_43up	5'-CTC GAG CTC <u>CCG CGG</u> ATG AGC GCA-3' ( <i>Sac</i> II)
A2_43dw	5'-GAC CGG TGG <u>TCT AGA</u> GGG CTC ACC-3' ( <i>Xba</i> I)
A2_45/53s	5'-GTC CGC CTA AGA GAC CCA GGG GAA GAC CCA AAG CCG-3'
A2_45/53as	5'-GAC CGG CTT TGG GTC TTC CCC TGG GTC TCT TAG GCG-3'
A2_45/63s	5'-GTC CGC CTA AGA GAC CCA GGG GAA GAC CCA AAG GCA GCA AAA ACA AGA GTC CCT CTA AAG CAG CCG-3'
A2_45/63as	5'-GAC CGG CTG CTT TAG AGG GAC TCT TGT TTT TGC TGC CTT TGG GTC TTC CCC TGG GTC TCT TAG GCG -3'
TRXup	5'-TCA ATG <u>CTA GCG</u> CCA CCA TGG GCG ATA- 3' ( <i>Nhe</i> I)
TRXdw	5'-TGA <u>TTC CGC GGC</u> TTT CCA AGT CGG TTC AT-3' ( <i>Sac</i> II)
A2_mS2up	5'-GGT GAG CCC TCT CCT GCC GCA CCC AGG GGA AGA CCC-3'
A2_mS2dw	5'-GGG TCT TCC CCT GGG TGC GGC AGG AGA GGG CTC ACC-3'
A2_mS3up	5'-CCT GCC GCA CCC GCG GGA AGA CCC AAA GG-3'
A2_mS3dw	5'-CCT TTG GGT CTT CCC GCG GGT GCG GCA GG-3'
A2_mS4up	5'-CGC ACC CGC GGG AGC ACC CAA AGG CAG C-3'
A2_mS4dw	5'-GCT GCC TTT GGG TGC TCC CGC GGG TGC G-3'
A2_mS5up	5'-CGC GGG AGC ACC CGC AGG CAG CAA AAA CAA-3'
A2_mS5dw	5'-TTG TTT TTG CTG CCT GCG GGT GCT CCC GCG-3'
A2_mT2up	5'-GCA GAA GCC ACT GGA GAA GCC GCG CCA AGA GCC AGA CCT AGG-3'
A2_mT2dw	5'-CCT AGG TCT GCC TCT TGG CGC GGC TTC TCC AGT GGC TTC TGC-3'
A2_mT3up	5'-AGA AGC CGC GCC AGC AGG CAG ACC TAG GA-3'
A2_mT3dw	5'-TCC TAG GTC TGC CTG CTG GCG CGG CTT CT-3'
A2_mT5up	5'-CGC CAG CAG GCA GAC CTG CCG CAT GGC CAC AAC AAG TTG TTC AG-3'
A2_mT5dw	5'-CTG AAC AAC TTG TTG TGG CCA TGC GGC AGG TCT GCC TGC TGG CG-3'
A2_mT6up	5'-CGC CAG CAG GCG CAC CTG CCG CAT GG-3'
A2_mT6dw	5'-CCA TAG GGC AGG TGC GCC TGC TGG CG-3'

Primers used in this work, as indicated in material and methods. The underlined sequences indicate the restriction site of the enzyme reported in brackets.

Plasmid pEGFP-HMGA2mT6 comes from subsequent site-directed mutagenesis steps of lysine and arginine residues to alanine: pEGFP-HMGA2mT2 (K74A/R75A), pEGFP-HMGA2mT3 (K74A/R75A/R77A) and

pEGFP-HMGA2mT5 (K74A/R75A/R77A/R81A/K82A). Mutagenesis products were obtained as follows: HMGA2 mS2 using primer A2\_mS2up and primer A2\_mS2dw; HMGA2 mS3 using A2\_mS3up and A2\_mS3dw; HMGA2 mS4 using A2\_mS4up and A2\_mS4dw; HMGA2 mS5 using A2\_mS5up and A2\_mS5dw; HMGA2 mT2 using A2\_mT2up and A2\_mT2dw; HMGA2 mT3 using A2\_mT3up and A2\_mT3dw; HMGA2 mT5 using A2\_mT5up and A2\_mT5dw; HMGA2 mT6 using A2\_mT6up and A2\_mT6dw, respectively (see Table 1). Nucleotide sequence of each construct was verified by automated sequencing.

### Transfections and luciferase assays

Transfections by the standard calcium phosphate precipitation method and reporter gene assays were performed essentially as previously described (15). Cells were plated at a density of  $0.5 \times 10^6$  cells per 60-mm diameter culture dish and processed 24 h after removal of the precipitates. For luciferase assay, cells were transfected with 1  $\mu$ g of the CycA reporter construct (pGL2CycAwt) and 3  $\mu$ g of the pEGFP HMGA2 wt, pEGFP-HMGA2mS5 and pEGFP-HMGA2mT6 vectors respectively. 0.1  $\mu$ g of pRL-CMV Renilla luciferase expression vector (Promega Corporation, Madison, WI) was included to normalize for transfection efficiencies. The assay was performed using the dual-luciferase reporter assay system (Promega) according to manufacturer's instructions. Luciferase activity was assayed on a Lumat LB 9501 Luminometer (Berthold, Bad Wildbad, Germany). For cell imaging, cells were transfected with 4  $\mu$ g of the indicated plasmid.

### Confocal microscopy and imaging

For cell imaging, cells grown onto coverslips were washed twice in PBS, fixed in ice cold 3% paraformaldehyde in PBS at room temperature for 20 min, treated with 100 mM glycine in PBS for 5 min at room temperature, permeabilized with 0.5% Triton X-100 in PBS for 5 min at room temperature, then washed twice in PBS. Nuclei were stained with propidium iodide at a final concentration of 5  $\mu$ g/ml in PBS containing 400  $\mu$ g/ml RNase A (Sigma-Aldrich, St. Louis, MO), for 20 min at 37°C. Coverslips were finally washed extensively with PBS and mounted onto glass slides with Mowiol (pH=8.5) solution. Cells were imaged with a Leica TCS SP microscope with a 488–543 nm Argon laser. Fluorescence relative intensity was quantitated by analyzing acquired TIFF-format images with the ImageJ software (version 1.36); at least one hundred cells from each different experiment were subjected to quantitation analysis.

