Published online 10 August 2010 *Nucleic Acids Research, 2010, Vol. 38, No. 22* ***8239–8256***
doi:10.1093/nar/gkq691

Critical lysine residues within the overlooked N-terminal domain of human APE1 regulate its biological functions

Damiano Fantini[1], Carlo Vascotto[1], Daniela Marasco[2,3], Chiara D'Ambrosio[4], Milena Romanello[1], Luigi Vitagliano[3], Carlo Pedone[2,3], Mattia Poletto[1], Laura Cesaratto[1], Franco Quadrifoglio[1], Andrea Scaloni[4], J. Pablo Radicella[5] and Gianluca Tell[1,*]

[1]Department of Biomedical Sciences and Technologies, University of Udine, 33100 Udine, [2]Department of Biological Sciences, University of Naples 'Federico II', [3]Institute of Biostructures and Bioimaging, National Research Council, 80134 Naples, [4]Proteomics and Mass Spectrometry Laboratory, ISPAAM, National Research Council, 80147 Naples, Italy and [5]CEA, Institut de Radiobiologie Cellulaire et Moléculaire, UMR217 CNRS, F-92265 Fontenay-aux-Roses, France

Received May 27, 2010; Revised July 20, 2010; Accepted July 22, 2010

ABSTRACT

Apurinic/apyrimidinic endonuclease 1 (APE1), an essential protein in mammals, is involved in base excision DNA repair (BER) and in regulation of gene expression, acting as a redox co-activator of several transcription factors. Recent findings highlight a novel role for APE1 in RNA metabolism, which is modulated by nucleophosmin (NPM1). The results reported in this article show that five lysine residues (K24, K25, K27, K31 and K32), located in the APE1 N-terminal unstructured domain, are involved in the interaction of APE1 with both RNA and NPM1, thus supporting a competitive binding mechanism. Data from kinetic experiments demonstrate that the APE1 N-terminal domain also serves as a device for fine regulation of protein catalytic activity on abasic DNA. Interestingly, some of these critical lysine residues undergo acetylation *in vivo*. These results suggest that protein–protein interactions and/or post-translational modifications involving APE1 N-terminal domain may play important *in vivo* roles, in better coordinating and fine-tuning protein BER activity and function on RNA metabolism.

INTRODUCTION

APE1/Ref-1 (APE1) is an essential protein acting as a master regulator of cellular responses to oxidative stress and contributing to the maintenance of genome stability (1–3). APE1 is involved both in the base excision repair (BER) of DNA lesions, functioning as the major DNA AP-endonuclease in mammals, and in transcription, acting as a redox co-activator of different transcription factors such as Egr-1, NF-κB and p53 (4,5). These biological activities are located in two functionally distinct protein regions. The N-terminal domain (residues 1–127), containing the nuclear localization signal (NLS), is mainly devoted to the transcriptional co-activating function, while the C-terminal portion (residues 61–318) exerts the enzymatic activity on abasic sites in DNA (4–7). A third function, regulated by K6/K7 acetylation (8), is its transcriptional repressor activity through indirect binding to the negative Ca^{2+}-response elements (nCaRE) (9,10). A completely unexpected new function has been recently unveiled and is associated to a role in RNA metabolism (11–13), which could explain some of the effects of APE1 on gene expression through post-transcriptional mechanisms.

Extensive structural investigations have shown that the first 42 amino acids of APE1 are disordered, while the remainder of the protein has a globular fold (14–16). APE1 homology with *Escherichia coli* ExoIII starts from residue 62; a Blast homology search against non-redundant databases clearly shows that the protein C-terminus is highly conserved, while the N-terminus (first ~40 residues) is not. Conservation of APE1 N-terminal portion is high in mammals (>90%) but almost absent in other organisms, with the exception of

*To whom correspondence should be addressed. Tel: +39 0432 494311; Fax: +39 0432 494301; Email: gianluca.tell@uniud.it

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

Danio, *Drosophila*, *Xenopus* and *Dictyostelium* (having a sequence identity <40%), suggesting that it may be a recent acquisition during evolution (17). Interestingly, the *Drosophila* homologue of mammalian APE1 (i.e. Rrp1), which also bears a bipartite arrangement, presents an unstructured N-terminal domain which is used to interact with Pol ζ (18,19). Accordingly, data on the human APE1 (hAPE1) clearly show that protein N-terminus (first ~35 amino acids) is used for interacting with other partners, such as XRCC1 (20), CSB (21), NPM1 and others (11), possibly modulating the different APE1 functions. These data underscore a common emerging theme in this protein: the presence of an unstructured N-terminus for mediating protein–protein interaction. Its truncation has been related to an enhanced immune cell death mediated by granzyme A and K (22,23), to a reduction of hAPE1 accumulation within nuclei (22,24) and to a loss of its interaction with XRCC1 (20) and other protein partners (11). A truncated form has also been detected in mitochondria, where it could exert a role in mtDNA repair (25–27). As the protein lacking the first 33 residues (hAPE1NΔ33) has been associated with an apoptotic phenotype (23), it cannot be excluded that its generation may promote a cytotoxic outcome, driving pro-apoptotic triggering directly within mitochondria.

