

A chicken embryo protein related to the mammalian DEAD box protein p68 is tightly associated with the highly purified protein–RNA complex of 5-MeC-DNA glycosylase

Jean-Pierre Jost*, Steffen Schwarz, Daniel Hess, Herbert Angliker, Frances V. Fuller-Pace¹, Hans Stahl², Stéphane Thiry and Michel Siegmann

Friedrich Miescher-Institute, PO Box 2543, CH-4002 Basel, Switzerland, ¹Department of Molecular and Cellular Pathology, University of Dundee, Ninewells Medical School, Dundee DD1 9SY, UK and ²Medizinische Biochemie und Molekularbiologie, Universität des Saarlandes, D-66421 Homburg/Saar, Germany

Received June 4, 1999; Revised and Accepted July 6, 1999

DDBJ/EMBL/GenBank accession no. AF158370

ABSTRACT

We have shown previously that DNA demethylation by chick embryo 5-methylcytosine (5-MeC)-DNA glycosylase needs both protein and RNA. Amino acid sequences of nine peptides derived from a highly purified 5-MeC-DNA glycosylase complex were identified by Nano-electrospray ionisation mass spectrometry to be identical to the mammalian nuclear DEAD box protein p68 RNA helicase. Antibodies directed against human p68 helicase cross-reacted with the purified 5-MeC-DNA glycosylase complex and immunoprecipitated the glycosylase activity. A 2690 bp cDNA coding for the chicken homologue of mammalian p68 was isolated and sequenced. Its derived amino acid sequence is almost identical to the human p68 DEAD box protein up to amino acid position 473 (from a total of 595). This sequence contains all the essential conserved motifs from the DEAD box proteins which are the ATPase, RNA unwinding and RNA binding motifs. The rest of the 122 amino acids in the C-terminal region rather diverge from the human p68 RNA helicase sequence. The recombinant chicken DEAD box protein expressed in *Escherichia coli* cross-reacts with the same p68 antibodies as the purified chicken embryo 5-MeC-DNA glycosylase complex. The recombinant protein has an RNA-dependent ATPase and an ATP-dependent helicase activity. However, in the presence or absence of RNA the recombinant protein had no 5-MeC-DNA glycosylase activity. *In situ* hybridisation of 5 day-old chicken embryos with antisense probes of the chicken DEAD box protein shows a high abundance of its transcripts in differentiating embryonic tissues.

INTRODUCTION

DNA methylation and demethylation reactions are important for the establishment of specific DNA methylation patterns required for normal embryonic development and imprinting (1–8). DNA can be either methylated *de novo* or the methylation pattern is maintained through cell cycles by the maintenance DNA methyltransferase (1,2 and references therein). The demodification of DNA can be either passive or active. Passive demethylation is obtained by an inhibition of the DNA maintenance methylation during replication (9–11), whereas active DNA demethylation reactions either remove the methyl group from the 5 position of methylcytosine (12) or it replaces 5-methylcytosine (5-MeC) by cytosine (13–17). Demethylation can also be obtained by a combination of both passive and active mechanisms. In this case, the product of the passive reaction is the formation of a hemimethylated DNA which becomes the substrate for either nucleotide excision or for the 5-MeC-DNA glycosylase (9). Active DNA demethylation has been observed in different systems (12–19). 5-MeC-DNA glycosylase activity has been measured in developing chicken and mouse embryos (17,20) in mouse embryonic stem cells (J.-P.Jost, unpublished results) and in G8 mouse myoblasts (21). In these systems, 5-MeC-DNA glycosylase used preferentially hemimethylated DNA as a substrate (20) and the reaction is specific for 5-MeCpGs (17). Recently we have shown that 5-MeC-DNA glycosylase reactions in developing chicken embryos (22,23) and in G8 mouse myoblasts (24) require both RNA and protein. RNA rich in CpGs is present in the highly-purified 5-MeC-DNA glycosylase from chicken embryos and it was shown that this RNA serves as a guide for the demethylation complex to hemimethylated sites (23). In the present work we show that one of the protein components of the DNA demethylation complex is related to the mammalian DEAD box protein p68.

*To whom correspondence should be addressed. Tel: +41 61 697 6688; Fax: +41 61 721 4091; Email: jost@fmi.ch

MATERIALS AND METHODS

Purification of 5-MeC-DNA glycosylase from 12-day-old chicken embryos

Four kilos of 12-day-old chicken embryos (1200 embryos) or 1 kilo of chicken embryo heads were processed as previously described (20) with the following modifications. Cell centrifugations and chromatographic separations were carried out at 0–4°C. The active fractions obtained after CM-Sepharose FPLC were dialysed against 1.7 M $(\text{NH}_4)_2\text{SO}_4$, 20 mM HEPES, pH 7.5, 1 mM dithiothreitol (DTT). Upon sedimentation of the insoluble proteins, the sample was applied onto a butyl-Sepharose FPLC column (10–15 mg protein per ml column). Elution proceeded by steps with decreasing concentrations of $(\text{NH}_4)_2\text{SO}_4$ in the same buffer. The enzyme eluted between 0.9 and 0.6 M $(\text{NH}_4)_2\text{SO}_4$. Active fractions were precipitated with a final concentration of 3.5 M $(\text{NH}_4)_2\text{SO}_4$. Active fractions were further separated on DNA-Dynabeads as previously described (20) and the active fractions were further purified on a 100 μl Mono S column (Smart system, Pharmacia). Between 0.5 and 1 mg protein was separated per run. The column was equilibrated in 20 mM HEPES, 10 mM EDTA, 10% glycerol, 1 mM DTT and 0.5 mM PMSF. The column was loaded and washed consecutively with the starting buffer and 100 mM NaCl (in the same buffer). Elution of proteins was carried out in steps with increasing concentrations of ammonium carbonate. 5-MeC-DNA glycosylase eluted between 0.4 and 0.5 M ammonium carbonate. Active fractions were lyophilised, and separated on a preparative SDS–polyacrylamide gel (10%) as previously described (20). Four kilos of embryos yielded a maximum of 5 μg of the highly purified enzyme (purification of >100 000-fold) that was used for microsequencing. 5-MeC-DNA glycosylase was assayed with a radioactively-labelled hemimethylated substrate as previously described (20).