### Western blot analysis

Sixteen hours after removal of the precipitates, transfected cells were washed twice with ice cold PBS, harvested from the plates and transferred to Eppendorf tubes, then lysed in 62 mM Tris (pH=6.8), 2% SDS, 5% beta-mercaptoethanol, 10% glycerol, 1 mM phenylmethylsulfonyl fluoride, and protease inhibitors (Sigma)



lysis buffer. The lysates were then clarified by centrifugation and resolved by sodium dodecyl sulfate polyacrylamide-gel electrophoresis (SDS-PAGE). Proteins were transferred from the gel to a nitrocellulose membrane using the Multiphor II semi-dry transfer apparatus (Amersham Biosciences, Uppsala, Sweden). Membranes were blocked at room temperature with 5% non-fat milk powder in PBS (pH=7.4) and 0.1% Tween-20 before incubating overnight at 4°C with polyclonal rabbit anti-GFP primary antibody (a kind gift from Dr. C. Brancolini) or polyclonal rabbit anti HMGA2 primary antibody (15). Membranes were then washed extensively with PBS and 0.1% Tween-20, and incubated for 2 h at room temperature with horseradish peroxidase-conjugated goat anti-rabbit secondary antibody (Sigma). Finally, membranes were washed extensively as described above, and processed using the SuperSignal West Femto chemiluminescence substrate (Pierce Biotechnology, Rockford, IL) following the manufacturer's protocol.

### *In vitro* translation (IVT) and GST pull-down assay

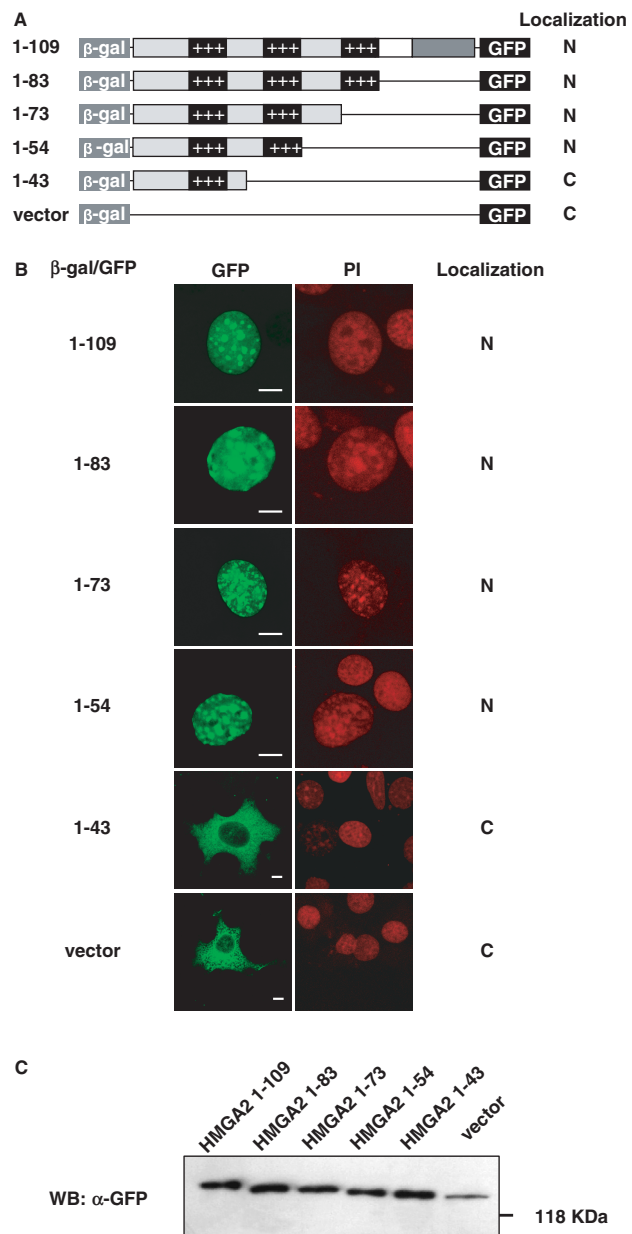
IVT of wild-type and deletion mutants of the human HMGA2 protein were performed using the TnT-coupled Transcription and Translation kit (Promega) with [<sup>35</sup>S]methionine (NEN Life Science, Boston, MA), according to manufacturer's instructions.

Expression and purification of recombinant GST fusion proteins were carried out with standard protocols. GST-pull-down assays were performed, essentially as previously described (15), by incubating 5 µg of recombinant GST importin-α2, GST importin-β and GST proteins bound to beads with 10 µl of the *in vitro* translated proteins in 200 µl of binding buffer (25 mM HEPES pH 7.9, 50 mM NaCl, 1 mM DTT, 0.01% NP-40). After 2 h of incubation at room temperature, the protein-bound beads were washed twice in complete binding buffer (binding buffer plus 0.25% BSA) and twice in binding buffer without BSA. The bound proteins were then eluted and analyzed by SDS-PAGE (T=8%) and autoradiography.

## RESULTS AND DISCUSSION

### Basic residues within the second AT-hook are necessary for nuclear localization

To identify regions of HMGA2 that could serve as a nuclear localization signal (NLS), we performed deletion mutagenesis analysis using a previously described *in vivo* assay designed specifically for the identification of NLSs (17). We generated four deletion mutants of the human HMGA2 removing progressively sequences from the C-terminal tail. These HMGA2 fragments result simultaneously fused at the N-terminal to β-galactosidase (β-gal) and at the C-terminal to green fluorescent protein (GFP) (Figure 1A). This system is particularly advantageous since it adds bulk (β-gal contributes with 86 kDa) preventing passive diffusion through the nuclear pore, since small GFP fusion proteins could be subjected to passive diffusion through nuclear pores and to proteolytic degradation. The different constructs were



**Figure 1.** Identification of HMGA2 nuclear localization signal by deletion mutagenesis. (A) Schematic representation of HMGA2 deletion mutants; each deletion mutant is fused to β-gal at the N-terminus, and to GFP at the C-terminus. Summary of intracellular localization is indicated at the right. (B) NIH-3T3 cells were transfected with HMGA2 deletion mutants constructs, and fusion proteins were visualized by confocal laser microscopy. Propidium iodide (PI) staining of the same nuclei is shown. At least one hundred cells per transfection were analyzed in three different experiments. Bars, 10 µm. (C) Western blot analysis performed with an anti-GFP antibody.

transiently transfected in NIH-3T3 cells and subcellular localization was monitored by laser confocal microscopy. As depicted in Figure 1B, full-length HMGA2 fusion protein (HMGA2 1-109) accumulates exclusively into the nucleus and is mostly localized into the large heterochromatic regions which can be positively stained with 4'-6-Diamidino-2-phenylindole (DAPI) (data not

shown). This intranuclear distribution appears very similar to that described for HMGA proteins and to that obtained by staining the endogenous HMGA2 protein with anti-HMGA2 specific antibodies (18,19 and data not shown).

