BIOCHEMISTRY
including biophysical chemistry & molecular biology

ARTICLE

pubs.acs.org/biochemistry

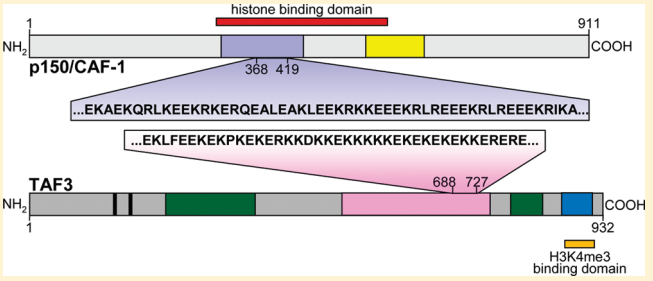
HMGA Interactome: New Insights from Phage Display Technology

Erika Malini,[†] Elisa Maurizio, Sara Bembich,[‡] Riccardo Sgarra, Paolo Edomi,* and Guidalberto Manfioletti*

Department of Life Sciences, University of Trieste, 34127 Trieste, Italy

Supporting Information

ABSTRACT: High mobility group A proteins (HMGA1 and HMGA2) are architectural factors involved in chromatin remodelling and regulation of gene expression. HMGA are highly expressed during embryogenesis and in cancer cells and are involved in development and cell differentiation as well as cancer formation and progression. These factors, by binding to DNA and interacting with other nuclear proteins, can organize macromolecular complexes involved in transcription, chromatin dynamics, RNA processing, and DNA repair. The identification of protein partners for HMGA has greatly contributed to our understanding of their multiple functions. He we report the identification of HMGA molecular partners using a gene fragment library in a phage display screening. Using an ORF-enriched cDNA library, we have isolated several HMGA1 interacting clones and for two of them, TBP associated factor 3 (TAF3) and chromatin assembly factor 1 p150/CAF-1, have demonstrated an *in vivo* association with HMGA1. The identification of these new partners suggests that HMGA can also influence general aspects of transcription and once more underlines their involvement in chromatin remodelling and dynamics.



High mobility group A (HMGA) proteins, including HMGA1 (isoforms HMGA1a and HMGA1b) and HMGA2, are chromatin architectural factors that are involved in several different biological processes, ranging from embryonic development to cell differentiation and transformation, cell cycle progression, apoptosis, and senescence. [For a comprehensive review of HMGA proteins, see the special issue "High Mobility Group (HMG) proteins: from Chromatin function to cellular phenotype" (2010) *Biochim. Biophys. Acta*, Jan−Feb (1−2), 1799]. HMGA expression is very high during embryogenesis, whereas it is undetectable or very low in differentiated adult tissues, but they are re-expressed at high levels following neoplastic transformation.[1] HMGA are in fact among the most widely expressed cancer-associated proteins; their overexpression represents a constant feature of human malignancies and correlates with a poor prognosis.[2] Literature results strongly demonstrate that HMGA play a critical role in tumorigenesis, since blocking their expression reduces the proliferation of cancer cells and in some cases reverts the transformed phenotype. For this reason HMGA proteins have been proposed as potential targets for cancer treatment.[2]

HMGA are modular factors characterized by the presence of three AT-hooks, highly conserved short basic repeats conferring the ability to bind AT-rich DNA sequences, and a constitutively phosphorylated acidic C-terminal tail involved in modulating HMGA interactions.[1] HMGA proteins are highly connected hubs in the chromatin network, establishing a huge number of protein−protein interactions. The utilization of mass spectrometry- and proteomic-based approaches has considerably extended our knowledge of the HMGA molecular network. Four different protein−protein interaction methodologies (affinity chromatography, antibody array, co-immunoprecipitation, and blot overlay) have been used in combination with conventional proteomic strategies leading to the discovery of several partners involved in RNA processing, in chromatin structuring and remodelling, and in DNA repair.[1]

In addition to the finding that the HMGA interactome is quite large, there are also other interesting aspects coming from these protein−protein studies. First of all, the region responsible for almost all HMGA protein−protein contacts has turned out to be confined to a region in the sequence comprising the second AT-hook up to the third. This protein−protein interaction domain therefore assumes a crucial role for the establishment of the entire HMGA network. Another aspect regards the heterogeneity of HMGA molecular partners and in particular the ability of HMGA to interact with molecular partners which apparently do not share any common sequence or structural features. Indeed, it appears that HMGA—thanks probably to their intrinsically disordered nature—can adapt to different regions and domains that, so far, do not seem to fall within specific categories. Therefore, a systematic identification of regions involved in protein−protein interactions with HMGA could provide some essential guidelines to develop short peptides to be potentially used to interfere with HMGA functions within tumor cells.

Among the *in vitro* tools for discovery of protein−protein interaction, the phage display has been amply demonstrated to be one of most powerful and long-lasting technologies,[3,4] whose principal advantage relies on the linking between phenotype and genotype in a selectable format. As a combinatorial technology, it

Received: January 21, 2011
Revised: March 16, 2011
Published: March 18, 2011

has been successfully employed in protein−protein interaction studies, mainly using peptide libraries, but the advantages of using ORF cDNA libraries has also recently been proved.[5,6]

Here we describe the production of an ORF-enriched cDNA library, and using HMGA1a as bait in phage display selection, we have isolated several HMGA1a interacting clones which however turned out to encode for just few proteins. Among them, chromatin assembly factor 1 (p150/CAF-1) and TBP associated factor 3 (TAF3) are of special interest since they are well-known chromatin factors, and for both we could demonstrate that this interaction also occurs *in vivo*. These results demonstrate the reliability of phage display to select for HMGA1a partners and moreover suggest that this approach could be fine-tuned for the identification of short peptide sequences to be used *in vivo* as possible inhibitors of the oncogenic functions of HMGA.

