A novel zinc finger transcriptional repressor, ZNF224, interacts with the negative regulatory element (AldA-NRE) and inhibits gene expression

L. Medugno[a], P. Costanzo[b], A. Lupo[c], M. Monti[d], F. Florio[a], P. Pucci[d,e], P. Izzo[a,e,*]

[a] *Dipartimento di Biochimica e Biotecnologie Mediche, Facoltà di Medicina e Chirurgia, Università di Napoli 'Federico II', Via S. Pansini 5, 80131 Naples, Italy*
[b] *Dipartimento di Medicina Sperimentale e Clinica, Facoltà di Medicina e Chirurgia, Università degli Studi di Catanzaro Magna Graecia, Via T. Campanella 115, 88100 Catanzaro, Italy*
[c] *Dipartimento di Scienze Biologiche ed Ambientali, Facoltà di Scienze, Università del Sannio, Via Port'Arsa 11, 82100 Benevento, Italy*
[d] *Dipartimento di Chimica Organica e Biochimica, Via Cinthia – Monte S. Angelo, 80126 Naples, Italy*
[e] *CEINGE, Via S. Pansini 5, 80131 Naples, Italy*

Received 15 November 2002; accepted 25 November 2002

First published online 11 December 2002

Edited by Jesus Avila

Abstract The interaction between the negative *cis*-element (AldA-NRE) and p97 repressor nuclear protein is a key step in modulating transcription of the human and mouse aldolase A (AldA) gene during the cell cycle and differentiation. In an attempt to clarify the role of transcriptional repression in regulating gene expression, we purified, from HeLa cells, the nuclear protein that specifically binds to the AldA negative regulatory element (NRE). Matrix-assisted laser desorption ionization-time of flight analysis and examination of protein profiles from the SwissProt database revealed that the previously defined p97 repressor is ZNF224, a zinc finger protein. We demonstrate that ZNF224, a Kruppel-like zinc finger transcription factor, is the repressor protein that specifically binds to the negative *cis*-element AldA-NRE and affects the AldA-NRE-mediated transcription.

Key words: Negative regulatory element; Matrix-assisted laser desorption ionization; Zinc finger; Transcriptional repressor; Kruppel-associated box domain; Aldolase A

1. Introduction

An interplay of interactions between nuclear proteins and DNA *cis*-elements within the promoter regions upstream of leader exons L1, M and F is responsible for modulating aldolase A (AldA) gene expression [1–5]. We previously identified a negative regulatory element (hAldA-NRE) that binds a nuclear protein and represses transcription in its own configuration and transcription of a reporter gene driven by a heterologous promoter [3,5]. In the murine system we detected a negative regulatory element (mAldA-NRE), highly homologous to the human silencer which binds a 97 kDa specific nuclear protein (p97) [6]. We also demonstrated that the interaction between p97 and AldA-NRE is involved in the regulation of aldolase A gene expression during the cell cycle and cell differentiation [7,8].

Transcriptional repression plays a crucial role in determining gene expression during development and differentiation. Over the years factors that repress gene expression have been isolated, and their mechanisms have been studied in great detail. For example, the transcriptional repressor Zeb shows a regulatory function in vertebrate myogenesis [9]; the zinc finger neuron-restrictive silencer factor is involved in regulation of neuronal genes [10]; mammalian CUT homeodomain protein is part of a network which controls the G_1-S transition during the cell cycle [11].

We have purified by affinity chromatography from HeLa nuclear extract a 82 kDa nuclear protein that is involved in the formation of repressor complex with AldA-NRE within the human aldolase A distal promoter (pL). Matrix-assisted laser desorption ionization (MALDI) technology and comparison of the profile of the purified protein with those of proteins from the SwissProt database revealed a zinc finger 224 (ZNF224) cDNA product.

ZNF224 cDNA encodes a protein belonging to the Kruppel-like zinc finger class, one of the largest families of transcriptional factors which is divided into many subclasses based on the number and type of zinc finger they contain [12–14]. ZNF224 protein contains a Kruppel-associated box (KRAB) domain, an evolutionarily conserved motif present at the NH_2-terminus of several Kruppel-like zinc finger proteins, and 19 Cys_2-His_2 zinc finger domains. Evidence suggests that KRAB zinc finger proteins are involved in transcriptional repression of gene expression and that the 75 amino acids long KRAB domain functions as a mediator of a composite repression apparatus [15–18]. Here we demonstrate that ZNF224 specifically binds to the AldA-NRE motif through its array of C_2H_2 zinc fingers and inhibits the transcription of a eukaryotic promoter bearing the AldA-NRE sequence. Therefore, ZNF224 is the repressor protein of aldolase A gene expression, previously defined as p97 [7].

2. Materials and methods

2.1. Cell cultures and nuclear extract preparation

Human HeLa cells were grown in Dulbecco's modified Eagle's medium (DMEM), supplemented with 10% fetal calf serum. Cells were grown to a density of 9×10^5 cells/ml and nuclear extracts were prepared essentially as described [19].

*Corresponding author. Fax: (39)-81-7464359.
E-mail address: izzo@dbbm.unina.it (P. Izzo).

Abbreviations: AldA-NRE, aldolase A negative regulatory element; KRAB, Kruppel-associated box; ZNF, zinc finger; MALDI, matrix-assisted laser desorption ionization; GST, glutathione *S*-transferase

0014-5793/02/$22.00 © 2002 Published by Elsevier Science B.V. on behalf of the Federation of European Biochemical Societies.
doi:10.1016/S0014-5793(02)03783-3

2.2. DEAE and DNA affinity chromatography

The nuclear extract from HeLa cells was applied onto a DEAE-Sepharose column (Pharmacia Biotech). The activity of p97 was monitored during each purification step with electrophoretic mobility shift assay (EMSA). The fractions containing the protein of interest eluted at 0.3–0.5 M KCl from DEAE-Sepharose chromatography were pooled, dialyzed in buffer Z (25 mM HEPES, pH 7.6, 50 mM KCl, 12.5 mM $MgCl_2$, 1 mM dithiothreitol (DTT), 20% glycerol, 0.1% NP40) and applied on DNA affinity chromatography, performed as described [20]. The double-stranded synthetic oligonucleotide Neg1 (TCCCCTTAGAGAGCAACAGACGTGT) (250 μg), which contains the AldA-NRE binding site [5], was annealed, phosphorylated, and