In-gel digestion

The SDS–PAGE separated proteins were excised from the gel, reduced with DTT, alkylated with iodoacetamide and cleaved with trypsin (Promega, sequencing grade) as described by Shevchenko *et al.* (25).

NanoESI mass spectrometry

The extracted tryptic peptides were desalted with 5% formic acid, 5% methanol in H_2O on a 1 μl Poros P20 column and concentrated to 1 μl with 5% formic acid, 50% methanol in H_2O directly into the Nanoelectrospray ionization (NanoESI) needle. NanoESI mass spectrometry (MS) was performed according to the published method of Wilm *et al.* (26). The mass spectra were obtained on an API 300 mass spectrometer (PE Sciex, Toronto, Ontario, Canada) equipped with a NanoESI source (Protana, Odense, Denmark).

LC-ESIMS and N-terminal sequence analysis

Fractionation of the peptides by LC-ESIMS was performed as described by Krieg *et al.* (27). N-terminal sequences analysis was carried out on a Model 477A protein sequencer (Applied Biosystems, Inc., Foster City, CA) according to the recommendations of the manufacturer.

Immunoblots with antibodies directed against human p68 helicase and SV-40 T antigen

Purified 5-MeC-DNA glycosylase (1–2 μg of post heparin-Sepharose or butyl-Sepharose) was separated on an SDS–polyacrylamide gel (10%). Upon electrotransfer of the proteins to Immobilon membranes (Pharmacia) the immunoreaction was carried out as described by Harlow and Lane (28). The second antibody was linked to alkaline phosphatase. The primary antibodies were incubated overnight at 4°C at the following dilutions: polyclonal antibodies A2 (1/100) B1 (1/200), 2907 (1/1000 and 1/500) monoclonal antibodies: PAb204 (1/3), C10 (1/1000), SV-40 T antigen Ab-2 (1/1000 and 1/500, the epitope is in the N-terminal region, Oncogene Research Product). The known peptide sequences reacting with the p68 antibodies are underlined in Figure 1.

Immunoprecipitation of the RNA–protein complex of 5-MeC-DNA glycosylase

Indirect immunoprecipitation with anti-p68 antibodies (B1) or anti-SV-40 T antigen Ab-2 antibodies (Oncogene Research Product) was carried out essentially as described by Liu *et al.* (29) with the following modifications. The incubation of the purified enzyme (5 μg of post butyl-Sepharose fraction) with the antibody was for 16 h on ice and the incubation with 50 mg protein A-Sepharose in a total volume of 100 μl was for 2–4 h on ice. The incubation was carried out in 0.15 M NaCl, 20 mM HEPES pH 7.5, 10 mM EDTA and 100 μg of ribonuclease-free BSA. The parallel controls were incubated without antibodies or with non-immune serum. The washing of the immunocomplexes was carried out with 0.15 M NaCl in the same buffer. 5-MeC-DNA glycosylase activity present in the sediments was measured directly with the bead suspension.

Since the polyclonal and monoclonal antibodies had a high activity of deoxyribonucleases and ribonucleases, they had to be first purified either on protein A-Sepharose as described by Harlow and Lane (28) or by microchromatography on carboxymethyl-Sepharose and DEAE-Sepharose (30). Direct immunoprecipitation with SV-40 T antigen antibodies Ab-2 linked to agarose beads was carried out by incubating 5–10 μg of purified 5-MeC-DNA glycosylase (post butyl-Sepharose fraction) with 50 mg of immunoglobulin beads in a total volume of 70–100 μl . The buffer contained 20 mM HEPES pH 7.5, 20 mM EDTA, 2 mM Pefabloc (Boehringer), 100 mM NaCl, 10% glycerol and 100 μg of ribonuclease-free BSA. Control beads were prepared by heating the antibody–agarose beads at 65°C for 30 min with 100 mM mercaptoethanol. Upon extensive washing with 100 mM mercaptoethanol, the beads were equilibrated in a vast excess of incubation buffer without BSA. The reaction was then assembled as for the positive test. Incubation was carried out on ice for 16 h. Beads were washed consecutively with cold PBS and 0.5 M NaCl in 20 mM HEPES pH 7.5, 10 mM EDTA. Upon exchange of the high salt buffer for the assay buffer, 5-MeC-DNA glycosylase activity was measured as above.

Construction and screening of a cDNA library

Total RNA from chicken embryos (5 or 12 days old) was isolated by the procedure of Chomczynski and Sacchi (31) and mRNA was separated on oligo dT-Dynabeads as recommended by the manufacturer. cDNA was synthesised and

cloned into the *EcoRI*–*XhoI* site of the Uni-ZAP, XR, λ insertion vector (cDNA synthesis Kit, ZAP-cDNA synthesis Kit and ZAP-cDNA Gigapack III gold cloning Kit from Statagene).

The screening of the cDNA library was carried out by conventional procedures, using synthetic, labelled oligonucleotides coding for the peptides sequences obtained from the highly purified 5-MeC-DNA glycosylase. The following antisense DNA sequences were used: 5'-GGCGATGGGCCTATTTGTTTGGTGCTGGCACCAACTCGG-3', 5'-GATGAAGCAGATAG-AATGCTTGATATGGGCTTTGAACCCCAA-3', 5'-AGCTCCTATTCTGATTGCTACAGATGTGGCCTCCAG-3' and 5'-GTTGCTTCAGTTGGTTCGAAGACAGAGGTTCA-3'. The same sequences were used for northern blots.

Positive clones were characterised by restriction mapping and DNA sequencing. DNA sequencing was carried out with an AB1 Prism 377 DNA sequencer from Perkin Elmer.