MATERIALS AND METHODS

Plasmids. Plasmids pAR3038 hHMGA2 wt, pAR3038 hHMGA2 1−83, pAR3038 hHMGA2 1−73, pAR3038 hHMGA2 1−54, and pAR3038 hHMGA2 1−43 expressing the human HMGA2 and its C-terminal deletion mutants have been previously described.[7,8] Plasmids pAR3038 hHMGA1a wt, pAR3038 hHMGA1a 1−90, pAR3038 hHMGA1a 1−80, pAR3038 hHMGA1a 1−62, and pAR3038 hHMGA1a 1−52 expressing the human HMGA1a and its C-terminal deletion mutants have been previously described.[8] Numbering refers to the cDNA coding sequence including the N-terminal methionine which is removed in the translated proteins. Amino acid numbering of the translated proteins is adjusted according to the absence of the N-terminal methionine.

Plasmids pGEX HMGA1b and pGEX HMGA2 expressing the murine HMGA1b and the human HMGA2 in fusion with the glutathione *S*-transferase (GST) were already described.[7] The pcDNA3MBP HMGA1a is a eukaryotic expression vector for the production of maltose binding protein (MBP) in fusion with human HMGA1a. It was generated by subcloning within the *Eco*RI and XhoI restriction sites of the pcDNA3MBP vector the ORF of HMGA1a.

The generation of plasmids pGDSV7-VSV-p150-A5, pcDNA-3HATAF3-C1, pcDNAHATAF3, pcDNA3MBP-TAF3$_{671-862}$, and pcDNA3MBP-TAF3$_{671-932}$ is described in the Supporting Information.

cDNA Library Construction, Selection, and Screening. The library construction is described in the Supporting Information. To obtain the phage particles for the panning process, the ORF-selected and recombined library was grown in the presence of 25 μg/mL chloramphenicol and 1% glucose. After infection with M13 K07 helper phage, bacterial cells were grown in 25 μg/mL chloramphenicol and 25 μg/mL kanamycin and the phages were PEG-purified. The biopanning was performed as described,[9] coating an immunotube (Nunc, Roskilde, Denmark) with the recombinant form of human HMGA1a at a concentration of 10 μg/mL. The following washing steps were performed: 10 times with PBS-Tween20 0.1% and 10 times with PBS for the first round, 10 times with PBS-Tween20 0.5% and 10 times with PBS for the second round, and 20 times with PBS-Tween20 0.5% and 20 times with PBS for the third round. Phage ELISA assays were performed as primary screening as described,[10] using peroxidase-conjugated anti-M13 monoclonal antibody (GE Healthcare). M13K07 helper phage and two phages displaying the a.a. 252−393 of p53 and 267−332 of USP11 were used as controls.

Sequence Analysis. All phage display library clones were sequenced using the primer pEPSEQ and analyzed using BLAST and MouseBLAST programs at the National Center for Biotechnology Information (NCBI, USA) and Mouse Genome Informatics (MGI, USA) Web sites, respectively.

Production and Purification of Recombinant Proteins. Recombinant hHMGA1a and hHMGA2, together with their deletion mutant proteins, were produced as already described.[7] HMGA proteins were extracted from bacterial cells by 5% (w/v) perchloric acid, precipitated by acetone−HCl, and purified by RP-HPLC on a Bio-Rad RP304 column (Hercules). Their purity and molecular masses were checked both by SDS-PAGE and MS. Concentration was calculated according to the Waddel method as previously described.[8]

GST Pull-Down and Far-Western Assays. Recombinant GST fusion proteins were expressed as previously described[8] using pGEX HMGA1b and pGEX HMGA2 plasmids and purified by glutathione−agarose affinity chromatography (Glutathione Sepharose 4B; GE Healthcare) following manufacturer's instructions. Their purity, molecular mass, and concentration were checked by SDS-PAGE.

Identified partners were IVT using a commercial *in vitro* translation kit (TNT Promega) and [^{35}S]methionine (NEN Life Science) according to the manufacturer's instructions. About 5 μg of GST and GST-fused HMGA proteins bound to glutathione sepharose beads were incubated with 10 μL of each of the IVT proteins in 200 μL of binding buffer (25 mM HEPES, pH 7.9, 50 mM NaCl, 1 mM DTT, and 0.01% v/v NP-40). After 2 h of incubation at room temperature, the beads were washed twice in complete binding buffer (binding buffer plus 0.25% w/v BSA) and twice in binding buffer without BSA. The bound proteins were then eluted, analyzed by SDS-PAGE, and visualized by fluorography.

In far-western experiments, 1 μg of the wild type HMGA1a and HMGA2 recombinant proteins and their deletion mutants were separated by SDS-PAGE and transferred to nitrocellulose membrane. Far-westerns were then carried out as previously described.[7]