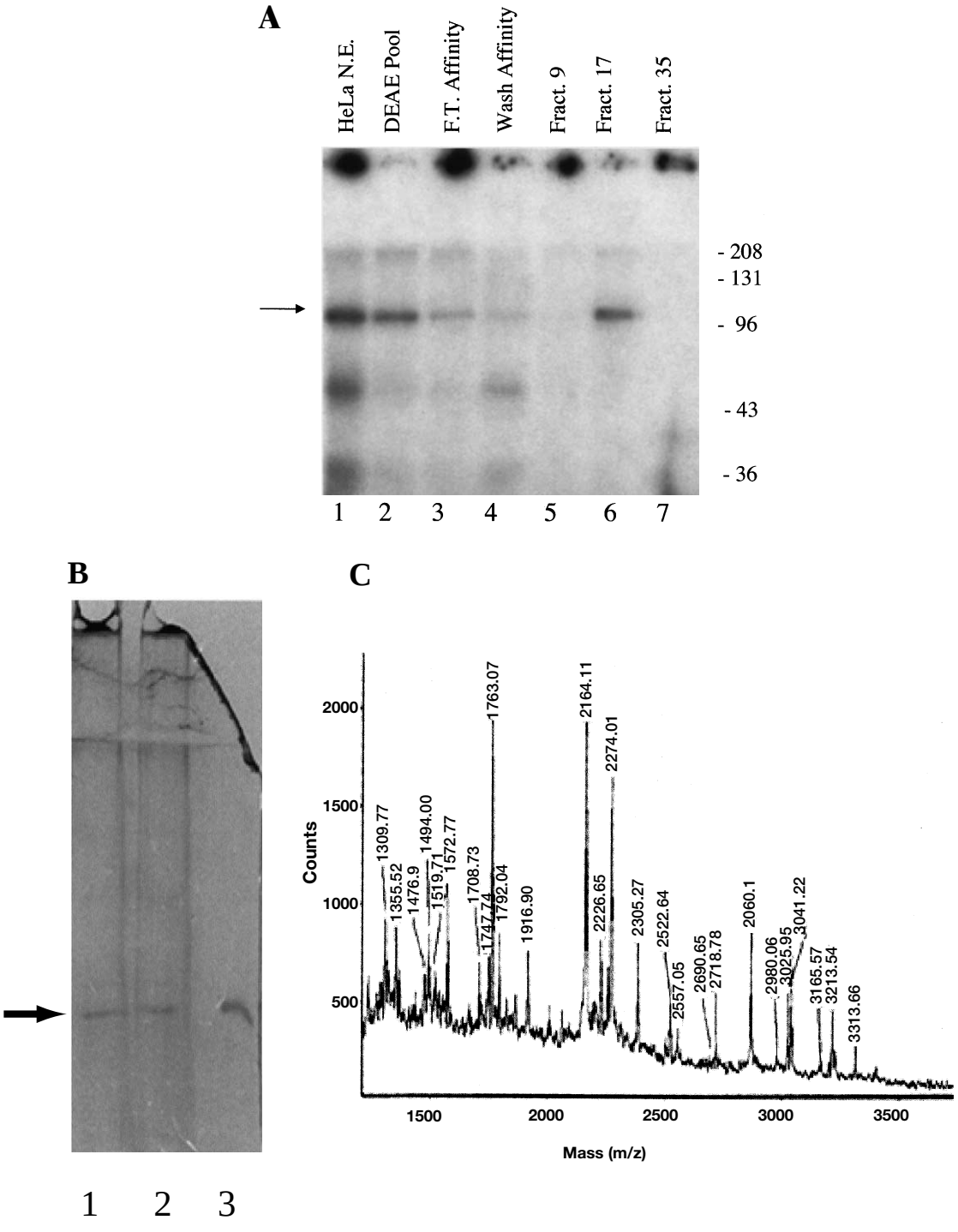

Fig. 1. Purification and identification of p97 as ZNF224. A: UV-induced cross-linking of protein–DNA complexes. DNA affinity-purified fractions were incubated with ^{32}P-labeled probe Neg1. After incubation, the reaction mixtures were irradiated with UV (see Section 2). The DNA–protein adducts were then resolved by SDS–PAGE and visualized by autoradiography. The positions of the molecular weight markers are indicated on the right. N.E. = nuclear extracts; F.T. = flow-through. The arrow indicates the specific protein–DNA complex. B: SDS–PAGE analysis. The arrow indicates the affinity-purified p97 visualized by staining with Coomassie. C: MALDI mass spectral analysis of the purified p97. D: Database analysis. On the left is indicated the probability score of unknown proteins identified as the best candidates by the Mascot and ProFound programs. On the right are the identified proteins and their relative molecular mass.

ligated. Neg1 oligonucleotide was added to cyanogen bromide-activated Sepharose 4B (Pharmacia Biotech). The resin was prepared as described by Kadonaga and Tjian [20], and applied on a 2 ml affinity chromatography column. Aliquots of fractions showing AldA-NRE binding activity, eluted by a linear gradient of KCl (0.1–1 M), were loaded onto SDS–PAGE gel and visualized by silver staining.

2.3. UV-induced cross-linking

In UV cross-linking experiments, the ^{32}P-labeled Neg1 probe was challenged with affinity-purified p97 fractions using reaction conditions as in EMSA (see above). The reaction mixture was then exposed to a UV transilluminator for 10 min. After electrophoresis on an 8% SDS–polyacrylamide gel, the protein–DNA complexes were identified by autoradiography. The molecular masses of the proteins were estimated considering the contribution to the mass of the oligonucleotide DNA probe.