hAPE1 is mainly a nuclear protein and is critical for controlling cellular proliferative rates (1,2,28). However, cytoplasmic, mitochondrial and ER localizations have also been reported (25–27,29,30). hAPE1 is an abundant and relatively stable protein in mammalian cells (4,5). Fine-tuning of the multiple hAPE1 functions may therefore depend on post-translational modifications (PTMs) and on its interactome modulation. While a functional role has been determined for some PTMs (K6/K7 acetylation and K24/K25/K27 ubiquitination) (8,31,32), the identity and the role of the interacting partners in modulating hAPE1 biological functions are still under investigation (11,31,33). By a proteomic approach, we recently discovered that hAPE1, through its 33N-terminal amino acids (33N-term), may functionally interact with different proteins involved in RNA metabolism, unveiling its unexpected function in RNA cleansing processes (11,12). We demonstrated that hAPE1 may affect cell growth by directly acting on RNA quality-control mechanisms, thus regulating gene expression through post-transcriptional mechanisms; in particular, hAPE1 interaction with NPM1 may affect its activity over rRNA molecules. However, many aspects of this new function are still unknown (12).

In this study, we explore the role of hAPE1 N-terminal domain by analysing the molecular details of its binding to both RNA and NPM1 and their impact on the DNA repair activity of the protein. Five lysine residues, located in the unstructured hAPE1 N-terminal domain made of 33 amino acids (33N-term) probably acquired during evolution, resulted crucial for hAPE1 interaction with RNA and NPM1. Interestingly, we found that some of these lysine residues resulted acetylated *in vivo*. Moreover, kinetic/functional studies showed that the unstructured 33N-term serves as a device for subtle regulation of hAPE1 catalytic activity on abasic DNA through regulation of product binding. These findings suggest that other proteins can modulate hAPE1 BER activity and ensure the correct processing of DNA intermediates and may have an impact on the RNA cleansing process. These data shed new light on the fine-tuning mechanisms responsible for hAPE1 regulation and open the way for developing new strategies for therapeutical intervention based on the modulation of hAPE1 functions.

MATERIALS AND METHODS

Biophysical and evolutionary analysis of APE1 sequences

Biophysical parameters of different APE1 peptides (molecular weight and pI) were calculated by the computing algorithms available at the Swiss Institute of Bioinformatics (http://www.expasy.ch/tools/pi_tool.html).

For phylogenetic analysis, a total of 17 APE1 orthologous metazoan sequences (*Homo sapiens, Pan troglodytes, Macaca mulatta, Equus caballus, Sus scrofa, Bos taurus, Canis lupus familiaris, Mus musculus, Rattus norvegicus, Ornithorhyncus anatinus, Xenopus laevis, Xenopus tropicalis, Salmo salar, Danio rerio, Gasterosteus aculeatus, Caenorhabditis elegans and Strongylocentrotus purpuratus*) were retrieved from databases using the BLAST algorithm (34). Out-group sequences were not included because of the almost complete lack of the N-terminal domain in other organisms. Alignment of the amino acid residues 1–60 and of the cDNA nucleotide 1–180 was made with the MAFFT program (v6.531b) (35) with default parameters, and evolutionary trees were inferred using the software MEGA version 4 (36). The minimum evolution tree-building method was used to reconstruct the phylogenetic tree (37); dendrograms' topologies were checked by the bootstrap method with 1000 replicates.

Cell culture and transient transfection experiments

HeLa cells were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% foetal calf serum (FCS), 100 U/ml penicillin and 10 μg/ml streptomycin sulphate. One day before transfection, cells were seeded in 10 cm plates at a density of 1.8×10^6 cells/plate. Cells were then transiently transfected with 2 μg of either pFLAG5.1 (empty vector) or pFLAG5.1-hAPE1 (wild-type and specific mutant forms) plasmids per dish, using Lipofectamine 2000 Reagent (Invitrogen, Milan, Italy) according to the manufacturer's instructions. Cells were harvested 48 h after transfection. Co-immunoprecipitation analysis was performed as described earlier (11).

Western blot analysis

The protein samples were electrophoresed onto a 10–12% sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS–PAGE). Proteins were then transferred to nitrocellulose membranes (Schleicher & Schuell, Keene, NH, USA), revealed as already described (11) using ECL chemiluminescence procedure (GE Healthcare, Milan, Italy) and quantified by using a Gel Doc 2000 video-densitometer (Bio-Rad, Milan, Italy).