Deletion of regions from the C-terminus tail that retain the second AT-hook (HMGA2 1-83, HMGA2 1-73, HMGA2 1-54) show an exclusive nuclear localization, as the full-length HMGA2 1-109 does, while the removal of the second AT-hook (HMGA2 1-43) results in a cytoplasmic localization of the fusion protein, as for the empty  $\beta$ -gal/GFP vector (Figure 1B). As shown in Figure 1C, all the fusion proteins show the expected molecular weight. Experiments were performed with NIH-3T3 cells but similar results were obtained with HeLa and 293 cells (data not shown).

With the deletion mutagenesis analysis, we restricted the NLS to an 11 amino acid region corresponding to the second AT-hook. HMGA2 has three AT-hooks which are highly conserved among HMGA proteins and which allow them to interact with the minor groove of AT-rich DNA sequences. Interestingly, even though their sequence is very similar, the first and the third AT-hook do not seem to be involved in nuclear localization since the removal of the third (construct HMGA2 1-73) does not prevent nuclear localization while the presence of the first (construct HMGA2 1-43) is not sufficient (Figure 1B).

NLSs are often characterized by clusters of basic amino acid residues, in particular the classical monopartite NLS is composed of a single cluster of basic residues (20). Within the second AT-hook there are five basic residues that could potentially serve as a monopartite NLS. To better define the NLS, we generated EGFP fusion constructs (HMGA2 mS2-EGFP, HMGA2 mS3-EGFP, HMGA2 mS4-EGFP and HMGA2 mS5-EGFP) in which we progressively mutated these basic residues to alanine and tested them for cellular localization (Figure 2). A construct with all the six basic residues of the third AT-hook converted to alanine (HMGA2 mT6-EGFP) was also included as control. As shown in Figure 2B, the mutation of the first two basic residues (K46, R47) to alanines (HMGA2 mS2-EGFP) disrupts normal sub-cellular localization, since a cytoplasmic fluorescence can be detected. To better quantify the influence of the mutated residues in nuclear localization, the ratio between nuclear and cytoplasmic fluorescence was quantified with a digital software analysis system. Figure 2B and C show that progressive mutation of basic residues within the second AT-hook leads to an increase in cytoplasmic fluorescence which is, with HMGA2 mS5-EGFP, comparable to that of the empty vector. Differently, the construct with all six basic residues of the third AT-hook mutated to alanines (HMGA2 mT6-EGFP) displays the same nuclear localization of wild type HMGA2. As shown in Figure 2D, mutations within the basic residues do not affect the integrity of the fusion proteins. To further confirm these results, the HMGA2 sequences of HMGA2 mS5 and HMGA2 mT6 were cloned in fusion with  $\beta$ -gal and GFP, as described above. Results did confirm complete nuclear localization for HMGA2

mT6 and cytoplasmic localization for HMGA2 mS5 (data not shown).

### **The second AT-hook (aminoacid residues 45–53) is necessary and sufficient for nuclear localization**

To directly test the second AT-hook for NLS functionality, we generated expression constructs (TrxA2 45-63 and TrxA2 45-53) in which the regions from amino acid 45 to 63 and 45 to 53 were inserted within the active site of the *E. coli* thioredoxin (Trx), flanked at the N-terminus with  $\beta$ -gal and at the C-terminus with GFP (Figure 3A). Trx is largely used as scaffold since it provides conformational constraint and stability to short peptide sequences (15). These constructs were transiently transfected in NIH-3T3 cells and both showed an exclusive nuclear localization in contrast with the empty vector containing Trx which localized into the cytoplasm (Figure 3B). The integrity of the expressed proteins was controlled by Western blot, as shown in Figure 3C.

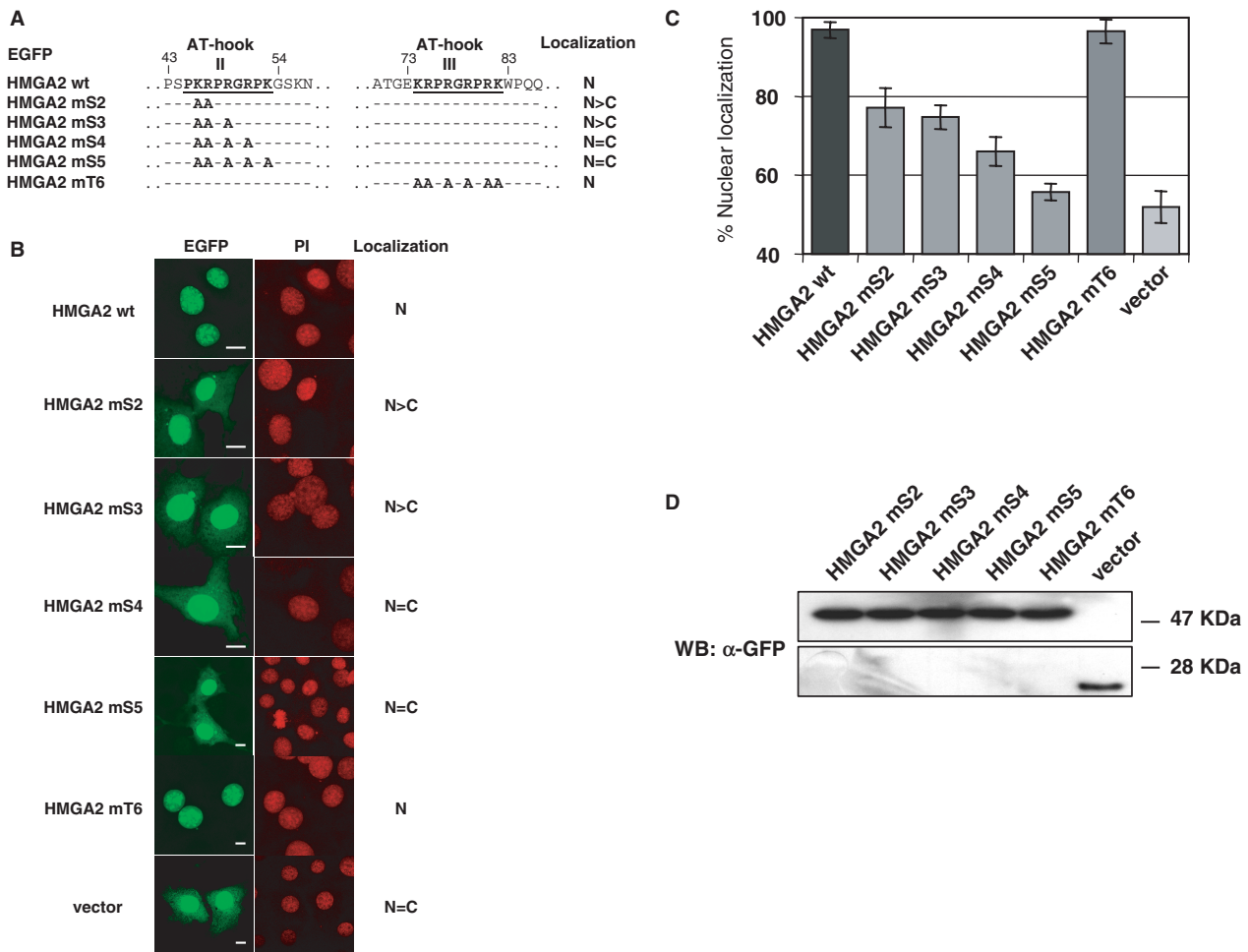
These experiments do demonstrate that the second AT-hook, in particular the region from amino acid 45 to 53, is necessary and sufficient to drive nuclear localization. This region is highly conserved in HMGA2 proteins of different species and it is also conserved in the other HMGA proteins – HMGA1a and HMGA1b – of different species (Figure 3D). Interestingly, our findings imply also a nuclear localization for HMGA1c, a less abundant and poorly characterized splice variant of HMGA1 whose sequence retains the first two AT-hooks and whose C-terminus differs completely from the other two isoforms HMGA1a and HMGA1b (21). We therefore believe that our results can be applied to the entire HMGA family and to different species of vertebrates.