Co-Affinity Purification. For p150/CAF-1 Co-AP experiments, HEK 293T cells were transfected by a conventional calcium phosphate method with pcDNA3MBP-HMGA1a or pcDNA3MBP plasmids as negative control. 36 h later, cells were harvested and lysed in Co-AP buffer (50 mM Tris/HCl pH 8, 150 mM NaCl, 5 mM EDTA, 2 mM DTT, 0.1% (v/v) NP40, 5% (v/v) glycerol) and supplemented with protease inhibitor cocktail, PMSF, NaF, and Na_3VO_4. After sonication, lysates were incubated with 45 μL of resuspended amylose resin (New England Biolabs) for 2 h at 4 °C. After incubation, beads were washed three times with Co-AP buffer and proteins were eluted by boiling in SDS sample buffer. Proteins were separated by 10% SDS PAGE and transferred onto nitrocellulose membrane for immunoblotting. P150 were recognized using α-CAF-1 p150 specific antibody (Santa Cruz Biotechnology). Co-AP for TAF3 was carried out cotransfecting 5 μg of pcDNA3HA-HMGA1a with 5 μg of pcDNA3MBP-TAF3$_{671-862}$, pcDNA3MBP-TAF3$_{671-932}$, or pcDNA3MBP as negative control. 36 h later, cells were harvested and lysed in Co-AP buffer (50 mM Tris/HCl pH 8, 50 mM NaCl, 5 mM EDTA, 2 mM DTT, 0.1% (v/v) NP40, 5% (v/v) glycerol) and processed as described above. HA-HMGA1a was recognized using α-HA antibody (Sigma).

RESULTS AND DISCUSSION

Production and Selection of a Murine CNS Phage Display cDNA Library. A cDNA library from murine brain was constructed in a new phagemid vector, named pEP2, that allows an ORF selection of cDNA fragments cloned in an oriented manner. The scheme for the construction of the murine brain cDNA phage display library is shown in Figure S1 of the Supporting Information, and the features of the library are reported in the Supporting Information. In addition to the well-known advantages of phage display technology, our system allows the enrichment of ORF fragments that could encode for protein domains. The chosen length of cloned cDNA fragments (200−500 bp) allows the local folding of the encoded polypeptide exposed on the phage surface and thus its interaction with the partner protein.[11]

A biopanning of the murine brain cDNA phage display library with the recombinant form of the HMGA1a protein was carried out using the same titer of input phages (1×10^8) for three rounds. The final enrichment, calculated as the ratio of output/input phages, was of 57-fold (Table 1). Both the increase in output phage number and the high enrichment indicate a high specificity of these selection conditions. We analyzed 144 clones, randomly picked from the last round, for their binding to HMGA1a in two independent phage-ELISA assays. A phage

Table 1. Enrichment of Phages after Selection of Total Murine Brain ORF-Enriched cDNA Library On HMGA1a

round	1st	2nd	3rd
output/input	6.00×10^{-5}	2.60×10^{-4}	3.40×10^{-3}
enrichment[a]		4.33	56.67

[a] Enrichment = ratio of output/input value for each round on the same value of first round.

Table 2. Summary of Phage-ELISA Results[a]

	O.D. > +ctrl	O.D. > −ctrl	O.D. > helper
total clones	41/144 (28%)	113/144 (78%)	131/144 (91%)
sequenced clones	27/42 (64%)	38/42 (90%)	42/42 (100%)

[a] Percentage of total and sequenced clones from phage display selection with O.D. value higher than controls.

Table 3. Sequence Data of Clones from Selection of Murine Brain cDNA Library

consensus/ clone[a]	*N* overlapped	consensus nt length	symbol	name	MGIId	homology[b] (%)	a.a.	Phage-ELISA[c]
HMG_cons0	10	243	Gm9861	putative uncharacterized protein	AK046 786	67	58−120	O.D. ≥ (+ctrl)
HMG _cons1	29	122	Taf3	TAF3 RNA polymerase II, (TBP)-associated factor	NM_027 748	100	688−727	O.D. ≥ (helper)
HMG_cons2	2	161	Chafla	chromatin assembly factor 1, subunit A (p150)	NM_013 733	100	368−419	O.D. ≥ (+ctrl)
HMG_1H5	1	134		*Mus musculus* hypothetical protein LOC100 046866	XM_001 476948	95	68−109	O.D. ≥ (+ctrl)

[a] Data of consensus sequence is indicated for overlapping clones. [b] Blast homology resulting from ratio between amino acids identical to recognized protein on total amino acids of identified consensus/clone sequence. [c] Data from Phage-ELISA assays.

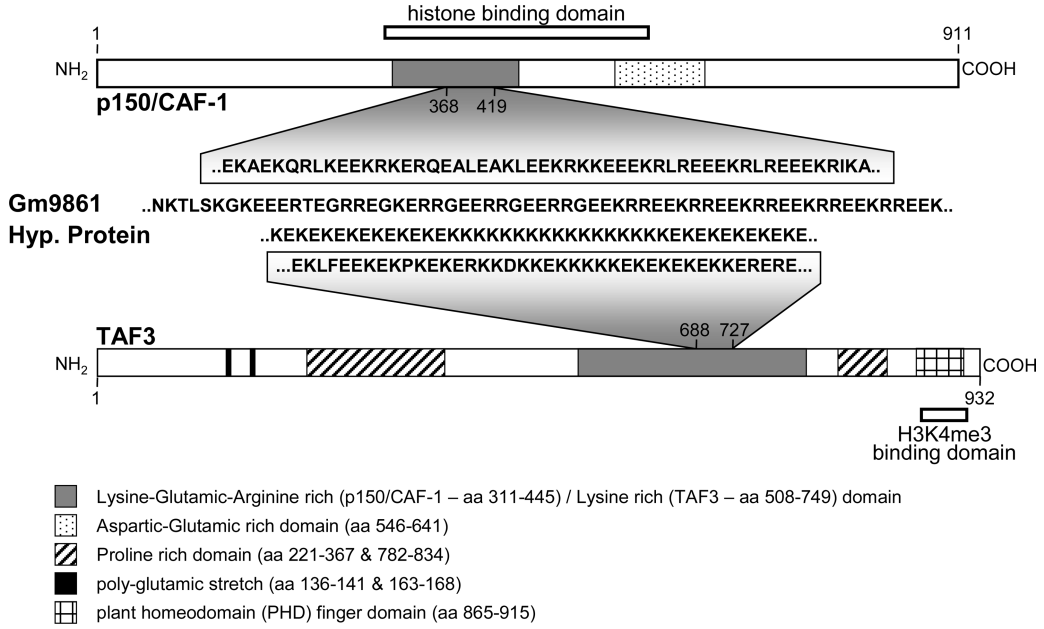


Figure 1. Schematic representation of the consensus sequences of the four interacting peptides identified from phage display library selection. For p150/CAF-1 and TAF3 the entire protein sequence is shown with relevant domain annotations. Numbers indicated correspond to amino acids positions (data from www.uniprot.org). Hyp. protein: hypothetical protein.