Plasmids and expression of recombinant proteins

The constructs pGex3X-hAPE1 and pGex-3X-zAPE1, coding for the glutathione S-transferase (GST)-hAPE1 and zebrafish zAPE1 full-length proteins, respectively, were kindly provided by Mark R. Kelley (Indiana University, IN, USA). The constructs pGex3X-NΔ33hAPE1, pGex3X-Δ22hAPE1 and pGex3X-Δ9hAPE1, coding for truncated hAPE1 forms were generated by subcloning from the previous one through PCR and confirmed by direct sequencing. Plasmids pGex3X-hAPE1K3 mutant (K24A, K25A, K27A) and pGex3X-hAPE1K5 mutant (K24A, K25A, K27A, K31A, K32A) were generated using the Quickchange II XL Mutagenesis kit (Stratagene, Cedar Creek, TX, USA) and the pGex3X-hAPE1 as template, following manufacturer's instructions. The plasmid coding for GST-NPM1 full-length was kindly provided by P. G. Pelicci (European Institute of Oncology, Milan, Italy). For expression of His_6-tagged NPM1, cDNA coding for full-length NPM1 was subcloned in the pET-15b-expressing plasmid (Invitrogen).

All proteins were expressed in *E. coli* BL21(DE3), induced with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG), and then purified on an AKTA Prime FPLC system (GE Healthcare) by using a GSTrap HP column (GE Healthcare) (for GST-tagged proteins) or a HisTrap HP column (GE Healthcare) (for His_6-NPM1). The quality of purification was checked by silver-stained SDS–PAGE analysis. Extensive dialysis against PBS was performed to remove any trace of imidazole from the HisTrap-purified proteins.

Accurate quantification of all recombinant proteins was performed by colorimetric Bradford assays (Bio-Rad) and confirmed by SDS–PAGE and western blotting analysis using anti-GST peroxidase-conjugated antibody (Sigma, Milan, Italy). To remove the GST tag, GST–hAPE1 and GST–hAPE1 mutant proteins were further hydrolyzed with factor Xa (five factor Xa units per milligram of recombinant GST-fused protein). The protease was removed from the sample using a benzamidine HiTrap FF column (GE Healthcare); the proteins were then purified on a GSTrap HP column (GE Healthcare).

The correct molecular mass of each recombinant protein was assessed by LC–MS analysis.

Peptide synthesis

Reagents for peptide synthesis were from Novabiochem (Laufelfingen, Switzerland) and Inbios (Pozzuoli, Italy). Solid phase peptide syntheses were performed on an automated synthesizer Syro I (MultiSynTech, Witten, Germany). Preparative RP–HPLC were carried out on a Shimadzu LC-8A, using a Phenomenex Luna-COMBI HTS C18 column (5 × 2.12 cm, 10 μm). LC–ESI–Ion Trap (IT)-MS analysis was carried out on a Surveyor HPLC system connected to a LCQ DECA XP IT (Thermo Scientific, USA), hAPE1 peptide 1–33 was purchased from EzBiolab (Westfield, IN, USA) and used without any further purification. Peptides corresponding to hAPE1 fragments 1–9 (in a free and acetylated version in K6 and K7) and 23–33 were prepared following standard Fmoc (9-fluorenylmethoxycarbonyl) chemistry protocols (38,39). Products were purified to homogeneity by RP–HPLC and their identity and purity were confirmed by LC–MS analysis.

GST pull-down assay with RNA or NPM1 and co-immunoprecipitation with NPM1 in transfected cells

An amount of 0.12 nmol of either GST–hAPE1 or each GST–hAPE1 mutant protein were added, together with 0.2 nmol of NPM1 protein or 10 μg of total RNA, to 10 μl of glutathione–Sepharose 4B beads (GE Healthcare). RNA was extracted using Trizol® Reagent (Invitrogen). Binding was performed in PBS supplemented with 5 mM DTT and 1 × protease inhibitor cocktail (Sigma) for 2 h, under rotation, at 4°C. The beads were washed three times with washing buffer (PBS) supplemented with 0.1% w/v Igepal CA-630 (Sigma), 5 mM DTT, 1 × protease inhibitor cocktail. Beads were then resuspended in 1 × Laemmli sample buffer containing 100 mM DTT for western blot analysis or in Trizol® reagent for RNA extraction and rRNA quantification. For western blot experiments, GST-tagged and NPM1 proteins were detected with an anti-GST peroxidase-conjugated antibody (Sigma) and an anti-NPM1 monoclonal antibody (NPM1: clone FC-61991; Invitrogen), respectively.

Co-immunoprecipitation analyses of hAPE1 and its different mutants with NPM1 were performed as described earlier (11).

Q-PCR analysis

Levels of 28S and 18S rRNA transcripts were determined as described earlier (11) by using the iScript cDNA synthesis kit (Bio-Rad) according to the manufacturer's instructions. Q-PCR was performed with an iQ5 multicolor real-time PCR detection system (Bio-Rad), according to the manufacturer's protocol. The following primers were used: 18S_for (5′-CTGCCCTATCAACTTTCGATGGTAG-3′) and 18S_rev (5′-CCGTTTCTCAGGCTCCCTCTC-3′), which amplified a region of 100 bp; 28S_for (5′-TGTCGGCTCTTCCTATCATTGT-3′) and 28S_rev (5′-ACCCAGCTCACGTTCCCTATTA-3′), which amplified a region of 81 bp.