In addition, despite the presence of several other basic residues within the HMGA2 sequence, our data indicate a basic monopartite NLS type for HMGA proteins.

We next tested whether cytoplasmic accumulation of HMGA2 mutants was due to removal of a NLS and was not due to decryption of a nuclear export signal (NES) which could be exposed as a consequence of the mutations. Cells transfected with constructs HMGA2 1-43 and HMGA2 mS5-EGFP were treated with the nuclear export pathway inhibitor ratjadone C. This compound, an analogue of leptomycin B, binds and inhibits the function of Crm1 protein, a critical mediator of nuclear export pathways (22). Results indicate that ratjadone C treatment had no effect on the cytoplasmic accumulation of both mutants (Figure 4). This result suggests that cytoplasmic accumulation of the mutant proteins results from NLS disruption and not from exposure of a cryptic NES. As a positive control for ratjadone C activity, we used a GFP-HDAC4 construct, which has been shown to contain an active NES (23). As expected, GFP-HDAC4 cytoplasmic localization is inhibited following drug treatment, resulting in nuclear accumulation (Figure 4).

### **Importin- $\alpha$ mediates HMGA2 nuclear transport**

The finding of a basic monopartite NLS suggested that HMGA2 might bind the conventional nuclear import

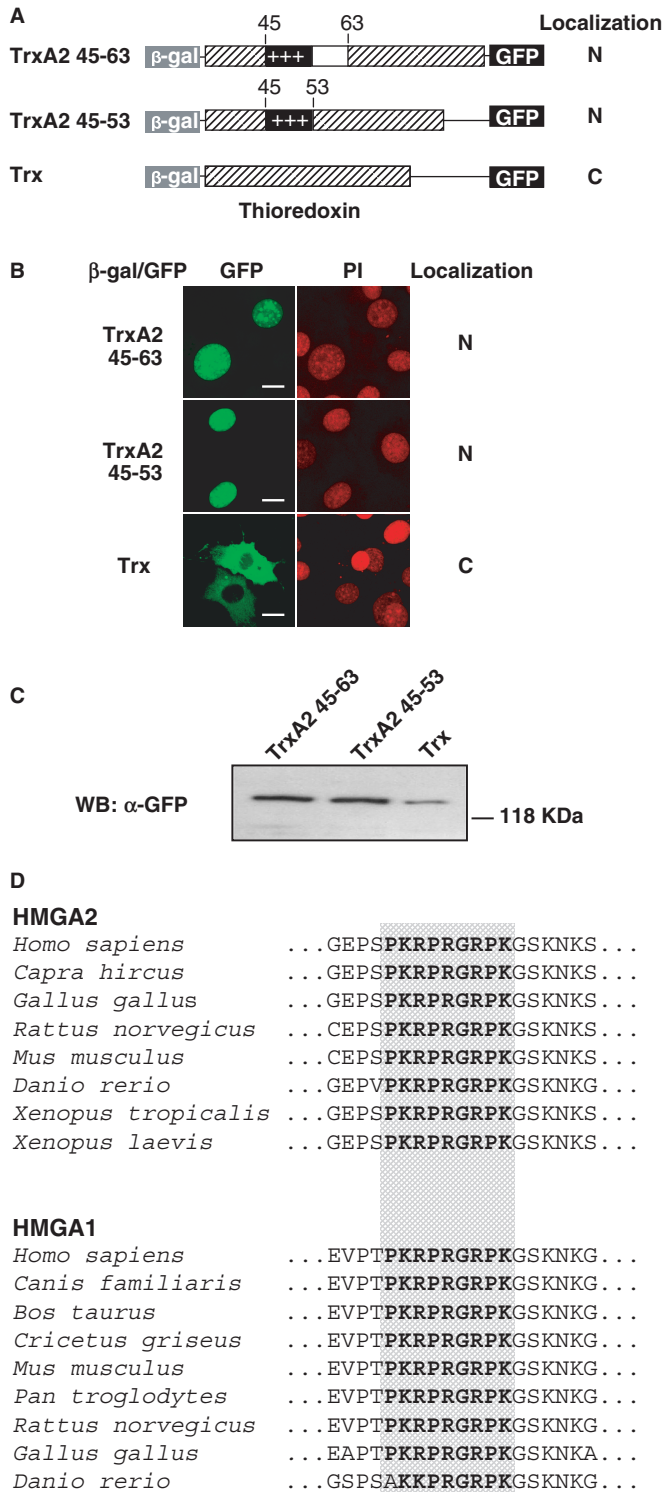


**Figure 2.** Basic residues of the second AT-hook of HMGA2 are responsible for nuclear localization of fusion proteins. **(A)** Partial sequences of point-mutated constructs of HMGA2. Basic residues were replaced with alanine residues as indicated. Each point-mutant is fused C-terminally with EGFP (not drawn); numbers identify positions of residues on HMGA2 protein sequence. Summary of intracellular localization is indicated at right. **(B)** NIH-3T3 cells were transfected with HMGA2 point-mutant constructs, and fusion proteins were visualized by confocal laser microscopy. Propidium iodide (PI) staining of the same nuclei is shown. At least one hundred cells per transfection were analyzed in three different experiments, and at least one hundred cells per transfection were subjected to fluorescence quantification. Bars, 10  $\mu$ m. **(C)** For each point-mutant construct, nuclear and cytoplasmic fluorescence were quantified, and percentage of nuclear localization was calculated. Bars indicate percentage  $\pm$  standard deviation. **(D)** The expression and integrity of the fusion proteins were analyzed by Western blot using anti-GFP antibody.