2.4. In situ digestion and MALDI analysis

Protein bands stained with Coomassie brilliant blue were excised from the gel and destained by repetitive washes with 0.1 M NH_4HCO_3 pH 7.5 and acetonitrile. Samples were reduced by incubation with 50 μl of 10 mM DTT in 0.1 M NH_4HCO_3 buffer pH 7.5 and carboxyamidomethylated with 50 μl of 55 mM iodoacetamide in the same buffer.

The alkylated samples underwent tryptic digestion at 37°C for 18 h. The resulting peptide mixtures were extracted from the gel by three separate washes using 50 μl of 50 mM NH_4HCO_3 pH 7.5, 50 μl of acetonitrile and finally 50 μl of 0.1% trifluoroacetic acid (TFA, Carlo Erba) in 50% acetonitrile. The three fractions were pooled together, dried down in a Speed-Vac centrifuge and dissolved in 20 μl of 0.2% TFA for mass spectral analysis. A linear Voyager DE MALDI-TOF mass spectrometer (Applied Biosystems) was used for mass spectrometric analyses of the tryptic peptide mixtures. The mass range was calibrated using bovine insulin (at *m/z* 5734.6) and a matrix peak (at *m/z* 379.1) as internal standards. About 1.0 μl of sample was applied to a sample slide and mixed with 1.0 μl of a 10 mg/ml α-cyano-4-hydroxycinnamic acid solution in acetonitrile/0.2% TFA (70:30, v/v) before air-drying.

Peptide mass values recorded in the MALDI spectra were used for database search conducted with the Mascot and the ProFound 4.10.5

D

Input Summary

Date & Time	: Thu Sep 7 10:49:37 2000
Database	: NCBInr (updated: 08/29/2000)
Taxonomic Category	: Homo Sapiens
Search for	: Single protein only
Protein Mass Range	: 75 - 90 kDa
Sample ID	:
Digest Chemistry	: Trypsin
Cys Modification	: Iodoacetamide
Max Missed Cut	: 4
Tolerance (Avg)	: 1.0 Da
Tolerance (Mono)	: 0.1 Da
Peptide Masses (MH+, average)	1235.52 1355.52 1476.90 1494.08 1519.71 1572.77 1708.73 743.20 1824.93 1864.14 2254.25 2522.64 2718.78 3025.95 041.22 3313.66 3400.53 4724.46 4975.14
Number of Peptides	: 19

Protein Candidates

Probability	Protein Description	MW (kDa)
1.0e+00	gi|7019591|ref|NP_037530.1| zinc finger protein 224	82
1.1e-04	gi|7243079|dbj|BAA92587.1| (AB037770) KIAA1349 protein [Homo sapiens]	87
2.3e-05	gi|5031779|ref|NP_005522.1| interferon, gamma-inducible protein 16	82
8.7e-06	gi|177960|gb|AAA58358.1| (M55602) amiloride-binding protein [Homo sapiens]	81
7.4e-06	gi|6433901|emb|CAA71414.2| (Y10388) Graf protein [Homo sapiens]	86
3.2e-06	gi|3417297|gb|AAC31673.1| (AC002310) Unknown gene product [Homo sapiens]	88
1.3e-06	gi|7706467|ref|NP_057350.1| GIOT-4 for gonadotropin inducible transcription repressor-4	75
1.0e-06	gi|4758632|ref|NP_004691.1| potassium voltage-gated channel, KQT-like subfamily, member 4	77
8.0e-07	gi|479805|pir||S35458 SNF2 protein homolog - human (fragment)	89
5.6e-07	gi|4455442|emb|CAB36862.1| (AL022067) dJ134E15.1 (Blimp-1) [Homo sapiens]	85

Fig. 1 (*Continued*).

software packages available on internet. Unknown proteins were identified from both the probability score associated with each putative candidate and the difference in the score between the first and the other protein candidates.

2.5. Plasmid construction and EMSA

Human ZNF224 cDNA was obtained by reverse transcription-polymerase chain reaction (RT-PCR) from total RNA of HeLa cells and was sequenced entirely. The cDNA was amplified using pfx platinum Taq (Life Technologies) and oligonucleotide primers were designed with attB1 or attB2 sites for the insertion into the Gateway vectors (Life Technologies) by homologous recombination. The PCR product was cloned into pDONR201 recipient vector and the resulting recombinant plasmid was used to transfer the gene sequence in *Escherichia coli* pDEST15 (glutathione *S*-transferase (GST) fusion). ZNF224 mutant constructs (ZNF224M1 and ZNF224M2) were generated using the QuikChange site-directed mutagenesis kit (Stratagene). The prokaryotic plasmids were used for overexpression of the fusion proteins in *E. coli* BL21-DE3 cells. Cells harboring the GST fusion plasmids were grown at 30°C up to an optical density, at 600 nm, of 1.0 in 1000 ml culture volume, and induced with 0.5 mM isopropyl-1-thio-β-D-galactopyranoside. The bacterial lysates and GST fusion protein extraction were obtained according to the Pharmacia Biotech instruction manual. Purified proteins (3 μg) were used in EMSA, carried out as previously described [5].

2.6. Antibodies and Western blot

Rabbit polyclonal antiserum anti-ZNF224 was raised against an N-terminal peptide (Y-15-I) designed on the amino acid sequence of region 169–182 of ZNF224 (Neo System groupe SNPE). Antiserum anti-ZNF224 was purified by using AminoLink immobilization kit (Pierce). GST mouse monoclonal antibody (Santa Cruz Biotechnology) and anti-ZNF224 were used in Western blot experiments to reveal the recombinant protein GST–ZNF224. Two Western immunoblotting experiments were carried out with the purified proteins (5 μg) and the resulting nitrocellulose membranes were probed with rabbit polyclonal antiserum anti-ZNF224 (1:1000 dilution) and GST mouse monoclonal antibody (1:5000 dilution), respectively.