Each sample analysis was performed in triplicate. As negative control, a sample without the template was used; as control for genomic DNA contamination, a sample with non-retro-transcribed RNA was included instead of template cDNA. The cycling parameters were denaturation at 95°C for 10 s and annealing/extension at 60°C for 30 s (repeated 40 times). In order to verify the specificity of the amplification, a melting-curve analysis was performed, immediately after the amplification protocol. A standard curve was generated by using a calibrating cDNA. This cDNA was obtained from a retrotranscription of 1 μg of total RNA from HeLa cells, which was serially diluted and analysed for 18S and 28S rRNA. We assumed that rRNA represented 90% of the total RNA, in a 28S/18S rRNA molar ratio of 2:1. By using the equation describing the plots of the $\log_{10}$ of the starting amount of five dilutions of the calibrator

cDNA versus the corresponding threshold cycle, the iQ5 optical system software calculated the amount of the template for each sample.

Immunofluorescence confocal analysis

For experiments on hAPE1-Flag/NPM1 colocalization, a double immunofluorescence procedure was carried out as described earlier (11), where hAPE1 was labelled by an anti-Flag antibody-FITC conjugate (Sigma), whereas NPM1 was labelled with the Zenon® Mouse IgG Labeling Kit (Invitrogen), using the Alexa Fluor 546-labelled Fab fragment directed against the Fc portion of the IgG primary anti-NPM1 antibody (Invitrogen), according to the manufacturer's instructions. Nuclei were counterstained with DAPI (Sigma); microscope slides were mounted and visualized through a Leica TCS SP laser-scanning confocal microscope.

AP site incision assays on abasic dsDNA

Cleavage assays were performed as described earlier (11,20) with some minor modifications and using a fluorescent abasic dsDNA probe called 34FDNA (11). When the effect of other proteins on the hAPE1 activity was investigated, proteins were first incubated in 5 mM Tris–HCl, pH 7.6, 5 mM KCl, 1 mM $MgCl_2$, 0.1 mg/ml bovine serum albumin and 0.005% Triton X-100, on ice, for 20 min. Then, 150 fmol of the labelled duplex containing a THF residue were added to each sample (11). Reactions were carried out at 37°C, for 30 min, and stopped by adding 4 µl of formamide dye (10 mM EDTA, 0.5% Bromophenol Blue, 80% formamide) followed by heating for 5 min, at 95°C, before loading. Reaction products were separated by denaturing 20% PAGE (19 : 1 acrylamide:bis-acrylamide). Gels were scanned and band intensities were quantified using a Storm Phosphoimager (Molecular Dynamics).

hAPE1–RNA-binding assays through EMSA

hAPE1–RNA complex formation was assessed as described earlier with a 34-mer synthetic RNA oligonucleotide containing an abasic site called 34FRNA (11), incubated on ice for 10 min and separated onto a native 8% w/v polyacrylamide gel at 120 V, for 2 h. Each reaction contained 2.5 pmol of 34FRNA labelled substrate and the indicated amount of proteins in a final volume of 10 µl.

Surface plasmon resonance experiments

Real time binding assays were performed on a Biacore 3000 Surface Plasmon Resonance (SPR) instrument (GE Healthcare). Full-length hAPE1 and hAPE1NΔ33 were immobilized on a CM5 Biacore sensor chip in 10 mM sodium acetate, pH 5.5, by using the EDC/NHS chemistry, with a reaction flow rate of 5 µl/min and a time injection of 7 min (40). Binding assays were carried out by injecting 90 µl of analyte, at 20 µl/min, with HBS (10 mM HEPES, 150 mM NaCl, 3 mM EDTA, pH 7.4), 0.1 mM TCEP as running-buffer. BIAevaluation analysis package (version 4.1, GE Healthcare), was used to subtract the signal of the reference channel, while RU_{max} data versus proteins concentrations were fitted with GraphPad Prism, vers. 4.00 (San Diego, CA), following a procedure reported earlier (41).

Mass spectrometry analysis

Identification of PTMs occurring on hAPE1 was performed on immunopurified hAPE1 obtained from hAPE1-Flag expressing HeLa cells (11). hAPE1 was resolved by 12% SDS–PAGE; corresponding protein band was excised, in-gel S-alkylated, digested with endoprotease AspN and analysed as reported earlier (11,42). Details on mass spectrometry analysis are described in the Supplementary Data.

Statistical analysis

Statistical analysis was performed using the Microsoft Excel data analysis program for Student's *t*-test analysis. $P < 0.05$ was considered as statistically significant.

RESULTS

Primary structure analysis of hAPE1N-terminal sequence

NPM1 and RNA interact with hAPE1 through its 33N-term (11). In order to map the specific residues responsible for binding, we performed a preliminary analysis of the 33N-term sequence. Based on the distribution of charged amino acids at physiological pH, the 33N-term can be subdivided into three distinct regions: (i) a first basic sequence comprising amino acids 1–9, (ii) an acidic sequence comprising the amino acids 10–22 and (iii) a second basic region, spanning amino acids 23–33 and containing five lysine residues (Figure 1A). Interestingly, the 33N-term confers a calculated basic pI (8.3) to the full-length protein, while its removal decreases the pI to 6.6. These calculations were supported by experimental evaluation of the pI of the two protein forms through 2-DE analysis (Supplementary Figure S1).