receptors (such as importin- $\alpha$  protein family) to achieve nuclear translocation. This pathway initiates with the binding of importin- $\alpha$  to the NLS of a substrate protein. Importin- $\alpha$  then binds importin- $\beta$  to the NLS and this heterotrimeric complex enters the nuclear pore complex and translocates into the nucleus. There are, however, many exceptions to this kind of import, derived from the presence of different kinds of cytoplasm receptors that can recognize the NLS, or from the ability of the NLS to bind directly to importin- $\beta$ , independently upon importin- $\alpha$ , or from the presence of proteins that can translocate through the nuclear pore complex in the absence of any soluble factors (20). To investigate whether HMGA2 directly interacts with importin- $\alpha$ , we performed GST pull-down experiments. Figure 5A shows that *in vitro* translated HA-tagged HMGA2 efficiently binds GST importin- $\alpha$ 2 but not GST importin- $\beta$  or GST alone. BSAP/Pax-5 transcription factor was used as a positive

control since it was previously reported to bind importin- $\alpha$ 2 (24). To map the region in HMGA2 which is recognized by importin- $\alpha$ 2, *in vitro* translated HMGA2 deletion constructs (HMGA2 1-83, HMGA2 1-73, HMGA2 1-54 and HMGA2 1-43) produced in fusion with  $\beta$ -gal and GFP were employed. As it can be seen in Figure 5B, deletions from the C terminal of HMGA2 did not affect binding to importin- $\alpha$ 2 as long as the second AT-hook is maintained (HMGA2 1-83, HMGA2 1-73, HMGA2 1-54). A further deletion removing the second AT-hook (HMGA2 1-43) abolishes almost completely importin- $\alpha$ 2 binding. To demonstrate that the region comprising the second AT-hook is necessary and sufficient for importin- $\alpha$ 2 binding, a GST pull-down experiment was performed using constructs TrxA2 45-63 and TrxA2 45-53. Results indicated that even the shortest construct consisting of nine residues of the second AT-hook is able to bind efficiently importin- $\alpha$ 2 (Figure 5C).





**Figure 3.** The second AT-hook of HMGA2 constitutes a nuclear localization signal. (A) Schematic representation of fusion proteins which include second AT-hook; each fusion protein is formed by the  $\beta$ -gal at the N-terminus, thioredoxin in which peptides coding for the second AT-hook (aminoacids 45–63 or 45–53) are expressed within the catalytic domain, and GFP at the C-terminus. Summary of intracellular localization is indicated on the right. (B) NIH-3T3 cells were transfected with HMGA2 peptide containing constructs, and fusion proteins were visualized by confocal laser microscopy. Fusion protein including the peptide sequence PKRPRGRPK (aminoacids

Taken together, therefore, these results demonstrate that sequence from aa 45 to 53 is necessary and sufficient to drive nuclear localization of HMGA2 and for binding to importin- $\alpha$ 2.

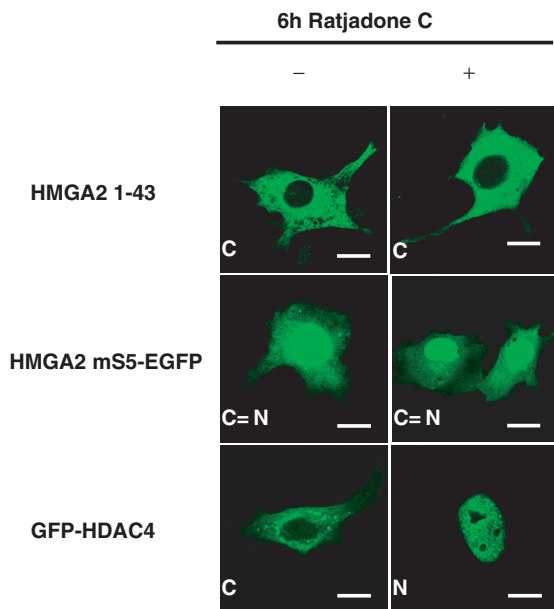
**The second AT-hook is critical for HMGA2 function**

To demonstrate that the basic residues within the second AT-hook are not only the critical determinants for HMGA2 nuclear localization but also are responsible for HMGA2 function, we tested HMGA2 mS5-EGFP in transient transfection experiments, and HMGA2 mT6-EGFP as a control, along with a cyclin A reporter for luciferase induction (Figure 6A). We have in fact previously demonstrated that HMGA2 is able to induce the cyclin A gene through the cyclic AMP-responsive element (CRE). This effect is mediated by HMGA2 association with transcription factors binding to CRE and, since the region involved included the second AT-hook, possibly also by HMGA2-DNA interaction (15).

Consistent with our previous report, wild type HMGA2 in fusion with EGFP is able to efficiently induce the reporter gene (Figure 6B). Mutations within the second AT-hook (construct HMGA2 mS5) completely abolish HMGA2-mediated induction, while mutations within the third AT-hook (construct HMGA2 mT6) do not alter the activity of HMGA2 on the reporter gene.

Nuclear localization signals of many transcription factors are localized onto DNA binding domains and electrostatic interactions between positively charged basic residues and the negatively charged phosphate DNA backbone are the predominant mediators of nucleic acid binding (25). Indeed NMR studies have demonstrated the relevant function played by basic residues within the second AT-hook of HMGA1 in establishing contacts with DNA bases (26). Therefore our work demonstrates that basic residues within the second AT-hook are critical determinants for HMGA2 nuclear localization and for HMGA2 function. Moreover, since the second AT-hook of HMGA proteins is the main region involved in protein–protein interactions with most of HMGA partners (16,27,28), our findings further stress the relevance of this domain in HMGA family.