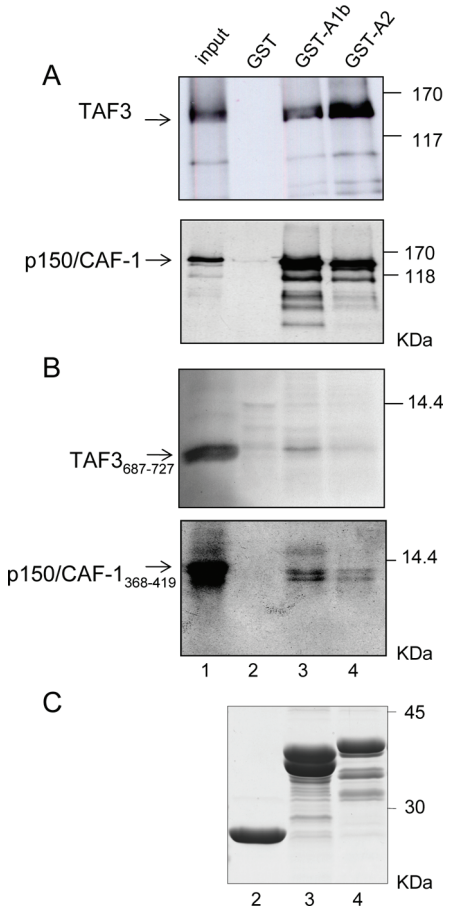


Figure 2. GST pull-down assays with recombinant HMGA proteins (A) ^{35}S-radiolabeled, *in vitro* translated full length TAF3 and p150/CAF-1 were incubated with GST-HMGA1b, GST-HMGA2, or GST immobilized on sepharose beads in GST pull-down assays. The bound proteins were resolved on SDS-PAGE and visualized by fluorography. 25% of the *in vitro* translated reactions used in the pull-down experiments were included (input). (B) ^{35}S-radiolabeled, *in vitro* translated TAF3$_{687-727}$ and p150/CAF-1$_{368-419}$ were incubated with GST-HMGA1b, GST-HMGA2, or GST immobilized on sepharose beads in GST pull-down assays. The bound proteins were resolved on SDS-PAGE and visualized by fluorography. 25% of the *in vitro* translated reactions used in the pull-down experiments were included (input). (C) A representative SDS-PAGE of the protein used in the experiments, stained with Blue Coomassie, is shown.

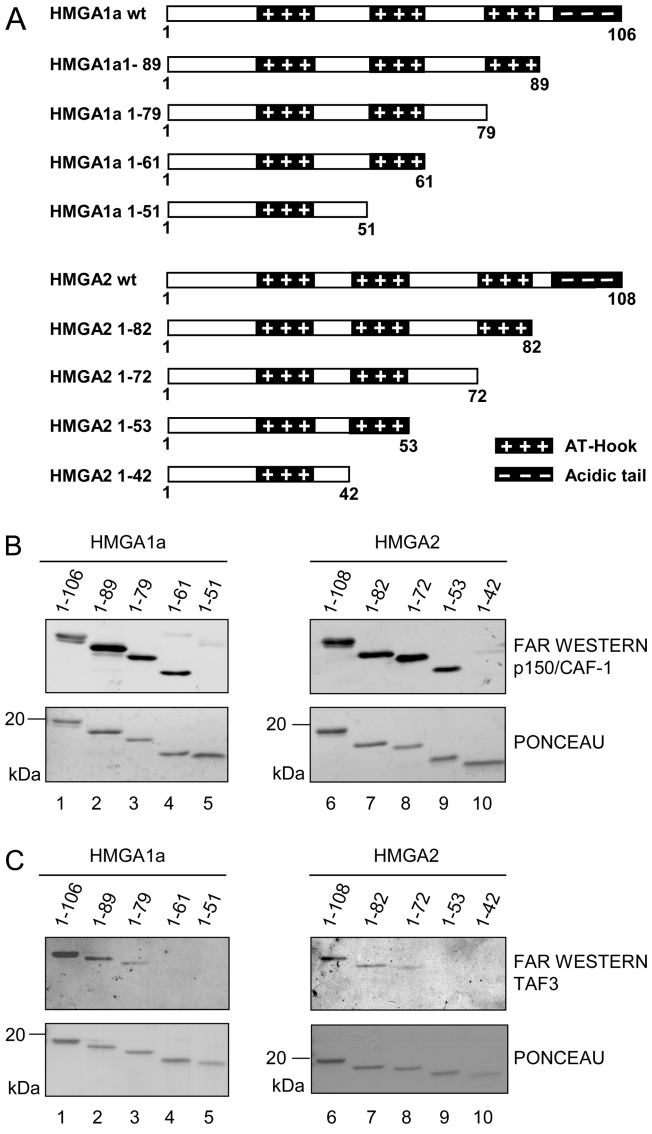


Figure 3. Far-western experiments with HMGA deletion mutants. (A) Schematic representation of the various HMGA1a and HMGA2 forms used in the far western experiments. (B) HMGA1a and HMGA2, with a series of their deletion mutants (HMGA1a1−89, 1−79, 1−61 and 1−51; HMGA21−82, 1−72, 1−53, 1−42) were SDS-PAGE separated and transferred to PVDF membranes. Far-western experiments were then performed incubating the membranes with ^{35}S-radiolabeled, *in vitro* translated full length p150/CAF-1. Bound proteins were visualized by fluorography. For quantification purposes, a ponceau stained membrane is shown. (C) Similar far-western experiments shown in (B) performed with TAF3.