2.7. Plasmid construction and transient transfection assay

In order to obtain the ZNF224 recombinant plasmid to use in transient transfection experiments in eukaryotic cells, we performed the standard procedure for the insertion of ZNF224 cDNA into the Gateway eukaryotic vectors by homologous recombination as described above (Life Technologies). The reporter chloramphenicol acetyltransferase (CAT) plasmid 5′HPNRE2x was constructed by cloning, upstream of the heavy chain ferritin basal promoter (5′HP), two copies of the core element of AldA-NRE (25-mer). Increasing amounts (0.5, 1 and 2 μg) of the eukaryotic expression vector, pDEST-26 ZNF224, and 2 μg of the reporter CAT plasmid, 5′HPNRE2x, were cotransfected in COS-7 cells using the Lipofectamine Plus kit (Life Technologies). After 48 h total extracts were produced and the CAT assay was carried out as previously described [5].

3. Results and discussion

3.1. Purification of ZNF224

We have previously found a negative regulatory element (AldA-NRE) in the human aldolase A promoter that is involved in the modulation of promoter transcriptional activity [3,5]. We have further identified a repressor protein (p97) that recognizes a specific AGAGAG motif and interacts with AldA-NRE in a cell cycle-dependent manner, thus regulating the transcriptional aldolase A promoter activity during the cell cycle [7]. We have demonstrated that the cell cycle-dependent p97/AldA-NRE interaction is modulated by protein kinase C-mediated phosphorylation of p97 [8].

In an attempt to elucidate the molecular mechanisms underlying the p97 binding activity and to define a model of p97-mediated transcriptional repression, we purified the repressor protein from HeLa nuclear extracts by ion exchange chromatography and DNA affinity chromatography as described in Section 2. The purified p97 protein had an apparent molecular mass of 82 kDa, as judged by silver staining on SDS–PAGE (data not shown). We performed UV cross-linking experiments to ascertain whether the 82 kDa protein was a candidate for the specific DNA binding activity. As shown in Fig. 1A, lane 6, the UV exposure of purified p97 and its target AldA-NRE resulted in the formation of a strong protein–DNA complex, with an adjusted molecular mass of approximately 110 kDa, that migrates as a single band in SDS–PAGE. The specific DNA–protein complex was present in the HeLa nuclear extract (Fig. 1A, lane 1), in the fractions eluted from DEAE at a concentration of 0.3 M KCl (lane 2), and in affinity fraction number 17 (lane 6). A very small amount of active protein was also present in the affinity column flow-through and wash (lanes 3 and 4); no band was detected in non-active fractions from the affinity column (lanes 5 and 7). These results demonstrate that the protein purified by DNA affinity chromatography indeed recognizes specifically the GA-rich motif present in AldA-NRE.

3.2. Identification of p97 as ZNF224

The protein fraction eluted from the affinity chromatography that specifically recognizes the AldA-NRE sequence was concentrated on a Centricon system and loaded onto an 8% SDS–PAGE gel. The Coomassie-stained gel showed a main band with an electrophoretic mobility corresponding to an apparent molecular mass of 82 kDa (Fig. 1B, lanes 1 and 2). The protein band was excised from the gel, further destained and digested in situ with trypsin. A protein-free gel slice (Fig. 1B, lane 3) was treated in the same way and used as control. The tryptic peptides extracted from the gel were directly subjected to MALDI mass spectral analysis and the resulting spectrum is shown in Fig. 1C. Fifteen peptide peaks, ranging in mass between *m*/*z* 1235.94 and *m*/*z* 3316.66, were selected from the mass spectra and used to search the non-redundant NCBI database using both the Mascot and the ProFound software packages.

Fig. 1D reports a summary of the computer output of the two searching procedures. In both cases, the query returned a highly significant match with ZNF224, whose predicted molecular mass is 82 kDa. The probability score of the second best candidate identified by the two programs was several orders of magnitude lower, i.e. random matching. Therefore, the protein specifically recognized by the AldA-NRE sequence is ZNF224.

The 15 mass signals used in the searching procedure accounted for a total of 34% of the entire ZNF224 sequence. The remaining peaks reported in the spectra were then mapped onto the anticipated sequence of ZNF224, thereby increasing the percentage of sequence coverage to 68% and confirming the identification.

3.3. Structural features of ZNF224

The ZNF224 protein sequence, deduced from the cDNA sequence submitted by Shannon et al. in GenBank (AF-187990), consists of a 707 amino acid open reading frame and contains 19 tandemly repeated C_2H_2-type zinc finger motifs, each of them separated by seven amino acid inter-finger space regions, often referred to as the H/C link (Fig. 2A). Alignment of the amino-terminal sequence of ZNF224 with the KRAB-A domain of several C_2H_2 zinc finger proteins