Production of recombinant chicken p68 helicase

The 2.7 kb *BamHI*–*XhoI* cDNA from positive λ clone was inserted in-frame at the *BamHI*–*SalI* site into the directional vector pQE-30 (Qiagen). *Escherichia coli* XL1-blue was transfected with the construct and selected on tetracycline–ampicillin plates. The presence of an insert was tested with restriction enzymes from single colony lysates. Recombinant protein was induced for 4 h with 2 mM IPTG. Cells were collected and broken with a French pressure cell. The majority of the recombinant protein was present in the soluble fraction. Recombinant protein was further fractionated on a Ni–NTA–agarose column (Qiagen). Recombinant protein was characterised by immunoblotting and MS.

Enzymatic test

5-MeC-DNA glycosylase was tested with a labelled hemimethylated oligonucleotide as previously described (24). RNA helicase and ATPse were tested as described by Fuller-Pace *et al.* (32) and Scheffner *et al.* (33). Single-stranded RNAs were prepared by run-off transcription of appropriate linearised pGEM-3 (Promega) clones. For description of used clones and experimental details see Scheffner *et al.* (33). RNA helicase activity was assayed in 40 μ l volumes containing 0.15 ng radioactively labelled RNA substrate in a reaction buffer composed of 30 mM Tris–HCl, pH 7.5, 75 mM NaCl, 5 mM MgCl₂, 4 mM ATP, 1.5 mM DTT, 30 mg/ml BSA and 0.5 U/ml RNasin. After incubation for 60 min at 37°C, reactions were stopped by addition of 0.1 vol of 3% SDS, 150 mM EDTA.

In situ hybridisation (ISH) on tissue sections

The probe used for ISH was part of the cloned chicken DEAD box protein covering the internal sequence from nucleotide position 88 to 1203 (*AvaI* and *EcoRI* restriction sites). The sequence was introduced into the vector pGEM-4Z (Promega) containing a T7 and SP6 promoter sequence flanking the polylinker. DIG-labelled RNA was obtained by using the SP6/T7 *in vitro* transcription kit (Boehringer Mannheim). ISH was essentially performed as described by Schaeren-Wiemers and Gerfin-Moser (34). The concentration of the probe was 400 ng/ml hybridisation mixture. The hybridisation was done overnight at 60°C. The cover slides were removed from the sections in 5 \times SSC at 68°C and slides were washed for 1 h at 68°C in 0.2 \times SSC. Immunological detection of DIG-labelled hybrids was performed

as recommended by the manufacturer (Boehringer Mannheim). Nuclear staining with Hoechst 33258 (DAPI) was performed on all slides.

Chemicals and enzymes

Heparin-Sepharose, CM-Sepharose, DEAE-Sepharose, Butyl-Sepharose and Mono-S were purchased from Pharmacia. Benzamidine was from Fluka AG (Buchs/SG, Switzerland). PMSF, Pefabloc, proteinase K, terminal transferase, biotin-dUTP, streptavidin and T4 DNA ligase were obtained from Boehringer Mannheim, whereas polynucleotide kinase, all restriction enzymes and ribonuclease-free BSA were purchased from Biofinex (Praroman, CH-1724, Switzerland). Streptavidin-Dynabeads (Qiagen) was from Milan Analytica AG, CH-1634 La Roche, Switzerland.

[α -³²P]dATP and [γ -³²P]ATP triethylammomium (3000 Ci/mmol) were purchased from Amersham. Some oligonucleotides DNA and RNA were synthesised by Microsynth, Balgach, CH-9436 Switzerland.

RESULTS

Presence of a DEAD box protein in the highly purified 5-MeC-DNA glycosylase

5-MeC-DNA glycosylase was purified >100 000-fold from either total embryos or from the embryonic brain and eyes, which have the highest specific activity of the enzyme. Nine peptides derived from the gel-purified 5-MeC-DNA glycosylase were identical to the human DEAD box protein p68 helicase (see sequences in bold italics of Fig. 1). The nine peptides were obtained from two separate 5-MeC-DNA glycosylase preparations, one from total 12 day-old embryos and the other one only from their brain and eyes. We have isolated the cDNA of the corresponding protein from a chicken embryonic cDNA library and a comparison of its deduced amino acid sequence is shown in Figure 1. In comparison to human p68 (35) the first 12 amino acids at the N-terminus are missing, but a close identity up to amino acid position 473 is observed (Fig. 1). The chicken protein contains all the essential conserved amino acid motifs of the DEAD box protein family, which are (36): I, AQTGSGKT (ATPase motif); Ia, APTRELA; Ib, ATPGRL; II, VLDEAD; III, SAT (RNA unwinding); IV, FVET; V, VASRGLD; and VI, YIHRIGRTAR (RNA binding). These motifs are boxed in Figure 1. In contrast, the sequence of the C-terminal region is less well conserved.

P68 is closely related to the human DEAD box protein p72 and we note that the amino acid sequence of both proteins also begins to diverge within this region (37). The cDNA sequences cloned were up to 3000 bp in length. A northern blot carried out with brain mRNA gave a major band at 4 kb and a minor band at 3800 bp (result not shown). Similar results were obtained for the mammalian p68 RNA helicase (38).