particle displaying the amino acidic portion 252−393 of human p53, involved in binding with HMGA1a,[12] was used as positive control in both cases. A helper phage that displays only the g3 protein, and a recombinant phage isolated from another panning and not interacting with HMGA1a (unpublished data), were used as negative controls. Since the optical density in the two assays were comparable for all these controls, all data were collectively analyzed (Table 2). Almost all clones showed an optical density higher than helper phage and about one-third of the isolated phages carried peptides with apparent binding capacity greater than the p53 fragment. Among these clones, 42 were subjected to sequencing to identify the original cDNA and the correct ORF-selection; 64% of the sequenced clones were chosen among those showing a reactivity higher than p53 (cf. Table 2). All clones were in frame with the g3 protein, confirming the ORF selection. Comparing the sequences of the clones, almost all were grouped in three consensus sequences (only one clone was single, see Table 3); the overlapping regions were identical among the clones of the same consensus. These consensus sequences were correctly oriented, encoded an ORF portion of murine cDNA, and showed from 67 to 100% nucleotide homology with murine sequences. Excluding the consensus HMG_cons0, that identified an uncharacterized protein, the homology of the other three sequences ranged from 95 to 100% (see Table 3). All four of these putative interactors of HMGA1a showed an optical density higher than helper phage in ELISA assays, and three were higher than the positive control

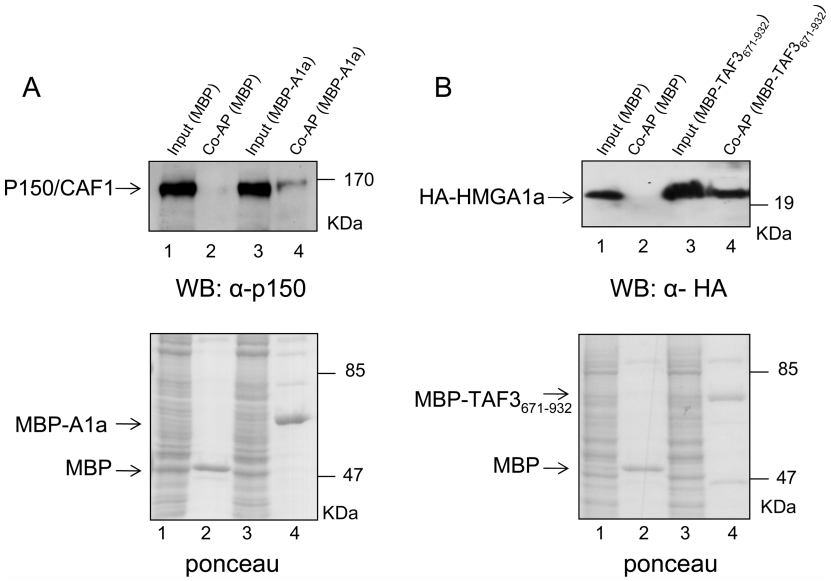


Figure 4. Co-affinity purification of p150/CAF-1 and TAF3. (A) HMGA1a fused to maltose binding protein (MBP-A1a) or the maltose binding protein (MBP) alone were produced by transient transfection in 293T cells. Cellular lysates (input, 1% of the amount used, lanes 1 and 3) were incubated with amylose resin and affinity captured MBP-A1a and MBP proteins recovered. Bound proteins were eluted by boiling in SDS sample buffer. Proteins were separated by SDS PAGE (T = 10%) and analyzed by WB using an α-p150/CAF-1 antibody. (B) TAF3 fused to maltose binding protein (MBP-TAF3) or the maltose binding protein (MBP) alone were produced by transient transfection in 293T cells together with HA-HMGA1a. Cellular lysates (input, 1% of the amount used, lanes 1 and 3) were incubated with amylose resin and affinity captured MBP-TAF3 and MBP proteins recovered. Bound proteins were eluted by boiling in SDS sample buffer. Proteins were separated by SDS PAGE (T = 10%) and analyzed by WB using an α-HA antibody. The ponceau stained membrane of representative experiments are shown below.

(Table 3). Two peptides are particularly interesting because they correspond to known nuclear proteins: the chromatin assembly factor 1 (p150/CAF-1) and the TBP associated factor 3 (TAF3). The isolated peptides match residues 368−419 of CAF-1 and 688−727 of TAF3 (Figure 1).

TAF3 belongs to the TFIID complex which is required by RNA polymerase II for transcription and interacts with several transcription factors like TAF10, TAF13, TBP, SAP130, and GCN5L2.[13] TAF3 is involved in the switching of core promoter complexes that allow to selectively turn on one transcriptional program while silencing many others.[14]

P150/CAF-1 is a component of the CAF-1 complex that is believed to mediate chromatin assembly in DNA replication and DNA repair.[15,16] P150/CAF-1 binds directly to histones H3 and H4, CAF-1B, and PCNA proteins.[17] Its interaction with MBD1 suggests a role in methylation-mediated transcriptional repression and the inheritance of epigenetically determined chromatin states.[18]

The other two candidate partners of HMGA1a are hypothetical proteins predicted from sequencing data but share with p150/CAF-1 and TAF3 the amino acid content of the region involved in the selection, suggesting that it could be characteristic of the HMGA binding domain.