The 33N-term peptide is able to bind RNA molecules by itself

Previous *in vitro* experiments demonstrated that hAPE1 is able to bind RNA (11). In this work, we verified that this process reasonably occurs also *in vivo*, as demonstrated by co-localization analysis of hAPE1 with RNA in HeLa cells (Supplementary Figure S2). Due to its basic features, the 33N-term peptide could potentially bind to RNA by itself. We therefore tested peptide binding to RNA by EMSA, as described earlier (11). Figure 1B (left) shows that the peptide 1–33 binds a 34-mer RNA oligonucleotide containing an abasic sites (34FRNA) (11) and competes for its binding with the full-length hAPE1 protein. More striking results were obtained with rRNA, through GST pull-down competition experiments (Figure 1B, right) and with t-RNA molecules of different origins (data not shown).

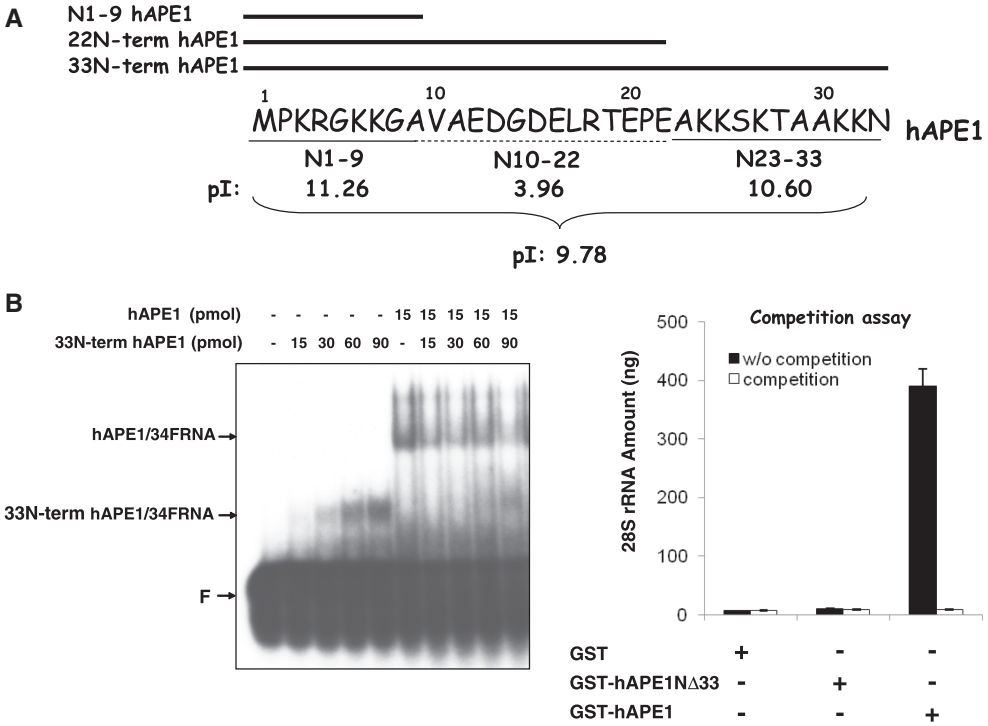

Figure 1. RNA-binding activity of 33N-term hAPE1. (**A**) Primary structure and biophysical parameters of 33N-term hAPE1. (**B**, left) 33N-term hAPE1 shows RNA-binding ability. EMSA experiments were performed as described in 'Materials and Methods' section with purified recombinant hAPE1 or a synthetic 1–33 hAPE1 peptide (33N-term hAPE1). Competition experiments were performed with the indicated amount of peptide co-incubated with the hAPE1 during the binding reaction. Mixtures were separated on a native 10% (w/v) polyacrylamide gels at 120 V for 2 h. F indicates the free 34FRNA oligonucleotide ^{32}P-labelled probe. The image is from a representative gel out of three independent experiments, (B, right) 33N-term hAPE1 competes with hAPE1 for 28S rRNA binding. GST pull-down experiments were performed with total RNA extracted from HeLa cells as prey, a 10-fold molar excess of the 33N-term hAPE1 synthetic peptide as competitor and GST, GST–hAPE1 or GST-hAPE1NΔ33 as baits (see 'Materials and Methods' section for details). Histograms represent average values with SDs of three independent experiments.