Purified 5-MeC-DNA glycosylase cross-reacts with antibodies directed against human p68 RNA helicase and SV-40 T antigen

Since the chicken DEAD box protein related to human p68 copurifies with CpG-rich RNA and the 5-MeC-DNA glycosylase complex (up to the preparative SDS–PAGE), it was of interest to first see whether a purified fraction of 5-MeC-DNA glycosylase

| | | |
|-----------|--|-----|
| chicken | MPGFGAPRFGGSRAGPLSGKKFGNPGKLTKKKWNLDLDPKFEKN | 45 |
| human p68 | MSGYSSDRDRGRDRGFGAPRFGGSRAGPLSGKKFGNPGKLVKKKWNLDLDPKFEKN | 57 |
| chicken | FYQEHDPVVRRTAQEVEQYRSSKEVTVRGHNCPKPIINFYEANFPANVMEVIQRQNF | 102 |
| human p68 | FYQEHFDLARRTAQEVEYRRSKEITVRGHNCPPVLFNFYEANFPANVMDVIARQNF | 114 |
| chicken | TEPTAIQAQGWVVALSGLDMVGV <u>AQTGSGKT</u> LSYLLPAIVHINHQPFLERGDGPICL | 159 |
| human p68 | TEPTAIQAQGWVVALSGLDMVGV <u>AQTGSGKT</u> LSYLLPAIVHINHQPFLERGDGPICL | 171 |
| chicken | <u>VLAPTRELAQQVQVAAEYSR</u> ACRLKSTCIYGGAPKGPQIRDLEGRVEICIA <u>ATPGRL</u> | 216 |
| human p68 | <u>VLAPTRELAQQVQVAAEYCRACRLKSTCIYGGAPKGPQIRDLEGRVEICIA</u> <u>ATPGRL</u> | 228 |
| chicken | <u>IDFLEAGK</u> TNLRRTYLVL <u>DEAD</u> <u>RMLDMGFEPQIR</u> KIVDQIRPDRQTLMW <u>SAT</u> WPKE | 273 |
| human p68 | IDFLECGKTNLRRTYLVL <u>DEAD</u> <u>RMLDMGFEPQIR</u> KIVDQIRPDRQTLMW <u>SAT</u> WPKE | 285 |
| chicken | VRQLAEDFLKEYVHINIGALELSANHNILQIVDVCHDVEKDDKLIIRLMEEIMSEKEN | 330 |
| human p68 | VRQLAEDFLKDYIHINIGALELSANHNILQIVDVCHDVEKDEKLIIRLMEEIMSEKEN | 342 |
| chicken | KTIV <u>FVET</u> KRRCDLTKMRRDGPAMPAGIHGDKSQQERDVLNEFKHGK <u>APILLIATD</u> | 387 |
| human p68 | KTIV <u>FVET</u> KRRCDLTKMRRDGPAMPAGIHGDKSQQERDVLNEFKHGK <u>APILLIATD</u> | 399 |
| chicken | <u>VASRGLD</u> VEDVKFVINYDYPNSSED <u>YIHRIGRTAR</u> STKTGTAYTFFTPNNIK <u>QVNDL</u> | 444 |
| human p68 | <u>VASRGLD</u> VEDVKFVINYDYPNSSED <u>YIHRIGRTAR</u> STKTGTAYTFFTPNNIK <u>QVSDL</u> | 456 |
| chicken | <u>ISVLR</u> EANQAINPKLL <u>QLIEDR</u> GSGRSRG...DRRDYSAGKRGGFSSFREREN | 495 |
| human p68 | <u>ISVLR</u> EANQAINPKLL <u>QLVEDR</u> GSGRSRGRGGMKDDRRDYSAGKRGGFNTFRDREN | 513 |
| chicken | FERTYGALGKRDFGAKAQNGAYSTQSFNSGTPFGNGFAAAGMQAGFRAGNPAGAYQN | 552 |
| human p68 | YDRGYSSLLKRDFGAKTQNGVYSAANYTNGS.FGSNFVSAGLQTSFRTGNPTGTQYN | 569 |
| chicken | GYD..QQYGSNIANMHNGMNQQYAYPATGAAPMIGYPMPASYSQ | 595 |
| human p68 | GYDSTQQYGSNVPNMHNGMNQQAYAYPATAAAPMIGYPMPTGYSQ | 614 |

2907, C10

Figure 1. Comparison of the deduced amino acid sequence of the chicken DEAD box protein (accession no. AF158370) with the human p68 RNA helicase (accession no. X52104). The underlined sequences are recognised by the antibodies A2, B1, PAb 204, 2907 and C10 directed against human p68 RNA helicase. The boxed sequences are the conserved motifs of the DEAD box family of RNA helicases (37). The peptide sequences obtained from purified 5-MeC-DNA glycosylase complex are in bold italics.

would cross-react with antibodies directed against human p68. A western blot analysis of the purified 5-MeC-DNA glycosylase complex (Fig. 2) shows clearly that two polyclonal antibodies, A2 and B1, which are specific for the N-terminal region of p68 (Fig. 1), cross-react specifically with a protein of the expected size. In contrast, the antibody PAb 204 covering a region poorly conserved in the chicken protein (Fig. 1) does not react at the correct position with the same preparation of the glycosylase. The polyclonal antibody 2907 specific for the C-terminal region of the human p68 helicase gave a signal, albeit not as strong as for the antibodies A2 and B1. The monoclonal antibody C10 raised against a peptide covering the same region as 2970 (Fig. 1) gave no reaction with the purified 5-MeC-DNA glycosylase which was not surprising because a closer look at its epitope revealed only the last five amino acids being recognised (H.Stahl, unpublished results). Finally, the

monoclonal antibody SV-40 T antigen (Ab-2) gave a specific signal with the purified enzyme (Fig. 2).

5-MeC-DNA glycosylase activity coprecipitates with antibodies directed against human DEAD box protein p68

Since a protein of the highly purified 5-MeC-DNA glycosylase complex cross-reacts in western blots with antibodies directed against p68, it was of interest to test whether these antibodies could also precipitate the 5-MeC-DNA glycosylase activity. For the glycosylase assay ribonucleases (pancreatic type) present in the sera had first to be eliminated (Materials and Methods) and the optimal ratios of antisera/antigen for immunoprecipitation were determined experimentally. Two antibodies were tested: the anti p68 B1 and the Ab-2. Activity of 5-MeC-DNA glycosylase was measured in both the supernatant fraction and the immunoprecipitates. High specificity of the reaction was ensured by stringent washing conditions of the immunoprecipitates

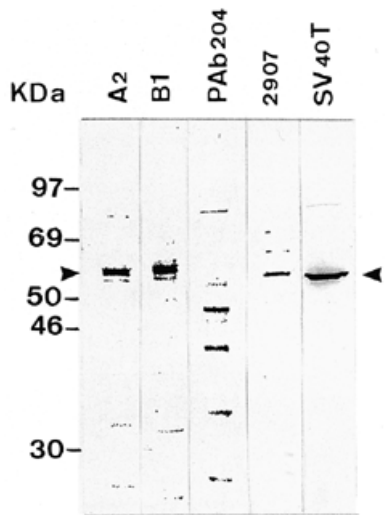


Figure 2. Western blot of purified 5-MeC-DNA glycosylase (2 µg per lane, post butyl-Sepharose fraction) with antibodies directed against human p68 RNA helicase and SV-40 T antigen. Experiments were carried out as outlined in Materials and Methods. Regions of the p68 RNA helicase reacting with the antibodies are underlined in Figure 1. Arrows show the position where 5-MeC-DNA glycosylase activity was located in the SDS-polyacrylamide gel.