HMGA1 and HMGA2 Interact *in Vitro* with Full Length TAF3 and p150/CAF-1. To confirm the interactions of HMGA1a with TAF3 and p150/CAF-1 identified as potential partners using phage display, we performed GST pull-down experiments. To this end we produced full length IVT [^{35}S]methionine-radiolabeled proteins and tested for their ability to interact with recombinant GST fused HMGA1b and HMGA2 proteins in order to confirm and extend the interactions to the entire HMGA family. In fact, HMGA1b is a splicing variant of HMGA1a, differing in having a deletion of only 11 amino acids between the first and the second AT-hook, while HMGA2 is the product of a different but highly related gene showing about 50% identity with HMGA1a.[1]

As shown in Figure 2A, both TAF3 and p150/CAF-1 were retained by both GST-HMGA1b and GST-HMGA2 (lanes 3 and 4) but not by GST alone (lane 2).

To confirm that the interaction was mediated directly by the peptide sequences identified by phage display, we performed GST pull-down experiments using the peptide sequences of TAF3 ($TAF3_{687-727}$) and p150/CAF-1 ($p150/CAF\text{-}1_{368-419}$) IVT [^{35}S]methionine-radiolabeled in fusion with HA and VSV tag, respectively. Figure 2B shows that both peptides were able to specifically interact with both HMGA1b and HMGA2 although the interaction is much weaker compared to that obtained with the full length protein. Equal amounts of the three proteins (GST, GST-HMGA1b, and GST-HMGA2) were used as shown in Figure 2C.

These results not only confirm the interactions between the short peptide sequences and the two nuclear factors but also demonstrate that the interaction exists in the context of fully structured full length proteins.

We then decided to investigate which HMGA protein regions were involved in the evidenced interactions. To this end, we performed farwestern experiments using a panel of HMGA1a and HMGA2 C-terminal deletion mutants, a scheme of these proteins is shown in Figure 3A. Equivalent amounts of these proteins were separated on SDS-PAGE and transferred to nitrocellulose membranes. These membranes were then incubated with the radiolabeled IVT HMGA partners, and bound proteins were revealed by fluorography. Representative farwestern experiments are reported in panels B and C of Figure 3 for both p150/CAF-1 and TAF3. As can be seen, p150/CAF-1

binds to full length HMGA1a and HMGA2 proteins (HMGA1a 1−106 and HMGA2 1−108) and to C-terminal deletion mutants lacking the acidic C-terminal tail and the third AT-hook (HMGA1a 1−89, 1−79, and 1−61 and HMGA2 1−82, 1−72, and 1−53). Removal of the second AT-hook (HMGA1a 1−51 and HMGA2 1−42), however, completely abolishes the binding of p150/CAF-1 to both HMGA1a and HMGA2.

TAF3 interacts with full-length HMGA1a and HMGA2 (HMGA1a 1−106 and HMGA2 1−108) and deletion mutants lacking the C-terminal acidic tail and the third AT-hook (HMGA1a 1−89, 1−79 and HMGA2 1−82 and 1−72). At variance with p150/CAF-1, TAF3 does not bind HMGA1a and HMGA2 deletion mutants still retaining the second AT-hook (HMGA1a 1−61 and HMGA2 1−53).

Taken together, these results demonstrate that the region comprising the second AT-hook in both HMGA1a and HMGA2 is indispensable for protein−protein interaction with p150/CAF-1 while the region between the second and third AT-hook seems to be more critical for the interaction with TAF3.

Both p150/CAF-1 and TAF3 Associate *in Vivo* with HMGA1a. To demonstrate that HMGA1a can also associate *in vivo* with p150/CAF-1 and TAF3, a co-affinity purification (Co-AP) was carried out. Vectors expressing HMGA1a in fusion with maltose binding protein (MBP-A1a) and MBP alone, as a negative control, were transfected into 293T cells. Soluble protein complexes were purified using amylose resin and then analyzed by WB using α-p150/CAF-1 antibody. As shown in Figure 4A, endogenous p150/CAF-1 was detected only in presence of MBP-A1a (lane 4) but not with MBP alone (lane 2). Lanes 1 and 3 show that the same amount of cell lysates (input) was used and that both MBP and MBP-A1a are expressed at the same levels (lower panel of Figure 4A, lanes 2 and 4, respectively).

To demonstrate the association of HMGA1a with TAF3, cells were cotransfected with the MBP expression vectors, expressing MBP-TAF3$_{671-932}$ and MBP alone, in combination with HA tagged HMGA1a. We could not perform the experiment using the endogenous TAF3 because suitable antibodies were not available, and since the full-length MBP-tagged TAF3 protein could not be expressed, we used a deletion mutant, TAF3$_{671-932}$, that includes the sequence involved in the interaction with HMGA1. Soluble protein complexes were purified using amylose resin and then analyzed by WB with α-HA antibody. As shown in Figure 4B, HA-tagged HMGA1a was detected only in presence of MBP-TAF3$_{671-932}$ (lane 4) but not with MBP alone (lane 2). Lanes 1 and 3 show that the same amount of cell lysates (input) was used and that both MBP proteins are expressed at the same levels (lower panel of Figure 4B, lanes 2 and 4, respectively). Similar results were obtained using a shorter deletion construct of TAF3, MBP-TAF3$_{671-862}$ (data not shown).

Taken together, these results demonstrate that the phage display approach can be used as a tool to successfully identify HMGA protein partners.