Mapping the hAPE1 amino acids involved in RNA binding

We mapped the amino acids responsible for RNA binding through deletion analyses and competition assays with peptides spanning the regions 1–9, 23–33 and 1–33. The binding abilities of hAPE1NΔ9, hAPE1NΔ22 and hAPE1NΔ33 (Figure 2A) were evaluated through GST pull-down experiments with total RNA samples purified from HeLa cells as prey (11). Figure 2B shows that amino acid residues 23–33 are required for efficient rRNA binding by hAPE1, as also confirmed by competition experiments with the specific 1–9, 23–33 and 1–33 peptides (Figure 2C). Acetylation of K6 and K7 residues, which regulates hAPE1 binding to nCaRE sequences (8), does not influence the hAPE1 RNA-binding activity, as determined by competition experiments with the acetylated peptide, i.e. 1-9ac hAPE1 (Figure 2C). The reported experiments, relative to hAPE1 binding to 28S rRNA, were also performed with 18S rRNA giving similar results (not shown).

Overall, these experiments demonstrate a leading role for the 23–33 basic amino acid sequence in mediating hAPE1 binding to RNA.

Specific lysine residues are required for hAPE1 binding to RNA

Since electrostatics may drive the RNA-binding activity of hAPE1, we tested the hypothesis that basic residues within the 23–33 hAPE1 sequence are responsible for mediating stable binding to RNA. Thus, we generated mutants of the protein bearing a triple (K24/K25/K27) and quintuple (K24/K25/K27/K31/K32) K to A substitutions, here referred to as hAPE1K3 and hAPE1K5, respectively. Recombinant proteins were expressed in *E. coli* and chromatographycally purified. SDS–PAGE analyses of purified proteins showed that each mutant displays an altered mobility (Figure 3A, left), possibly due to alteration of its overall charge since LC–MS analyses confirmed the correctness of their masses and the difference in their apparent mobility observed in SDS–PAGE was abolished when separating proteins in urea-containing denaturing gels (not shown). Then, their binding activity to 34FRNA was measured through EMSA analysis (Figure 3A, right). The obtained results clearly demonstrate that the hAPE1 binding to this small RNA strictly depends on the presence of these lysine residues. GST pull-down experiments with 28S and 18S rRNA (Figure 3B and data not shown) were in good agreement with the EMSA experiments. As expected, hAPE1K5 mutant was as deficient in binding to both 28S (Figure 3B) and 18S rRNA (not shown) as the hAPE1NΔ33 deletion mutant. On the contrary, the hAPE1K3, which was proficient in binding to 28S RNA, as shown in Figure 3B, showed only a slightly reduced ability to bind 18S RNA compared to the entire hAPE1 (not shown).