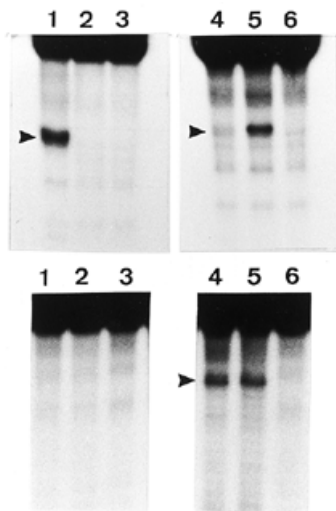


Figure 3. Immunoprecipitation of 5-MeC-DNA glycosylase complex. Upper panels show the direct immunoprecipitation of 5-MeC-DNA glycosylase with SV-40 T Ab-2-agarose (identical results were obtained with anti-p68 B1). Lane 1, activity of 5-MeC-DNA glycosylase in the immunoprecipitates sediments; lane 2, activity in the control sediment; lane 3, blank; lane 4, activity of the enzyme in the supernatant fraction post-immunoprecipitation; lane 5, activity in the supernatant of control beads; lane 6, blank. In the lower panels is the indirect immunoprecipitation of 5-MeC-DNA glycosylase with antibodies directed against DNA methyltransferase. Lanes 1-6 are the same as for the upper panels.

with 0.5 M NaCl. The results of Figure 3 (the two upper panels) show the direct immunoprecipitation of 5-MeC-DNA glycosylase with Ab-2 linked to agarose (direct immunoprecipitation). As

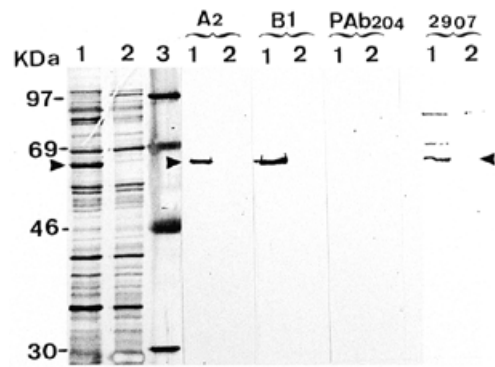


Figure 4. Analysis of the recombinant chicken DEAD box protein produced in *E.coli*. Left panel, silver-stained SDS-PAGE of 2 µg of recombinant fraction (lane 1) and control plasmid (lane 2). Lane 3 shows the size standards. The arrows indicate the correct position of the recombinant protein. The western blot of lanes 1 (recombinant protein) and 2 (control proteins) was carried out with the same p68 antibodies as in Figure 2.

shown in lane 1 of Figure 3, the activity of 5-MeC-DNA glycosylase can be precipitated with the antibody. Sediments of control agarose beads (lane 2) had no activity. Lane 4 of Figure 3 (upper panel) shows that the antibody completely depleted the incubation mixture of 5-MeC-DNA glycosylase, whereas in the supernatant of control beads (lane 5 of Fig. 3) all activity could be recovered. Identical results were also obtained by the indirect immunoprecipitation with the antibody B1 directed against human p68 (result not shown).

We have also observed that DNA methyltransferase co-chromatographed with 5-MeC-DNA glycosylase and it was thought that the two enzymes were also closely associated. To test this possibility we used the same approach as for Figure 3 (upper panel) but with specific antibodies directed against the C- and the N-terminal parts of DNA methyltransferase (29). The lower panel of Figure 3 (lane 1) shows that the antibodies directed against DNA methyltransferase did not precipitate 5-MeC-DNA glycosylase and all activity remained in the supernatant fraction (lane 4).

The specific precipitation of 5-MeC-DNA glycosylase with antibodies directed against human p68 strongly suggests that either the chicken DEAD box protein and 5-MeC-DNA glycosylase are tightly associated and together with the CpG-rich RNA (23) form a complex of DNA demethylation or alternatively 5-MeC-DNA glycosylase is part of the DEAD box protein. This possibility will be investigated with recombinant chicken DEAD box protein.

Production of recombinant chicken DEAD box protein in *E.coli*

Since the upstream non-coding region of chicken DEAD box protein cDNA had no stop signal and no Shine-Dalgarno sequences, it was directly inserted in-frame at the *Bam*HI-*Sal*I restriction sites of the expression vector pQE30. In this construct there is, in addition to the 595 amino acids of the chicken DEAD box protein, a stretch of 45 amino acids (including 6 His) giving an apparent molecular weight of 60 000 on an SDS-polyacrylamide gel (10%) at pH 8.1. Figure 4 lane 1, shows a silver stain of a fraction post Ni-Sepharose containing