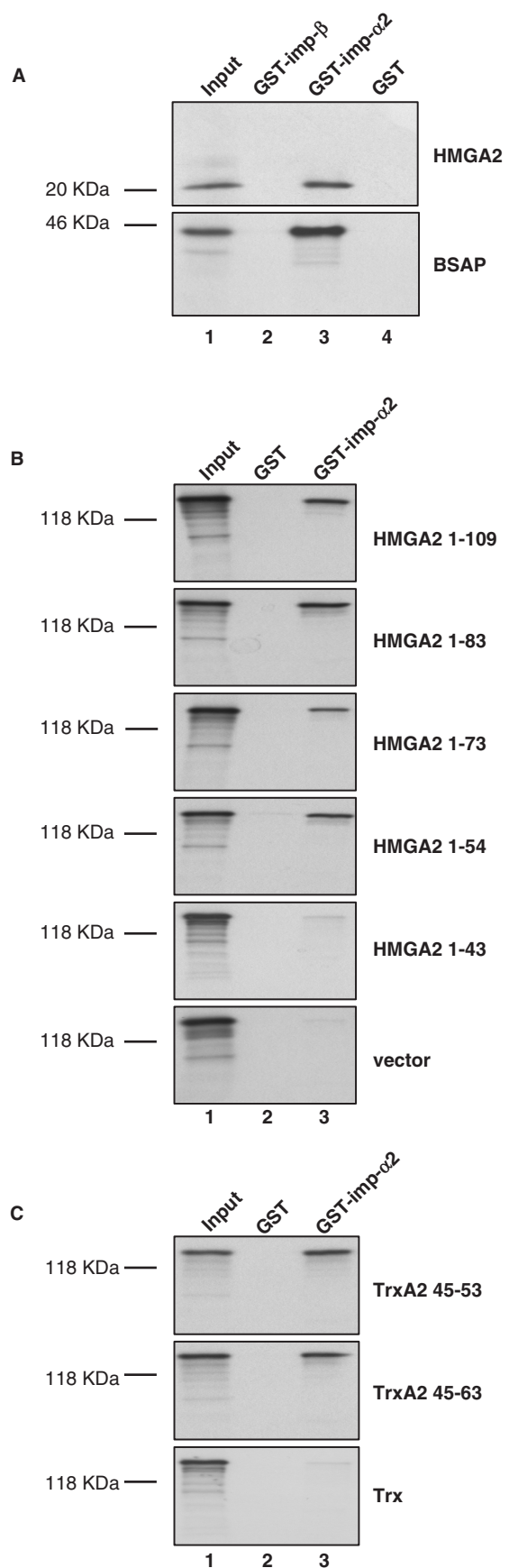
from 45 to 53 of HMGA2) localizes into the nucleus (more than 95% of transfected cells). At least one hundred cells per transfection were analyzed in three different experiments. Bars, 10  $\mu$ m. (C) The expression and integrity of the fusion proteins were confirmed by Western blot using anti-GFP antibody. (D) Alignment of aminoacid sequences of the second AT-hook of HMGA2 and HMGA1 proteins, and flanking residues. Sequences show very high conservation, and residues of the second AT-hook are identical (bolded and dot boxed) in all of the HMGA2 proteins from *H. sapiens* to *D. rerio*, as well as in the HMGA1 proteins from *H. sapiens* to *G. gallus* (while *D. rerio* sequence has one difference in a basic residue [R>K]). Sequences were obtained from public databases (HMGA2: *H. sapiens* NP\_003474; *C. hircus* BAB64336; *G. gallus* NP\_990332; *M. musculus* NP\_835158; *R. norvegicus* NP\_114459; *X. tropicalis* AL955898; *X. laevis* AAH82363; *D. rerio* NP\_997845; HMGA1: *H. sapiens* NP\_665911; *C. familiaris* NP\_001003387; *B. Taurus* XP\_888488; *C. griseus* AAF06666; *M. musculus* NP\_057869; *P. troglodytes* XP\_530918; *R. norvegicus* NP\_647543; *G. gallus* AAQ63840; *D. rerio* AAH44554).



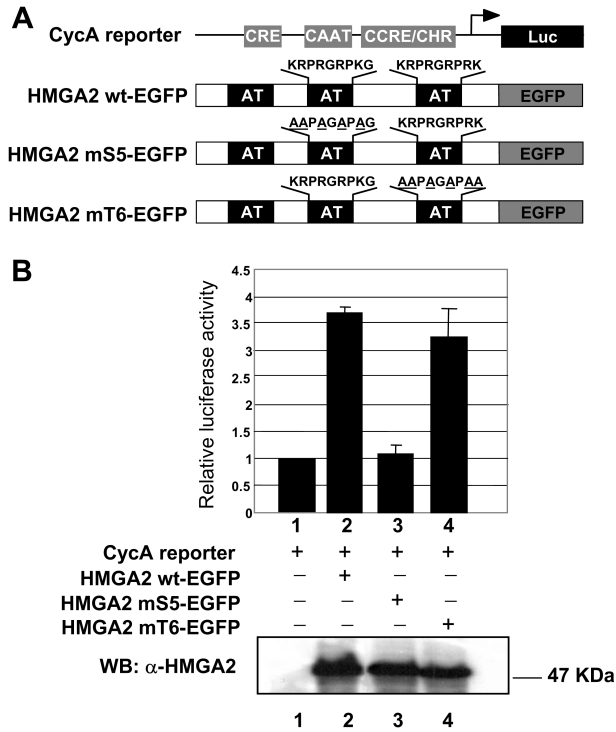
**Figure 4.** Cytoplasmic localization of  $\beta$ -gal/GFP HMGA2 1-43 and HMGA2 point-mutant mS5-EGFP does not result from exposure of a cryptic nuclear export signal. NIH-3T3 cells were transfected with either HMGA2 1-43, HMGA2 mS5-EGFP and GFP-HDAC4 constructs, and fusion proteins were visualized by confocal laser microscopy. Cells were either untreated (-) or treated (+) with 5 ng/ml ratjadone C for 6 h. On the contrary of GFP-HDAC4, localization of  $\beta$ -gal/GFP HMGA2 1-43 and HMGA2 mS5-EGFP does not change following treatment with the nuclear export inhibitor. At least one hundred cells per transfection were analyzed in three different experiments, and more than 95% of transfected cells showed the representative staining. Bars, 10  $\mu$ m.

HMGA proteins are known to be subjected to extensive post-translational modifications which affect their function (1). The second AT-hook of HMGA2 is flanked by two residues, S44 and S55, which can be phosphorylated. The first one was shown to be phosphorylated *in vitro* by p34/cdc2 while the second bears a consensus for PKC. Both consensus are conserved in HMGA1 and the residues were reported to be phosphorylated *in vitro* and *in vivo* (1). Phosphorylation within – or adjacent to – a classic NLS sequence can negatively modulate nuclear localization interfering with binding of cargo proteins

**Figure 5.** The second AT-hook of HMGA2 interacts directly with importin- $\alpha$ 2. (A)  $^{35}$ S-radiolabeled *in vitro* translated HA-tagged HMGA2 was incubated with equivalent amounts of GST, GST importin- $\alpha$ 2 and GST importin- $\beta$  immobilized on Sepharose beads. 20% of the *in vitro* translated reactions used in the pull-down experiments was included as input. As a control, the same experiment was performed with HA-tagged BSAP. (B) To determine the HMGA2 domain involved in importin- $\alpha$ 2 interaction,  $^{35}$ S-radiolabeled *in vitro* translated HMGA2, together with the indicated C-terminal deletion mutants cloned in fusion with  $\beta$ -gal and GFP, were incubated with Sepharose immobilized GST importin- $\alpha$ 2 or GST in GST pull-down assays as described in A. Only the deletion mutant HMGA2 1-43 that lacks the second AT-hook has lost the ability of binding importin- $\alpha$ 2. (C)  $^{35}$ S-radiolabeled *in vitro* translated peptides corresponding to the second AT-hook within the thioredoxin scaffold and in fusion with  $\beta$ -gal and GFP were incubated with GST-importin- $\alpha$ 2 or GST immobilized on Sepharose beads in GST pull-down assays as described in A. The size of all the IVT products are in agreement with the expected molecular weights.