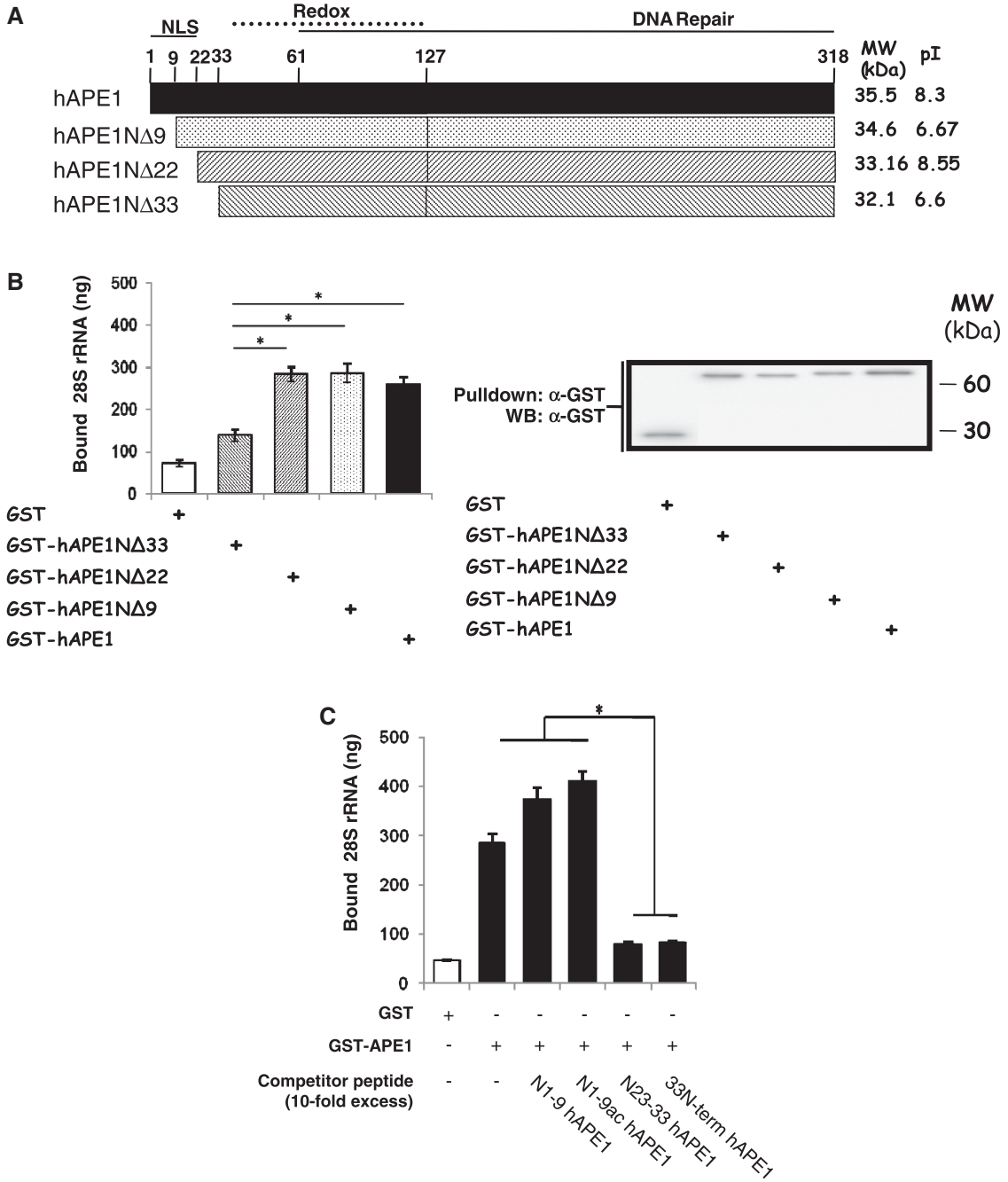

Figure 2. Mapping of the RNA-binding domain of hAPE1. (**A**) Schematic representation of the hAPE1 deletion mutants used. (**B–D**) An amount of 120 pmol of each bait protein was incubated with 10 μg of total RNA purified from HeLa cells, as described in 'Materials and Methods' section. (B, right) After GST pull-down, samples underwent either western blot analysis using an anti-GST antibody or RNA extraction and quantification by reverse transcription and Q-PCR analysis using 28S rRNA specific primers. The amounts of 28S rRNA recovered were then normalized versus the amount of bait proteins and the resulting values are plotted in the histograms in which average values with SDs of three independent experiments are plotted ($*P < 0.05$) (B, left). (C) The same procedure as in (B) was used in the presence of 10-fold molar excess competitor peptides: N1-9 hAPE1 spanning the first 9N-terminal aminoacids of APE1; 1–9ac hAPE1 indicates the same peptide bearing acetylation on residues 6 and 7; N23–33 hAPE1 spanning the region 23–33 of hAPE1, and 33N-term hAPE1 spanning the first 33N-terminal amino acids of the protein. Histograms represent the normalized amount of recovered 28S rRNA and the bars represent SDs of three independent experiments ($*P < 0.05$).

Lysine residues within the 23–33 region control the hAPE1 endonuclease activity on abasic DNA

We (11) and others (25) previously reported that the enzymatic activity on abasic dsDNA (34FDNA) of the hAPE1NΔ33 protein was significantly higher than that of the wild-type protein. We therefore evaluated the effect of the K to A mutations on the endonuclease activity of APE1 over abasic dsDNA, by measuring the kinetic parameters (K_M and k_{CAT}) of hAPE1, hAPE1NΔ33 and the two K to A mutants (Table 1).