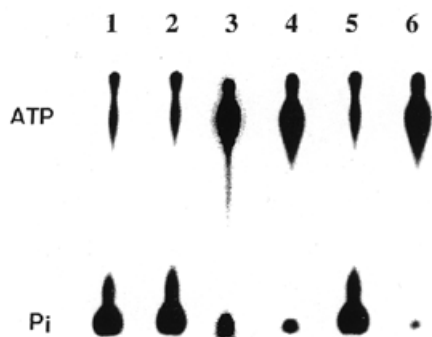


Figure 5. Presence of ATPase in the 5-MeC-DNA glycosylase complex purified from chicken embryos and in the recombinant chicken DEAD box protein. Two micrograms of the 5-MeC-DNA glycosylase (post butyl-Sepharose fraction) were incubated without (lane 1) and with (lane 2) 5 μ g of recombinant RNA (24). Lane 3, as lane 1 except that the enzyme was pre-incubated for 10 min at 37°C in the presence of 50 μ g of ribonuclease A; lane 4, 2 μ g of the recombinant chicken DEAD box protein; lane 5, the same protein plus 5 μ g of recombinant RNA; and lane 6, blank.

the recombinant protein, whereas lane 2 is the extract from bacteria with the control vector. The identity of the 60 kDa recombinant protein was further established by analysing peptides derived from the recombinant by MS and by western blots. MS analysis indicated the presence of the p60 peptides and, in addition, a large quantity of the heat shock protein p60 (a stress induced protein) co-migrating with the recombinant protein. The western blot shown in Figure 4 indicates that only the extract containing the recombinant protein (compare lanes 1 and 2) reacted specifically with p68 antibodies. As for the native protein purified from chicken embryos, the recombinant protein only reacted with antibodies A2, B1 and 2907 but not with PAb 204. Some weak reaction was also obtained with Ab-2 (results not shown).

The fraction obtained from Ni-Sepharose chromatography was further purified by DEAE-Sepharose FPLC and tested for enzyme activities. As seen in Figure 5, the recombinant protein had an RNA-dependent ATPase activity. In addition, as shown in Figure 6, the recombinant protein has also an ATP-dependent helicase activity. However, under our experimental conditions, in the presence or absence of RNA, the recombinant DEAD box protein did not have any trace of 5-MeC-DNA glycosylase activity (results not shown).

Presence of the DEAD box protein mRNA in developing chicken embryos

A comparative enzyme activity study carried out with different organs of 5 and 12 day-old chicken embryos indicated that the highest specific activity of 5-MeC-DNA glycosylase is in the developing brain and eyes (J.-P.Jost, unpublished results). Furthermore, ISH experiments carried out with antisense RNA complementary to one of the cloned CpG-rich RNAs associated with the DNA demethylation complex showed, for example, a specific location of the RNA in the neuroepithelium of the eye and in the developing brain (S.Schwarz, M.Frémont and J.-P.Jost, manuscript in preparation).

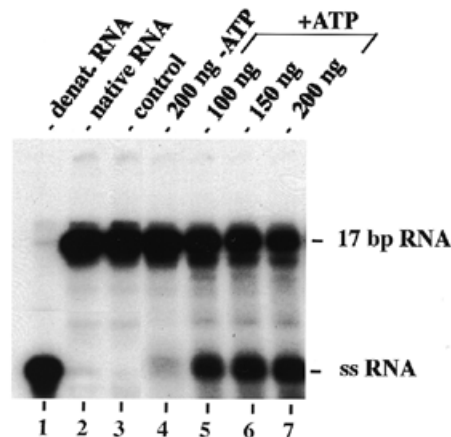


Figure 6. RNA unwinding activity of the recombinant chicken protein. Analysis of the unwinding reaction was done by SDS-PAGE and autoradiography. We used an intermolecular hybrid with a 17 bp double-stranded region as substrate RNA. Lane 1, heat denatured substrate RNAs; lane 2, native substrate RNA; lane 3, substrate RNA incubated without protein; lane 4, substrate RNA incubated with 200 ng recombinant protein under helicase reaction conditions but without ATP; lanes 5–7, substrate RNA incubated with indicated amounts of recombinant protein under helicase reaction conditions in the presence of ATP. All reactions contained 500 ng of poly(C) RNA in addition to the substrate.

Similarly, the results in Figure 7 show clearly that the chicken DEAD box protein is also located in the neuroepithelium of the eye as the CpG-rich RNA associated with DNA demethylation complex. Control experiments carried out with sense RNA probe of the DEAD box protein show a very low background of hybridisation thus ruling out artefactual hybridisation of the antisense probe.

DISCUSSION

The p68 is a nuclear DEAD box protein which is immunologically related to SV-40 T antigen (39). It is developmentally regulated (32) and its presence in vertebrates is correlated with organ differentiation and maturation (32). *In situ* studies carried out with cells in tissue culture indicate that p68 is transiently associated with the nucleoli during the telophase (32,40) and then returns to the nucleoplasm. In addition, cell fractionation studies also indicate a possible close association with the nuclear matrix (41). In this context it is noteworthy that elements of the nuclear matrix have been shown to be involved in the active DNA demethylation (16). Previous work from our laboratory showed that the upstream region of avian Vitellogenin II gene becomes demethylated upon estradiol treatment (42). This region of DNA demethylation was tightly associated with the nuclear matrix (43) suggesting that the nuclear matrix may be associated with DNA demethylation. The question now is to find out how p68 and its chicken homologue could possibly be functionally associated with DNA demethylation. As we have previously shown, DNA demethylation by 5-MeC-DNA glycosylase requires both RNA and proteins (22). An RNA very rich in CpGs was isolated from gel-purified 5-MeC-DNA glycosylase and cloned (23). One function of this RNA in the single-stranded form is to target the site of DNA demethylation (23).