A

```
1    GAGTCCAAACATTCGGTGGAGTCGCGGACACTTCCGCTCGGGGACTGAGGTTGCTGCAGTTTTTCCGCGATAGTTTGGGGCAGTTCCGCGCGTTGCAGGC   100
101  CCTTCTGAATTTCCTGGACCTACGCATTGGATCCTCAAAGAACTGCTGAATACCACTAGAAACATACCTGTAACCAGAGACAGCTGATTATAGCTTTCTG   200
201  CAGCAAGGAAGCCCACGTACCAGGGGCTGTCTTGGCACAATTCTGCTTTCCCAGGAACTGCATCACTCAGGACTCTGCAAGTTTCCAGAAGTAAGAGGGA   300
301  AAATGACCACGTTCAAGGAGGCAATGACCTTCAAGGACGTGGCTGTGGTCTTCACTGAGGAAGAGCTGGGGCTGCTGGACCTTGCTCAGAGGAAGCTGTA   400
       M  T  T  F  K  E  A  M  T  F  K  D  V  A  V  V  F  T  E  E  E  L  G  L  L  D  L  A  Q  R  K  L  Y    33
401  TCGAGATGTGATGCTGGAGAACTTCAGGAACCTGCTCTCAGTGGGACATCAAGCATTCCACAGGGATACTTTCCACTTCCTAAGGGAAGAAAAGATTTGG   500
       R  D  V  M  L  E  N  F  R  N  L  L  S  V  G  H  Q  A  F  H  R  D  T  F  H  F  L  R  E  E  K  I  W    66
501  ATGATGAAGACAGCAATCCAAAGGGAAGGGAATTCAGGAGACAAGATCCAAACTGAGATGGAGACTGTTTCAGAAGCAGGAACACATCAAGAGTGGTCCT   600
       M  M  K  T  A  I  Q  R  E  G  N  S  G  D  K  I  Q  T  E  M  E  T  V  S  E  A  G  T  H  Q  E  W  S  F   100
601  TCCAGCAAATCTGGGAAAAAATTGCAAGTGATTTAACCAGGTCTCAAGACTTGGTGATAAATAGCTCTCAGTTCTCCAAAGAAGGTGATTTCCCCTGCCA   700
       Q  Q  I  W  E  K  I  A  S  D  L  T  R  S  Q  D  L  V  I  N  S  S  Q  F  S  K  E  G  D  F  P  C  Q   133
701  GACTGAGGCAGGACTATCTGTAATTCACACAAGACAGAAATCTTCCCAGGGCAATGGATATAAACCATCCTTCAGTGATGTCTCCCACTTTGATTTTCAT   800
       T  E  A  G  L  S  V  I  H  T  R  Q  K  S  S  Q  G  N  G  Y  K  P  S  F  S  D  V  S  H  F  D  F  H   166
801  CAACAATTACACTCAGGAGAGAAATCTCATACGTGTGATGAGTGTGGAAAGAACTTTTGTTACATCTCAGCCCTTCGTATTCATCAGAGAGTCCACATGG   900
       Q  Q  L  H  S  G  E  K  S  H  T  C  D  E  C  G  K  N  F  C  Y  I  S  A  L  R  I  H  Q  R  V  H  M  G   200
901  GAGAGAAATGCTATAAGTGTGACGTGTGTGGTAAGGAATTCAGTCAGAGTTCACATCTGCAAACTCATCAGAGAGTCCACACTGGAGAGAAACCGTTCAA  1000
       E  K  C  Y  K  C  D  V  C  G  K  E  F  S  Q  S  S  H  L  Q  T  H  Q  R  V  H  T  G  E  K  P  F  K   233
1001 ATGTGTGGAATGTGGGAAAGGCTTCAGTCGTAGATCAGCACTTAATGTTCATCACAAATTACACACAGGAGAGAAACCTTATAATTGTGAGGAATGCGGG  1100
       C  V  E  C  G  K  G  F  S  R  R  S  A  L  N  V  H  H  K  L  H  T  G  E  K  P  Y  N  C  E  E  C  G   266
1101 AAGGCCTTCATTCACGATTCCCAGCTTCAAGAACATCAGAGAATCCATACGGGGGAGAAGCCATTCAAATGTGATATATGTGGTAAGAGCTTCTGTGGTA  1200
       K  A  F  I  H  D  S  Q  L  Q  E  H  Q  R  I  H  T  G  E  K  P  F  K  C  D  I  C  G  K  S  F  C  G  R   300
1201 GATCAAGACTTAATAGGCATTCCATGGTTCACACGGCAGAGAAACCATTCCGATGTGATACGTGTGATAAGAGCTTTCGTCAGAGATCAGCACTTAATAG  1300
       S  R  L  N  R  H  S  M  V  H  T  A  E  K  P  F  R  C  D  T  C  D  K  S  F  R  Q  R  S  A  L  N  S   333
1301 TCATCGCATGATCCACACAGGAGAGAAACCATACAAATGTGAGGAGTGTGGAAAAGGCTTTATTTGTAGGCGAGATCTTTATACGCATCATATGGTCCAC  1400
       H  R  M  I  H  T  G  E  K  P  Y  K  C  E  E  C  G  K  G  F  I  C  R  R  D  L  Y  T  H  H  M  V  H   366
1401 ACGGGAGAAAAGCCATATAATTGTAAAGAGTGTGGGAAGAGCTTCAGATGGGCCTCGTGTCTTTTGAAACATCAGCGAGTCCACAGTGGAGAAAAACCAT  1500
       T  G  E  K  P  Y  N  C  K  E  C  G  K  S  F  R  W  A  S  C  L  L  K  H  Q  R  V  H  S  G  E  K  P  F   400
1501 TCAAATGTGAAGAATGTGGGAAAGGATTTTACACAAATTCACAATGCTATTCCCACCAGAGATCCCATAGTGGAGAAAAACCATACAAATGTGTGGAGTG  1600
       K  C  E  E  C  G  K  G  F  Y  T  N  S  Q  C  Y  S  H  Q  R  S  H  S  G  E  K  P  Y  K  C  V  E  C   433
1601 TGGGAAGGGCTACAAAAGGAGGTTGGATCTTGACTTTCACCAGCGCGTCCATACAGGAGAGAAACTGTATAATTGTAAGGAATGTGGGAAGAGCTTTAGT  1700
       G  K  G  Y  K  R  R  L  D  L  D  F  H  Q  R  V  H  T  G  E  K  L  Y  N  C  K  E  C  G  K  S  F  S   466
1701 CGGGCCCCATGTCTTTTGAAACATGAGAGACTCCACAGTGGAGAAAAACCATTCCAATGTGAAGAGTGTGGGAAGAGATTTACTCAAAATTCACATCTTC  1800
       R  A  P  C  L  L  K  H  E  R  L  H  S  G  E  K  P  F  Q  C  E  E  C  G  K  R  F  T  Q  N  S  H  L  H   500
1801 ATTCCCATCAGAGAGTTCACACTGGAGAAAAGCCATACAAATGTGAGAAGTGTGGAAAGGGCTACAATAGTAAGTTTAATCTTGATATGCACCAGAAGGT  1900
       S  H  Q  R  V  H  T  G  E  K  P  Y  K  C  E  K  C  G  K  G  Y  N  S  K  F  N  L  D  M  H  Q  K  V   533
1901 CCACACAGGAGAGAGACCATACAATTGTAAGGAATGTGGGAAGAGTTTTGGCTGGGCCTCGTGTCTTTTGAAACATCAGAGACTGCGCAGTGGGGAAAAA  2000
       H  T  G  E  R  P  Y  N  C  K  E  C  G  K  S  F  G  W  A  S  C  L  L  K  H  Q  R  L  R  S  G  E  K   566
2001 CCTTTCAAATGTGAAGAGTGTGGGAAAAGATTTACTCAGAATTCACAGCTTCATTCTCATCAAAGAGTGCACACTGGAGAAAAGCCATACAAATGTGATG  2100
       P  F  K  C  E  E  C  G  K  R  F  T  Q  N  S  Q  L  H  S  H  Q  R  V  H  T  G  E  K  P  Y  K  C  D  E   600
2101 AGTGTGGGAAGGGCTTCAGCTGGTCCTCAACTCGTCTGACCCATCAGAGACGCCACAGCAGAGAAACACCTCTCAAATGTGAGCAGCATGGGAAGAACAT  2200
       C  G  K  G  F  S  W  S  S  T  R  L  T  H  Q  R  R  H  S  R  E  T  P  L  K  C  E  Q  H  G  K  N  I   633
2201 TGTACAGAATTCATTCTCTAAAGTGCAAGAAAAAGTTCACAGTGTAGAAAAGCCATACAAATGTGAGGACTGTGGGAAGGGCTACAACAGGCGCTTGAAT  2300
       V  Q  N  S  F  S  K  V  Q  E  K  V  H  S  V  E  K  P  Y  K  C  E  D  C  G  K  G  Y  N  R  R  L  N   666
2301 CTTGATATGCATCAGAGGGTCCACATGGGAGAGAAAACATGGAAGTGTAGGGAGTGTGATATGTGCTTTAGTCAGGCCTCAAGCCTTCGACTTCATCAGA  2400
       L  D  M  H  Q  R  V  H  M  G  E  K  T  W  K  C  R  E  C  D  M  C  F  S  Q  A  S  S  L  R  L  H  Q  N   700
2401 ATGTTCATGTTGGAGAAAAACCTTAGTGATGTGATGGTGCAATAAAGTCTTCACTCAGTCTTCATG                                      2466
       V  H  V  G  E  K  P  *  *
```