**Figure 6.** HMGA2-mediated transcriptional activation is dependent upon basic residues within the second AT-hook. (A) Diagram of luciferase reporter gene under the transcriptional regulation of the cyclin A promoter and the HMGA2 wild type and mutants expressed by the vectors used. (B) CHO cells were transiently cotransfected with 1 µg of the CycA luciferase reporter plasmid (bars 1–4) and with 3 µg of HMGA2 wt-, mS5- and mT6- EGFP expression vectors (bars 2–4 respectively). 0.1 µg of pRL-CMV Renilla luciferase expression vector was included to normalize for transfection efficiencies. Values are reported as relative luciferase activity. Standard deviations are indicated for experiments repeated three times. The amount of HMGA2 wt-, mS5- and mT6-EGFP expression was assayed by Western blot analysis using a polyclonal α-HMGA2 antibody. The subcellular localization of the expressed proteins in CHO cells was confirmed by confocal microscopy (data not shown).

to importin (29). We therefore tested this possibility by producing a double mutant (S44E, S55E) where serine residues were replaced by glutamic acid that should mimic a constitutive phosphorylation. However, repeated experiments in NIH-3T3 cells always showed an exclusive nuclear localization of this double mutant, even when fluorescence was monitored at early stages after transfection, suggesting therefore a strict nuclear localization for HMGA2 protein which is independent of phosphorylation (data not shown).

Although some works reported also a cytoplasmic localization for HMGA proteins (30–32), HMGA localization seems to be strictly nuclear, as this family of proteins performs several functions at nuclear level, the most relevant being the involvement in transcription regulation. The HMGA2 gene has been described as the most frequent target of cytogenetic rearrangements in human neoplasms. In fact both *HMGA* genes, in particular *HMGA2*, are mutated in several tumors of mesenchymal origin such as lipomas, leiomyomas and hamartomas and have been implicated in the

pathogenesis of these tumors (1,9). *HMGA* alterations result in production of deleted or chimeric proteins that have lost, in the vast majority of the cases, the C-terminal domain but retain the three AT-hooks, while only in some cases also the third AT-hook is lost. An important consequence of our study is therefore that all these deleted or chimeric proteins retain their NLS and are therefore nuclear although the function of the protein could be altered due to ectopic or deleted sequences.

The function of many key proteins involved in cell growth, such as p53 and NF-κB, can be regulated by their cellular localization and nuclear transport has been suggested as a potential target for alteration of cell growth (33). HMGA proteins are both overexpressed and rearranged in several tumors; therefore, molecular strategies able to interfere with their nuclear localization could be important tools to interfere with their transforming activity. The definition of the NLS, reported in this work, could lay the bases for further studies towards this aim.

**ACKNOWLEDGEMENTS**

We are particularly indebted to C. Brancolini for kindly providing reagents and for helpful discussions. We thank B. Birshtein for the pGEX3X importin-α2, I. Mattaj for pGEX4T1 importin-β and T. Stamminger for pHM829 plasmids. We also thank R. Sgarra for comments and suggestions to the manuscript. This work was supported by grants from Associazione Italiana per la Ricerca sul Cancro (AIRC), Milan, Italy to G.M.; Università degli Studi di Trieste, Italy to G.M. and V.G.; and Ministero dell'Università e della Ricerca (PRIN 2004) to V.G and C.P. Funding to the pay the Open Access publication charge was provided by AIRC.

**REFERENCES**

- Sgarra,R., Rustighi,A., Tessari,M.A., Di Bernardo,J., Altamura,S., Fusco,A., Manfioletti,G. and Giancotti,V. (2004) Nuclear phosphoproteins HMGA and their relationship with chromatin structure and cancer. *FEBS Lett.*, **574**, 1–8.
- Johnson,K.R., Lehn,D.A., Elton,T.S., Barr,P.J. and Reeves,R. (1988) Complete murine cDNA sequence, genomic structure, and tissue expression of the high mobility group protein HMGI(Y). *J. Biol. Chem.*, **263**, 18338–18342.
- Manfioletti,G., Giancotti,V., Bandiera,A., Buratti,E., Sautiere,P., Cary,P., Crane-Robinson,C., Coles,B. and Goodwin,G.H. (1991) cDNA cloning of the HMGI-C phosphoprotein, a nuclear protein associated with neoplastic and undifferentiated phenotypes. *Nucleic Acids Res.*, **19**, 6793–6797.
- Zhou,X., Benson,K.F., Ashar,H.R. and Chada,K. (1995) Mutation responsible for the mouse pygmy phenotype in the developmentally regulated factor HMGI-C. *Nature*, **376**, 771–774.
- Fedele,M., Fidanza,V., Battista,S., Pentimalli,F., Klein-Szanto,A.J., Visone,R., De Martino,I., Curcio,A., Morisco,C. et al. (2006) Haploinsufficiency of the Hmga1 gene causes cardiac hypertrophy and myelo-lymphoproliferative disorders in mice. *Cancer Res.*, **66**, 2536–2543.
- Foti,D., Chiefari,E., Fedele,M., Iuliano,R., Brunetti,L., Paonessa,F., Manfioletti,G., Barbetti,F., Brunetti,A. et al. (2005) Lack of the architectural factor HMGA1 causes insulin resistance and diabetes in humans and mice. *Nat Med.*, **11**, 765–773.
- Miyazawa,J., Mitoro,A., Kawashir,S., Chada,K.K. and Imai,K. (2004) Expression of mesenchyme-specific gene HMGA2 in