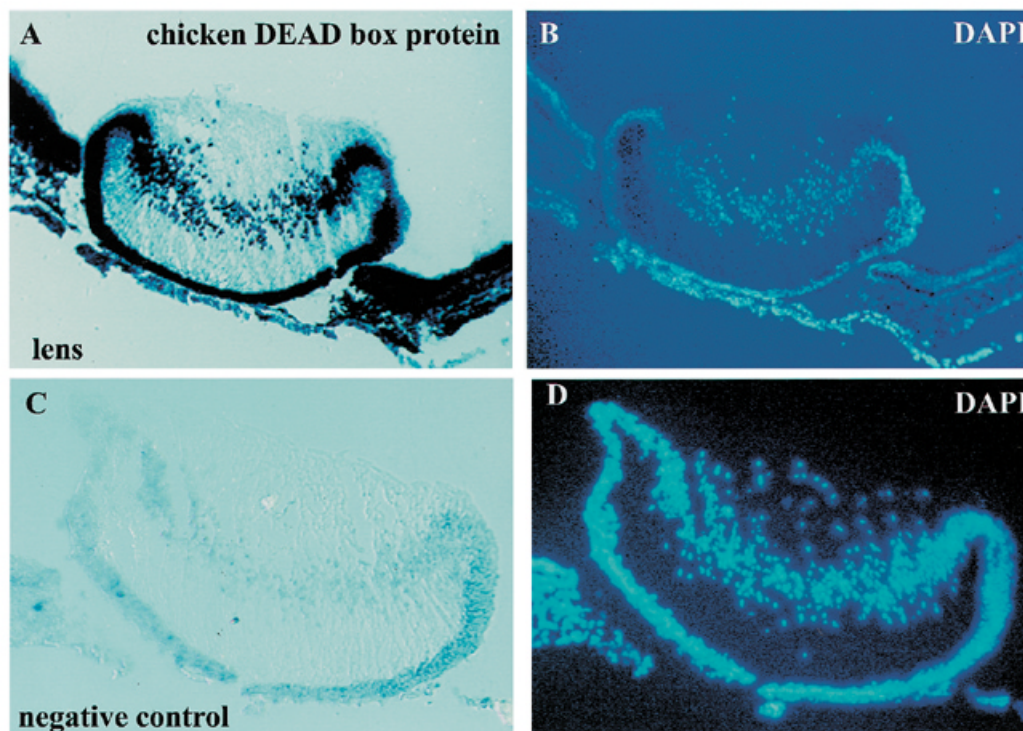


Figure 7. Expression of chicken DEAD box protein in 5 day-old chicken embryos. Experiments were carried out as outlined in the Materials and Methods. (A) Section through the optical cup and lens. The probe was an antisense chicken DEAD box protein (magnification 10 \times). (C) As (A) except hybridisation was done with sense DEAD box protein probe. (B and D) The DAPI staining of the corresponding sections (A) and (C).

Involvement of RNA is not unique to DNA demethylation; it has also been observed for *de novo* RNA methylation (44–46) and DNA methylation in plants (47,48). RNA isolated from purified 5-MeC-DNA glycosylase was shown to have a very complex secondary structure of stem loops under normal salt conditions (S.Schwarz and C.D.Nager, unpublished results). Since many of the possible targeting sequences are in these stems it is conceivable that one role of the chicken p68 homologue is to melt the secondary structure and make the single-stranded sequences available for 5-MeC-DNA glycosylase. We have already shown that the preferred substrate of 5-MeC-DNA glycosylase is hemimethylated DNA that is formed during DNA replication. This hemimethylated DNA is the preferred substrate of two enzymes, which have opposite function: the DNA maintenance methyltransferase and the 5-MeC-DNA glycosylase. The presence of one of the two enzymes in its active form could decide whether a particular CpG site will be methylated or demethylated. In the model a DEAD box protein could possibly have a decisive role for the DNA demodification which is occurring during cell and tissue differentiation. Recent experiments indicate that both the CpG-rich RNA associated with the DNA demethylation complex and the DEAD box protein are localised in the same differentiating chicken embryonic tissues (S.Schwarz, unpublished results). In these tissues, where high levels of RNA helicase and CpG-rich RNA are expressed, it would be interesting to see whether there is also a change in the level of DNA methylation. As far as we can tell,

the DEAD box protein from chicken embryos is not 5-MeC-DNA glycosylase. In experiments where the recombinant protein was first denatured in guanidium hydrochloride and then slowly renatured in the presence or absence of RNA did not show any trace of 5-MeC-DNA glycosylase activity. We could argue that the addition of 45 amino acid residues to the N-terminal part of the DEAD box recombinant protein may influence the folding of the protein and selectively inhibit the 5-MeC-DNA glycosylase activity. Alternatively, the protein may need some specific covalent modifications that are necessary for the 5-MeC-DNA glycosylase activity. Since the DEAD box protein is part of an RNA–protein complex of DNA demethylation (proteins of very similar molecular weight) it is more likely that 5-MeC-DNA glycosylase is a different protein. We have now sequenced some other peptides derived from the highly purified DNA demethylation complex. These peptides are different from the DEAD box protein and they are being presently characterised.

ACKNOWLEDGEMENTS

We would like to thank Mrs Renate Matthies for micro-sequencing the peptides and Mrs Y. C. Jost for typing the manuscript. Many thanks also to Mr P. Müller for the synthesis of the oligonucleotides. This work was sponsored in part by Novartis AG, Basel, Switzerland. F.V.F.-P. is an MRC Senior Fellow.