B

KRAB-A

ZNF224	5	KEAMTFKDVAVV FTEEELGLLD LAQRK LYRDV MLENFRNLL SVG
GIOTi	1	MDAVTYEDVYV SFTQEEWALLN PSQKN LYKDVMMQ TCRNLIAIG
DBP	11	NAVTY YDVLV SFTQEEWALLD PSQK SLYKDVM LETY RNLT AIG
ZNF1 36	1	MDSVA FEDVDVN FTQEEWALLD PSQKNLY RDVM WETMRNLA SIG
ZNF14 1	1	MELLT FRDVA IEFSPP EEWKCLD PDQQNLY RDVML ENYR NLVSLG
ZNF1 77	12	NSVTFQEVAVDFSQEEWALLD PAQKN LYXDVML ENFRNLA SVG
Zim1	48	EPVIFKDVAV YFSQK EWQLLE PAQKD LYKDVML ENYENL ISVE
KS1	9	G PVSFKDVTVD FTQEEW QRLD PAQKALY RDVML ENYCH FIS VG
Kid-1	21	VSVTFEDVAVL FTRDEWKK LDL SQRSLYREVML ENY SNLA SMA
Kzf-1	3	MEQLT FRDVAVD FSP DEWECLD PPQQ SLY RDVMVENY RNLV SVA
consensus		**VTF DVAV F EEW LD QR LYR VMLENY NLVS G**

Fig. 2. Structure of ZNF224 cDNA. A: Nucleotide and predicted amino acid sequences of ZNF224. The 19 C_2H_2-type zinc finger motifs are underlined, the conserved KRAB-A domain is in bold print. B: Amino acid alignment of the KRAB-A domain of ZNF224 with that of other zinc finger proteins.

showed that several residues within the N-terminal 42 amino acids of ZNF224 are fully conserved and that these residues represent a typical consensus domain of the KRAB-A subfamily of Kruppel-type zinc finger proteins (Fig. 2B). In the human genome, about 700 different genes encode C_2H_2 zinc finger proteins and one-third of these contain a KRAB domain that is subdivided into box A and box B domains. Box A in the KRAB domain, containing 45 amino acids, is usually a potent repression module [13,14]. The features of ZNF224 strongly suggest that it functions as a DNA binding protein endowed with transcriptional regulatory properties.

3.4. ZNF224 is the AldA-NRE binding protein

To examine whether ZNF224 protein interacts directly with AldA-NRE, we subjected the ZNF224 recombinant protein to gel shift assay. We cloned the full-length ZNF224 cDNA and two deletion mutants, M1 and M2, which have eight and five zinc fingers, respectively (Fig. 3A) in the prokaryotic expression vector pDEST15 (Gateway cloning system). The fusion proteins (ZNF224WT, ZNF224M1 and ZNF224M2) were then analyzed by Western blot carried out using an anti-ZNF polyclonal antibody (Fig. 3B) and anti-GST antibody (Fig. 3C). Both antibodies specifically recognize ZNF224WT and the deleted M1 and M2 proteins.