- squamous cell carcinomas of the oral cavity. *Cancer Res.*, **64**, 2024–2029.
8. Thuault, S., Valcourt, U., Petersen, M., Manfioletti, G., Heldin, C.H. and Moustakas, A. (2006) Transforming growth factor-beta employs HMGA2 to elicit epithelial-mesenchymal transition. *J. Cell. Biol.*, **174**, 175–183.
  9. Fedele, M., Battista, S., Manfioletti, G., Croce, C.M., Giacotti, V. and Fusco, A. (2001) Role of the high mobility group A proteins in human lipomas. *Carcinogenesis*, **22**, 1583–1591.
  10. Fedele, M., Pentimalli, F., Baldassarre, G., Battista, S., Klein-Szanto, A.J., Kenyon, L., Visone, R., De Martino, I., Ciarmiello, A. *et al.* (2005) Transgenic mice overexpressing the wild-type form of the HMGA1 gene develop mixed growth hormone/prolactin cell pituitary adenomas and natural killer cell lymphomas. *Oncogene*, **24**, 3427–3435.
  11. Zaidi, M.R., Okada, Y. and Chada, K.K. (2006) Misexpression of full-length HMGA2 induces benign mesenchymal tumors in mice. *Cancer Res.*, **66**, 7453–7459.
  12. Narita, M., Narita, M., Krizhanovsky, V., Nunez, S., Chicas, A., Hearn, S.A., Myers, M.P. and Lowe, S.W. (2006) A novel role for high-mobility group A proteins in cellular senescence and heterochromatin formation. *Cell*, **126**, 503–514.
  13. Reeves, R. (2001) Molecular biology of HMGA proteins: hubs of nuclear function. *Gene*, **277**, 63–81.
  14. Munshi, N., Agalioti, T., Lomvardas, S., Merika, M., Chen, G. and Thanos, D. (2001) Coordination of a transcriptional switch by HMGI(Y) acetylation. *Science*, **10**, 1133–1136.
  15. Tessari, M.A., Gostissa, M., Altamura, S., Sgarra, R., Rustighi, A., Salvagno, C., Caretti, G., Imbriano, C., Mantovani, R. *et al.* (2003) Transcriptional activation of the cyclin A gene by the architectural transcription factor HMGA2. *Mol. Cell. Biol.*, **23**, 9104–9116.
  16. Fedele, M., Visone, R., De Martino, I., Troncone, G., Calmieri, D., Battista, S., Ciarmiello, A., Pallante, P., Arra, C. *et al.* (2006) HMGA2 induces pituitary tumorigenesis by enhancing E2F1 activity. *Cancer Cell*, **9**, 459–471.
  17. Sorg, G. and Stamminger, T. (1999) Mapping of nuclear localization signals by simultaneous fusion to green fluorescent protein and to beta-galactosidase. *Biotechniques*, **26**, 858–862.
  18. Martelli, A.M., Riccio, M., Bareggi, R., Manfioletti, G., Tabellini, G., Baldini, G., Narducci, P. and Giacotti, V. (1998) Intranuclear distribution of HMGI/Y proteins. An immunocytochemical study. *J. Histochem. Cytochem.*, **46**, 863–864.
  19. Harrer, M., Luhrs, H., Bustin, M., Scheer, U. and Hock, R. (2004) Dynamic interaction of HMGA1a proteins with chromatin. *J. Cell. Sci.*, **117**, 3459–3471.
  20. Jans, D.A., Xiao, C.Y. and Lam, M.H. (2000) Nuclear targeting signal recognition: a key control point in nuclear transport? *Bioessays*, **22**, 532–544.
  21. Nagpal, S., Ghosh, C., DiSepio, D., Molina, Y., Sutter, M., Klein, E.S. and Chandraratna, R.A. (1999) Retinoid-dependent recruitment of a histone H1 displacement activity by retinoic acid receptor. *J. Biol. Chem.*, **274**, 22563–22568.
  22. Meissner, T., Krause, E. and Vinkemeier, U. (2004) Ratjadone and leptomycin B block CRM1-dependent nuclear export by identical mechanisms. *FEBS Lett.*, **576**, 27–30.
  23. Paroni, G., Mizzau, M., Henderson, C., Del Sal, G., Schneider, C. and Brancolini, C. (2004) Caspase-dependent regulation of histone deacetylase 4 nuclear-cytoplasmic shuttling promotes apoptosis. *Mol. Biol. Cell.*, **15**, 2804–2818.
  24. Kovac, C.R., Emelyanov, A., Singh, M., Ashouian, N. and Birshtein, B.K. (2000) BSAP (Pax5)-importin alpha 1 (Rch1) interaction identifies a nuclear localization sequence. *J. Biol. Chem.*, **275**, 16752–16757.
  25. LaCasse, E.C. and Lefebvre, Y.A. (1995) Nuclear localization signals overlap DNA- or RNA-binding domains in nucleic acid-binding proteins. *Nucleic Acids Res.*, **23**, 647–1656.
  26. Huth, J.R., Bewley, C.A., Nissen, M.S., Evans, J.N., Reeves, R., Gronenborn, A.M. and Clore, G.M. (1997) The solution structure of an HMGI(Y)-DNA complex defines a new architectural minor groove binding motif. *Nat. Struct. Biol.*, **4**, 657–665.
  27. Sgarra, R., Tessari, M.A., Di Bernardo, J., Rustighi, A., Zago, P., Liberatori, S., Armini, A., Bini, L., Giacotti, V. *et al.* (2005) Discovering high mobility group A molecular partners in tumour cells. *Proteomics*, **5**, 1494–1506.
  28. Frasca, F., Rustighi, A., Malaguarnera, R., Altamura, S., Vigneri, P., Del Sal, G., Giacotti, V., Pezzino, V., Vigneri, R. *et al.* (2006) HMGA1 inhibits the function of p53 family members in thyroid cancer cells. *Cancer Res.*, **66**, 2980–2989.
  29. Harreman, M.T., Kline, T.M., Milford, H.G., Harben, M.B., Hodel, A.E. and Corbett, A.H. (2004) Regulation of nuclear import by phosphorylation adjacent to nuclear localization signals. *J. Biol. Chem.*, **279**, 20613–20621.
  30. Bamberger, A.M., Makrigiannakis, A., Roser, K., Radde, J., Carstens, T., Flohr, A.M., Bamberger, C.M., Bullerdiek, J. and Loning, T. (2003) Expression of the high-mobility group protein HMGI(Y) in human trophoblast: potential role in trophoblast invasion of maternal tissue. *Virchows Arch.*, **443**, 649–654.
  31. Chiappetta, G., Tallini, G., De Biasio, M.C., Manfioletti, G., Martinez-Tello, F.J., Pentimalli, F., de Nigris, F., Mastro, A., Botti, G. *et al.* (1998) Detection of high mobility group I HMGI(Y) protein in the diagnosis of thyroid tumors: HMGI(Y) expression represents a potential diagnostic indicator of carcinoma. *Cancer Res.*, **58**, 4193–4198.
  32. Masciullo, V., Baldassarre, G., Pentimalli, F., Berlingieri, M.T., Boccia, A., Chiappetta, G., Palazzo, J., Manfioletti, G., Giacotti, V. *et al.* (2003) HMGA1 protein over-expression is a frequent feature of epithelial ovarian carcinomas. *Carcinogenesis*, **24**, 1191–1198.
  33. Kau, T.R., Way, J.C. and Silver, P.A. (2004) Nuclear transport and cancer: from mechanism to intervention. *Nat. Rev. Cancer.*, **4**, 106–117.