REFERENCES

1. Russo, V.E.A., Martienssen, R.A. and Riggs, A.D. (1996) *Epigenetic Mechanisms of Gene Regulation*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
2. Jost, J.P. and Bruhat, A. (1997) *Prog. Nucleic Acids Res. Mol. Biol.*, **57**, 217–248.
3. Li, E., Bestor, T.H. and Jaenisch, R. (1992) *Cell*, **69**, 915–926.
4. Tucker, K.L., Beard, C., Dausmann, J., Jackson-Grusby, J., Laird, P.W., Lei, H., Li, E. and Jaenisch, R. (1996) *Genes Dev.*, **10**, 1008–1020.
5. Brannan, C.I. and Bartolomei, M.S. (1999) *Curr. Opin. Genet. Dev.*, **9**, 164–170.
6. Jones, P. and Laird, P.W. (1999) *Nature Genet.*, **21**, 163–167.
7. Holliday, R. (1994) *Dev. Genet.*, **15**, 453–457.
8. Shemer, R., Birger, Y., Dean, W.L., Reik, W., Riggs, A.D. and Razin, A. (1996) *Proc. Natl Acad. Sci. USA*, **93**, 6371–6376.
9. Hsieh, C.L. (1999) *Mol. Cell Biol.*, **19**, 46–56.
10. Chesnokov, I.N. and Schmid, C.W. (1995) *J. Biol. Chem.*, **270**, 18539–18542.
11. Matsuo, K., Silke, J., Georgiev, O., Marti, P., Giorannini, N. and Rungger, D. (1998) *EMBO J.*, **17**, 1446–1453.
12. Bhattacharya, S.K., Ramchandani, S., Cervoni, N. and Szyf, M. (1999) *Nature*, **397**, 579–583.
13. Paroush, Z., Keshet, I., Yisraeli, J. and Cedar, H. (1990) *Cell*, **63**, 1229–1237.
14. Shemer, R., Kafri, T., O'Connell, A., Eisenberg, S., Breslow, J.L. and Razin, A. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 11300–11304.
15. Razin, A., Szyf, M., Kafri, T., Roll, M., Giloh, H., Scarpa, S., Carotti, D. and Cantoni, G.L. (1986) *Proc. Natl Acad. Sci. USA*, **83**, 2827–2831.
16. Kirillov, A., Kistler, B., Mostoslavsky, R., Cedar, H., Wirth, T. and Bergman, Y. (1996) *Nature Genet.*, **13**, 435–441.
17. Jost, J.P. (1993) *Proc. Natl Acad. Sci. USA*, **90**, 4684–4688.
18. Collas, P. (1998) *Nucleic Acids Res.*, **26**, 4454–4461.
19. Schmitz, A., Short, M., Ammerpohl, O., Asbrand, C., Nickel, J. and Renkawitz, R. (1997) *J. Biol. Chem.*, **272**, 20850–20856.
20. Jost, J.P., Siegmann, M., Sun, L. and Leung, R. (1995) *J. Biol. Chem.*, **270**, 9734–9739.
21. Jost, J.P. and Jost, Y.C. (1994) *J. Biol. Chem.*, **269**, 10040–10043.
22. Frémont, M., Siegmann, M., Gaulis, S., Matthies, R., Hess, D. and Jost, J.P. (1997) *Nucleic Acids Res.*, **25**, 2375–2380.
23. Jost, J.P., Frémont, M., Siegmann, M. and Hofsteenge, J. (1997) *Nucleic Acids Res.*, **25**, 4545–4550.
24. Jost, J.P., Siegmann, M., Thiry, S., Jost, Y.C., Benjamin, D. and Schwarz, S. (1999) *FEBS Lett.*, **449**, 251–254.
25. Shevchenko, A., Wilm, M., Vorm, O. and Mann, M. (1996) *Anal. Chem.*, **68**, 850–858.
26. Wilm, M. and Mann, M. (1996) *Anal. Chem.*, **68**, 1–8.
27. Krieg, J., Hartmann, S., Vicentini, A., Gläsner, W., Hess, D. and Hofsteenge, J. (1998) *Mol. Biol. Cell*, **9**, 301–309.
28. Harlow, E. and Lane, D. (1998) *Antibodies a Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
29. Liu, Y.L., Sun, L. and Jost, J.P. (1996) *Nucleic Acids Res.*, **24**, 2718–2722.
30. Palacios, R., Palmiter, R.D. and Schimke, R.T. (1972) *J. Biol. Chem.*, **247**, 2316–2321.
31. Chomczynski, P. and Sacchi, N. (1987) *Anal. Biochem.*, **162**, 156–159.
32. Fuller-Pace, F.V., Nicol, S.M., Reid, A.D. and Lane, D.P. (1993) *EMBO J.*, **12**, 3619–3626.
33. Scheffner, M., Knippers, R. and Stahl, H. (1998) *Cell*, **57**, 955–963.
34. Schaeren-Wiemers, N. and Gerfin-Moser, A. (1993) *Histochemistry*, **100**, 431–440.
35. Hloch, P., Schiedner, G. and Stahl, H. (1990) *Nucleic Acids Res.*, **18**, 3045.
36. Aubourg, S., Kreis, M. and Lecharny, A. (1999) *Nucleic Acids Res.*, **27**, 628–636.
37. Lamm, M.G., Nicol, S.M., Fuller-Pace, F.V. and Lamond, A.I. (1996) *Nucleic Acids Res.*, **24**, 3739–3747.
38. Stevenson, R.J., Hamilton, S.J., MacCallum, D.E., Hall, P.A. and Fuller-Pace, F.V. (1998) *J. Pathol.*, **184**, 351–359.
39. Lane, D.P. and Hoeffler, W.K. (1980) *Nature*, **288**, 167–170.
40. Iggo, R.D., Jamieson, D.J., MacNeil, S.A., Southgate, J., McPheat, J. and Lane, D.P. (1991) *Mol. Cell Biol.*, **11**, 1326–1333.
41. Hirling, H., Scheffner, M., Restle, T. and Stahl, H. (1989) *Nature*, **339**, 562–564.
42. Saluz, H.P., Jiricny, J. and Jost, J.P. (1986) *Proc. Natl Acad. Sci. USA*, **83**, 7167–7171.
43. Jost, J.P. and Seldran, M. (1984) *EMBO J.*, **3**, 2005–2008.
44. Kiss-László, Z., Bachelier, H.Y., Caizergues-Ferrer, M. and Kiss, T. (1996) *Cell*, **85**, 1077–1088.
45. Cavaillé, J., Nicoloso, M. and Bachelier, J.P. (1996) *Nature*, **383**, 732–735.
46. Tollervey, D. (1996) *Science*, **273**, 1056–1057.
47. Wassenegger, M., Heimes, S., Riedel, L. and Sänger, H.L. (1994) *Cell*, **76**, 567–576.
48. Péliissier, T., Thalmeir, S., Kempe, D., Sänger, H.L. and Wassenegger, M. (1999) *Nucleic Acids Res.*, **27**, 1625–1634.