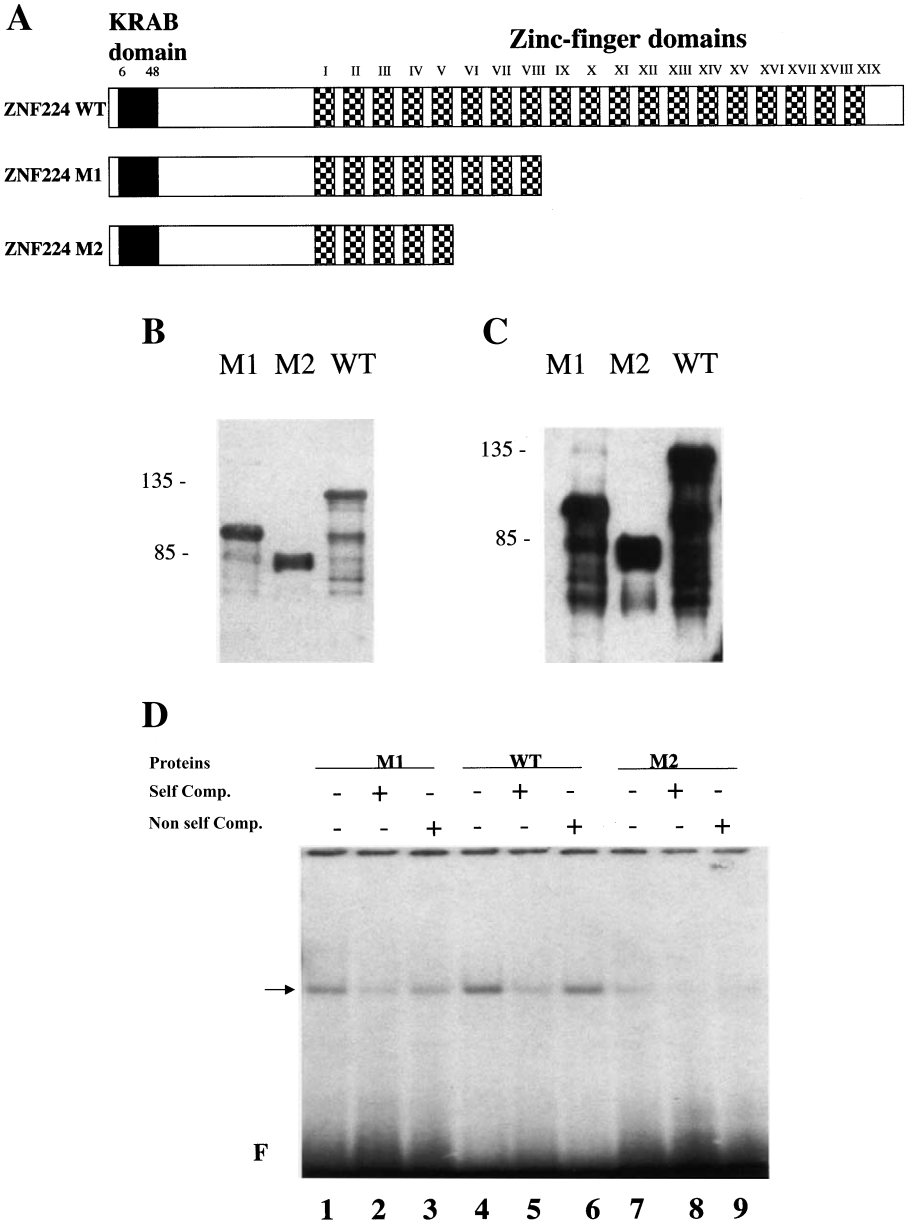

Fig. 3. ZNF224 recombinant protein analysis. A: Schematic representation of the structure of the full-length protein (WT) and of two deletion mutants, M1 and M2, which have only eight and five zinc fingers, respectively, in the C-terminal region. The purified recombinant proteins (M1, M2, WT) were analyzed by Western blot experiments with ZNF224 antibody (B) and with GST antibody (C). Molecular weight markers are indicated on the left. D: Binding of ZNF224 to AldA-NRE. Purified proteins M1, WT and M2 were incubated with Neg1 ^{32}P-labeled probe (lanes 1, 4 and 7, respectively). Homologous competition was performed with a 100-fold molar excess of Neg1 (lanes 2, 5 and 8) and unrelated unlabeled oligonucleotides (lanes 3, 6 and 9). The arrow indicates the specific DNA–protein complex; F indicates the free DNA probe.

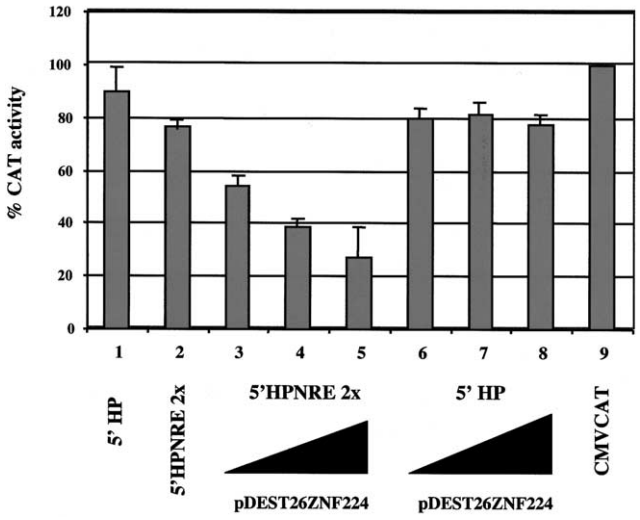

Fig. 4. Transcriptional repression mediated by ZNF224 protein. Lane 1 shows the CAT activity of the heavy chain ferritin promoter. In lane 2 is shown the effect of the twice repeated core element (AldA-NRE) on the transcriptional activity of reporter CAT. In lanes 3–5 increasing amounts of pDEST26ZNF224 (0.5, 1, 2 μg) were cotransfected in COS cells in the presence of the reporter CAT p5′HPNRE2x. In lanes 6–8 increasing amount of pDEST26ZNF224 (0.5, 1, 2 μg) were cotransfected in COS cells in the presence of the reporter CAT 5′HP. CMVCAT was used as a positive control of transcriptional CAT efficiency (100%) (lane 6). The results are presented as an average of three experiments.

In EMSA experiments with the bacterial expressed ZNF224WT protein, we demonstrate that ZNF224 specifically binds to Neg1 oligonucleotide (Fig. 3D, lane 4). In fact, the ZNF224/Neg1 complex can be competed by adding an excess amount of the unlabeled Neg1 oligonucleotide (Fig. 3D, lane 5), but not by non-specific DNA (Fig. 3D, lane 6). This is the first demonstration that the ZNF224 protein, whose function was hitherto unknown, is indeed the DNA binding protein that specifically interacts with the negative regulatory element AldA-NRE.