CONCLUSIONS

One of the concepts that has been firmly established in the proteomic era is that proteins operate within molecular networks and that the peculiar functions they perform depend strictly on the molecular partners they interact with. This justifies the growing demand for strategies to reveal protein−protein interactions. Previous large-scale proteomic approaches identified several HMGA molecular partners with a very low overlap.[7,8,19] This is probably due to the fact that the HMGA interactome is quite large, but also that each approach is based on different techniques each with advantages and disadvantages. For example, high molecular weight proteins (above 100 kDa) are usually difficult to visualize in 2D gel-based approaches,[7,8] and this is probably the reason why in these approaches p150/CAF-1 and TAF3 were not identified. Therefore, the use of different approaches based on different techniques is invaluable to explore the protein interactome and can give therefore complementary results. It is of note, for example, that the smaller subunit of CAF-1 (p48/CAF1—also known as RBBP-4) was found among the proteins in complex with HMGA2 in a previous approach.[7]

The phage display approach provides the well-known advantage of handling larger numbers of clones, compared to other types of libraries. In addition the use of an ORF-enriched cDNA library enhances the possibility of displaying folded domains.[20] Using such an ORF fragment library for the display of functional domains, we have demonstrated that the phage display approach is successful in discovering new HMGA partners. With this approach we were not only able to evidence interactions involving specific amino acid regions, such as the E, R, and K rich domains, but also to directly identify the proteins to which those regions belong to. Moreover, it demonstrates the efficiency and sensitivity of real ORF displaying in mapping protein−protein interactions.[5] Considering the very high stringency conditions used in the biopanning, it is understandable that only a few proteins have been identified, but in the future, by playing with the parameters that influence the interactions it will be possible to identify a larger number of clones.

Given that HMGA proteins are nuclear chromatin factors, it is significant that (i) the two identified interactors (TAF3 and p150/CAF-1) are nuclear proteins and that (ii) they have been confirmed to interact with HMGA1, both *in vitro* and *in vivo*, using different strategies. The interaction with both TAF3 and p150/CAF-1 suggests new functions for HMGA1.

TAF3 is a TATA-binding protein associated factor associated with the TFIID complex, part of the "basal transcription machinery". Despite being described as part of the core promoter recognition apparatus, TAF3, in association with TRF3, has been demonstrated to be able to disrupt and replace the canonical holo-TFIID complex leading to cell differentiation.[13] This mechanism of interference with core promoter complexes has been shown to constitute a simple yet effective means to selectively turn on one transcriptional program while silencing many others, representing a more general mechanism for regulating cell type-specific terminal differentiation.[14] The association of HMGA1 with TAF3 could therefore suggest an involvement of HMGA in regulating general aspects of transcription, a hypothesis that is consistent with its relevance regarding both embryogenesis and neoplastic transformation, biological processes in which HMGA proteins have been demonstrated to play crucial roles.[2,21]

The interaction of HMGA1 with the histone chaperone p150/CAF-1 underlines once more the involvement of HMGA in chromatin remodelling and dynamics, as suggested by our previous finding regarding the interaction of HMGA proteins with nucleophosmin (NPM) and nucleolin (NCL).[7,8] Indeed, NPM and NCL, as well as p150/CAF-1, are histone chaperones involved in the proper assembly of protein/DNA complexes and in particular in the formation of nucleosomes and higher order chromatin structures.[22] The functional significance of HMGA1/p150/CAF-1 association still needs to be unravelled, but whether

HMGA1 modulates p150/CAF-1 histone chaperone activity or p150/CAF-1 modulates HMGA1 DNA binding properties, this could lead to relevant changes at the chromatin structure level.

The identification of E-R-K rich peptide sequences binding with high affinity to HMGA points for the first time to an HMGA binding consensus, laying the basis for the development/discovery of structurally related peptides able to interact with HMGA, potentially to be used *in vivo* as inhibitors of the HMGA oncogenic functions.

■ ASSOCIATED CONTENT

Supporting Information. Materials and methods for plasmids and ORF-selected cDNA phage display library production, results about the production of an ORF-selected phage display cDNA library from murine CNS, table on selected and unselected "in frame" clones, and figure of the flowchart of phage display ORF library construction. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Tel +39-040-5583560, fax +39-040-5583694, e-mail edomi@units.it (P.E.); Tel +39-040-5583675, fax +39-040-5583694, e-mail manfiole@units.it (G.M.).

Present Addresses

†Regional Coordination Centre for Rare Diseases, University Hospital Santa Maria della Misericordia, 33100 Udine, Italy.

‡International Centre for Genetic Engineering and Biotechnology, Area Science Park, 34149 Trieste, Italy.

Funding Sources

This work was supported by grants from Associazione Italiana per la Ricerca sul Cancro (AIRC), Milan (Italy), Progetto AITT 2007 Regione FVG, Università degli Studi di Trieste, MIUR (PRIN 2007), project "Rete Nazionale Proteomica" FIRB 2008-RBRN07BMCT to G.M. and partially from Fondazione Italiana Sclerosi Multipla (FISM) 2004R6 to P.E.

■ ACKNOWLEDGMENT

We thank Dr. I. Davidson (Institut de Génétique et de Biologie Moléculaire et Cellulaire, Cédex, France) for kindly providing pXJ41 mTAF3-FLAG plasmid.

■ ABBREVIATIONS

ORF, open reading frame; CNS, central nervous system; GST, glutathione-S-transferase; SDS-PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; Co-AP, Co-affinity purification; MBP, maltose binding protein; WB, Western blot; IVT, *in vitro* translated; HA, human influenza hemagglutinin; VSV, vesicular stomatitis virus.