To test the hypothesis that ZNF224 binds to DNA via its multiple zinc finger domains we examined the M1 and M2 recombinant proteins in EMSA experiments. The progressive deletions of zinc fingers in the carboxy-terminal region resulted in a decreased protein amount of the DNA–protein complex. In fact, as shown in Fig. 3D, lane 1, the deletion of 11 zinc fingers (M1 recombinant protein) leads to a decreased binding activity, and the deletion of 14 zinc fingers (M2 recombinant protein) results in a very weak binding to DNA (Fig. 3D, lane 7). These results confirm that the C_2H_2 motifs, which generally confer sequence-specific DNA binding activity to the zinc finger proteins, are involved in the ZNF224–AldA-NRE interaction. Moreover, we suggest that the binding specificity of ZNF224 to AldA-NRE is dependent on the presence of several zinc finger domains.

3.5. ZNF224 functions as a transcriptional repressor

The zinc finger DNA binding proteins harboring a KRAB domain usually show a powerful ability to repress gene transcription. To test the ZNF224-mediated transcriptional repression, we cotransfected COS cells with a reporter CAT plasmid (5′HPNRE2x) containing two core elements of AldA-NRE upstream of a heterologous basal promoter and the recombinant plasmid encoding ZNF224 (pDEST26ZNF224). As shown in Fig. 4, the expressed ZNF224 protein inhibits the reporter promoter activity in a dose-dependent manner (lanes 3–5). The plasmid 5′HPNRE2x transfected alone (lane 2) exhibits a basal CAT activity due to a moderate effect of AldA-NRE on transcription of heterologous basal promoter (lane 1). These results demonstrate that ZNF224 is able to specifically repress AldA-NRE-mediated transcription of a heterologous promoter. The dose-dependent decrease in expression of the CAT reporter gene suggests the specific role of ZNF224 in binding AldA-NRE motifs and in causing the negative modulation of reporter CAT transcription.

In this work, we have purified a 82 kDa nuclear protein from HeLa cells and have identified it as the C_2H_2 zinc finger DNA binding protein, ZNF224. The obtained results are the first demonstration that ZNF224 functions as a DNA binding protein recognizing specifically the AldA-NRE sequences and represses AldA-NRE-mediated transcription of a heterologous promoter. These findings strongly suggest that ZNF224 is involved in the control of transcription through a general mechanism of repression. Northern blot analysis suggests that ZNF224 cDNA in several human and mouse tissues is ubiquitously expressed, thus confirming the general function of ZNF224 in repressing transcription in mammals (data not shown). Moreover, we have demonstrated that AldA-NRE is the DNA module which could specifically target AldA-NRE-mediated repression also to other genes. Isolation of ZNF224 target genes will further help to elucidate the biological functions of this transcriptional factor in various cellular and tissue contexts during the cell cycle and development.

Acknowledgements: We are grateful to Jean Ann Gilder (Scientific Communication) for editing the text. This work was supported by Grants from MIUR (PRIN '99), from CNR (Target project Biotechnology), and from BIOGEM, SCARL.

References

[1] Izzo, P., Costanzo, P., Lupo, A., Rippa, E. and Salvatore, F. (1989) FEBS Lett. 257, 75–80.
[2] Stauffer, J.K., Colbert, M.C. and Ciejek-Baez, E. (1990) J. Biol. Chem. 265, 11773–11782.
[3] Costanzo, P., Lupo, A., Rippa, E., Grosso, M., Salvatore, F. and Izzo, P. (1993) Biochem. Biophys. Res. Commun. 195, 935–944.
[4] Salminen, M., Maire, P., Concordet, J.-P., Moch, C., Porteu, A., Khan, A. and Daegelen, D. (1994) Mol. Cell. Biol. 14, 6797–6800.
[5] Lupo, A., Costanzo, P., Medugno, L., Salvatore, F. and Izzo, P. (1995) Biochem. Biophys. Res. Commun. 216, 69–77.
[6] Lupo, A., Costanzo, P., De Rosa, M., Russo, T., Salvatore, F. and Izzo, P. (1994) Exp. Cell Res. 213, 359–364.
[7] Lupo, A., Costanzo, P., Medugno, L., Romeo, I., Salvatore, F. and Izzo, P. (1997) J. Biol. Chem. 272, 31641–31647.
[8] Costanzo, P., Lupo, A., Medugno, L., D'Agostino, P., Zevino, C. and Izzo, P. (1999) FEBS Lett. 454, 61–66.
[9] Postigo, A. and Douglas, C.D. (1997) EMBO J. 16, 3935–3943.
[10] Palm, K., Belluardo, N., Metsis, M. and Timmusk, T. (1998) J. Neurosci. 18, 1280–1296.
[11] Coqueret, O., Berubé, G. and Nepveu, A. (1998) EMBO J. 17, 4680–4694.
[12] Turner, J. and Crossley, M. (1999) Trends Biochem. Sci. 24, 236–240.
[13] Philipsen, S. and Suske, G. (1999) Nucleic Acids Res. 27, 2991–3000.
[14] Klug, A. and Schwabe, J.W.R. (1995) FASEB J. 9, 597–604.

[15] Witzgall, R., O'Leary, E., Leaf, A., Onaldi, D. and Bonventre, J.V. (1994) Proc. Natl. Acad. Sci. USA 91, 4514–4518.
[16] Margolin, J.F., Friedman, J.R., Meyer, K.H.K., Vissing, H., Thiesen, H. and Rauscher III, F. (1994) Proc. Natl. Acad. Sci. USA 91, 4509–4513.
[17] Agata, Y., Matsuda, E. and Shimizu, A. (1999) J. Biol. Chem. 274, 16412–16422.
[18] Schultz, D., Friedman, J. and Rauscher III, F. (2001) Genes Dev. 15, 428–443.
[19] Dignam, J., Lebovitz, R.M. and Roeder, R.G. (1983) Nucleic Acids Res. 11, 1475–1489.
[20] Kadonaga, J.T. and Tijan, R. (1986) Proc. Natl. Acad. Sci. USA 83, 5889–5893.