■ REFERENCES

(1) Sgarra, R., Zammitti, S., Lo Sardo, A., Maurizio, E., Arnoldo, L., Pegoraro, S., Giancotti, V., and Manfioletti, G. (2010) HMGA molecular network: From transcriptional regulation to chromatin remodeling. *Biochim. Biophys. Acta 1799*, 37–47.

(2) Fedele, M., and Fusco, A. (2010) HMGA and cancer. *Biochim. Biophys. Acta 1799*, 48–54.

(3) Sidhu, S. S., and Koide, S. (2007) Phage display for engineering and analyzing protein interaction interfaces. *Curr. Opin. Struct. Biol. 17*, 481–487.

(4) Rhyner, C., Kodzius, R., and Crameri, R. (2002) Direct selection of cDNAs from filamentous phage surface display libraries: potential and limitations. *Curr. Pharm. Biotechnol. 3*, 13–21.

(5) Li, W., and Caberoy, N. B. (2010) New perspective for phage display as an efficient and versatile technology of functional proteomics. *Appl. Microbiol. Biotechnol. 85*, 909–919.

(6) Di Niro, R., Sulic, A. M., Mignone, F., D'Angelo, S., Bordoni, R., Iacono, M., Marzari, R., Gaiotto, T., Lavric, M., Bradbury, A. R., Biancone, L., Zevin-Sonkin, D., De Bellis, G., Santoro, C., and Sblattero, D. (2010) Rapid interactome profiling by massive sequencing. *Nucleic Acids Res. 38*, e110.

(7) Sgarra, R., Tessari, M. A., Di Bernardo, J., Fustighi, A., Zago, P., Liberatori, S., Armini, A., Bini, L., Giancotti, V., and Manfioletti, G. (2005) Discovering high mobility group A molecular partners in tumour cells. *Proteomics 5*, 1494–1506.

(8) Sgarra, R., Furlan, C., Zammitti, S., Lo Sardo, A., Maurizio, E., Di Bernardo, J., Giancotti, V., and Manfioletti, G. (2008) Interaction proteomics of the HMGA chromatin architectural factors. *Proteomics 8*, 4721–4732.

(9) Zacchi, P., Sblattero, D., Florian, F., Marzari, R., and Bradbury, A. R. (2003) Selecting open reading frames from DNA. *Genome Res. 13*, 980–990.

(10) Marks, J. D., Griffiths, A. D., Malmqvist, M., Clackson, T. P., Bye, J. M., and Winter, G. (1992) By-passing immunization: building high affinity human antibodies by chain shuffling. *Biotechnology 10*, 779–783.

(11) Yunjia, C., Shihong, Q., Chi-Hao, L., and Ming, L. (2007) Domain selection combined with improved cloning strategy for high throughput expression of higher eukaryotic proteins. *BMC Biotechnol. 7*, 45–58.

(12) Frasca, F., Rustighi, A., Malaguarnera, R., Altamura, S., Vigneri, P., Del Sal, G., Giancotti, V., Pezzino, V., Vigneri, R., and Manfioletti, G. (2006) HMGA1 inhibits the function of p53 family members in thyroid cancer cells. *Cancer Res. 66*, 2980–2989.

(13) Juven-Gershon, T., and Kadonaga, J. T. (2010) Regulation of gene expression via the core promoter and the basal transcriptional machinery. *Dev. Biol. 339*, 225–229.

(14) Deato, M. D., and Tjian, R. (2007) Switching of the core transcription machinery during myogenesis. *Genes Dev. 21*, 2137–2149.

(15) Li, Q., Fazly, A. M., Zhou, H., Huang, S., Zhang, Z., and Stillman, B. (2009) The elongator complex interacts with PCNA and modulates transcriptional silencing and sensitivity to DNA damage agents. *PLoS Genet. 5*, e1000684.

(16) Li, Q., Zhou, H., Wurtele, H., Davies, B., Horazdovsky, B., Verreault, A., and Zhang, Z. (2008) Acetylation of histone H3 lysine 56 regulates replication-coupled nucleosome assembly. *Cell 134*, 244–255.

(17) Rolef Ben-Shahar, T., Castillo, A. G., Osborne, M. J., Borden, K. L., Kornblatt, J., and Verreault, A. (2009) Two fundamentally distinct PCNA interaction peptides contribute to chromatin assembly factor 1 function. *Mol. Cell. Biol. 29*, 6353–6365.

(18) Reese, B. E., Bachman, K. E., Baylin, S. B., and Rountree, M. R. (2003) The methyl-CpG binding protein MBD1 interacts with the p150 subunit of chromatin assembly factor 1. *Mol. Cell. Biol. 23*, 3226–3236.

(19) Pierantoni, G. M., Esposito, F., Giraud, S., Bienvenut, W. V., Diaz, J. J., and Fusco, A. (2007) Identification of new high mobility group A1 associated proteins. *Proteomics 20*, 3735–3742.

(20) Di Niro, R., Ferrara, F., Not, T., Bradbury, A. R. M., Chirdo, F., Marzari, R., and Sblattero, D. (2005) Characterizing monoclonal antibody epitopes by filtered gene fragment phage display. *Biochem. J. 388*, 889–894.

(21) Ashar, H. R., Chouinard, R. A., Jr., Dokur, M., and Chada, K. (2010) *In vivo* modulation of HMGA2 expression. *Biochim. Biophys. Acta 1799*, 55–61.

(22) Frehlick, L. J., Eirín-López, J. M., and Ausió, J. (2007) New insights into the nucleophosmin/nucleoplasmin family of nuclear chaperones. *Bioessays 29*, 49–59.