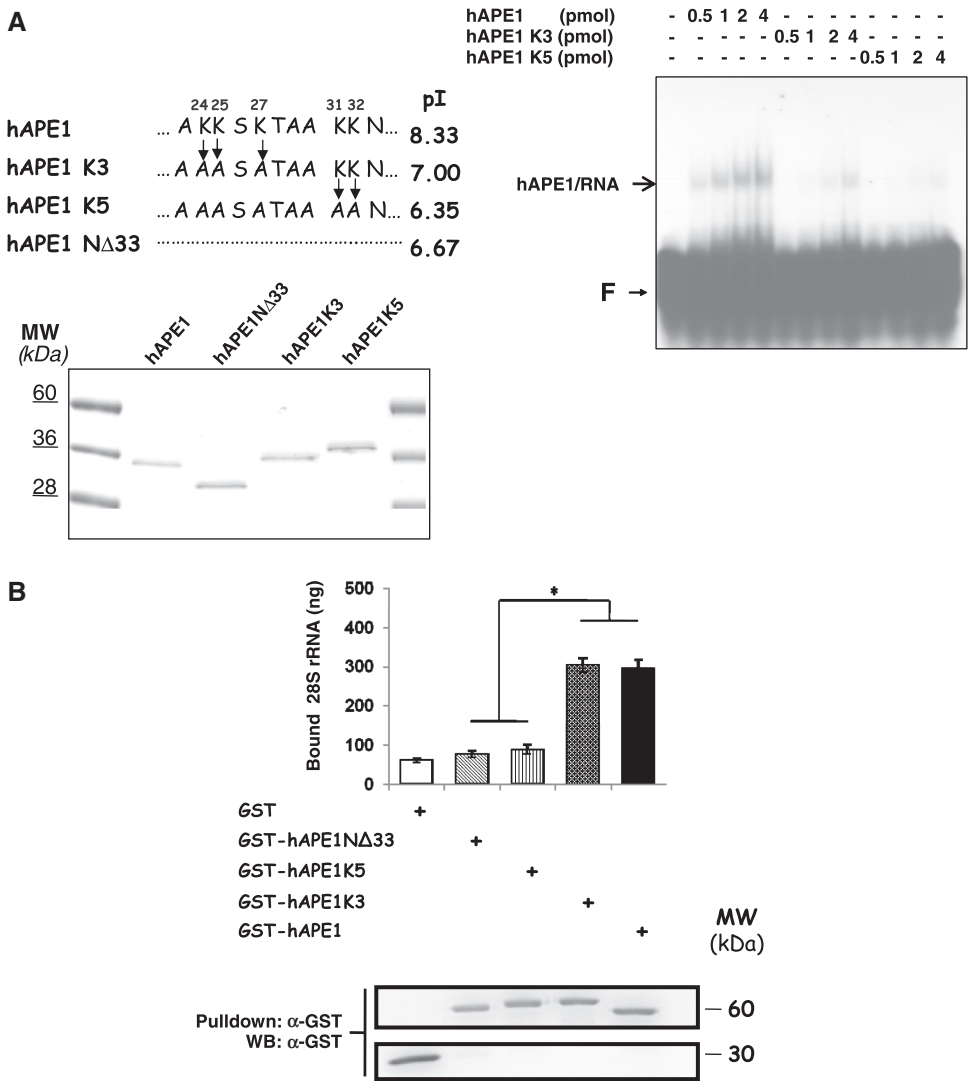

Figure 3. Lysine residues in the 23–33 basic region are responsible for hAPE1 binding to RNA. (**A**, left) Lysine to alanine mutants of hAPE1 were generated through site-directed mutagenesis. Recombinant proteins were expressed in *E. coli*, purified to homogeneity and used for further assays. The correct molecular mass of each recombinant protein was assessed by LC–MS analysis. Representative Coomassie-stained 10% SDS–PAGE of 1 μg of each recombinant protein is shown. (A, right) Purified hAPE1, hAPE1-K3 and hAPE1-K5 proteins were assayed for their ability to bind to 34FRNA by EMSA. A representative EMSA gel image of three independent experiments is shown. (**B**) GST pull-down experiments were performed, as described in 'Materials and Methods' section, with total RNA extracted from HeLa cells as prey to test the RNA-binding ability of the hAPE1-K3 and hAPE1-K5 mutant proteins in comparison with the wild-type hAPE1 and the hAPE1NΔ33. After pull-down, samples underwent either western blot analysis (using an anti-GST antibody) to confirm the amount of bait protein present in each sample, or 28S rRNA quantification. The amounts of 28S rRNA recovered after pull-downs were then normalized versus the amount of bait proteins. The resulting histogram shows the average values with SD of three independent experiments ($*P < 0.05$).

Since the different proteins exhibited very different endonuclease activities, an initial set of experiments (Figure 4A) was performed in order to select the most suitable enzyme concentration for further analysis. An enzyme concentration capable of cleaving ~7 nM 34FDNA over the time of reaction, indicated by the red spots, was selected for time-lapse experiments and its value is indicated in the plots. K_M and k_{CAT} values for each enzyme are shown in Table 1, and are in good agreement with the parameters reported earlier by other authors (25,43).

Our findings (Table 1) show that the K_M value is similar for all protein variants, therefore suggesting that the enzyme affinity for the abasic site in the 34FDNA substrate is not significantly affected by the removal of the 33N-term or by mutation of the critical lysine residues within the region 23–33. On the contrary, k_{CAT} values are very different, indicating that the enzymatic catalysis is faster for the NΔ33 or the K5 mutants than for the full-length protein. These experiments were performed with different preparations of recombinant proteins confirming the reliability of our data.

These findings suggest that molecular interactions and electrostatics involving the amino acid region 23–33 of hAPE1, though not appreciably changing the binding affinity to a double-stranded abasic substrate, participate in the modulation of hAPE1 endonuclease activity.

Table 1. Kinetic parameters for endonuclease activity on the abasic DNA of hAPE1 and APE1 mutant proteins

	hAPE1	hAPE1 K3	hAPE1 K5	hAPE1 NΔ33
K_M (nM)	40.57 ± 11.62	31.48 ± 8.19	27.72 ± 5.67	39.29 ± 9.20
k_{CAT} (min^{-1})	5.99 ± 0.70	20.22 ± 2.34	67.29 ± 5.61	57.97 ± 8.98
k_{CAT}/K_M ($nM^{-1}\,min^{-1}$)	0.15	0.64	2.62	1.48

Kinetic parameters (K_M and k_{CAT}) were calculated from the measurement of the endonucleolytic reaction rates for hAPE1 and the different mutant hAPE1 proteins using a Cy5-labelled 34FDNA as substrate, at 37°C, as described in (11) and detailed in 'Materials and Methods' section, and using a Lineweaver–Burk plot analysis. In order to evaluate K_M and k_{CAT} of the different hAPE1 forms, we performed time-lapse experiments using increasing concentrations of the substrate. The reactions were monitored at 0, 5, 10, 15 and 20 min. For each protein, the concentration of the product increased in a linear manner over time by using the selected enzyme concentration (refer to Figure 4A). Values represent the mean ± SD of three independent experiments performed with different batches of purified proteins.

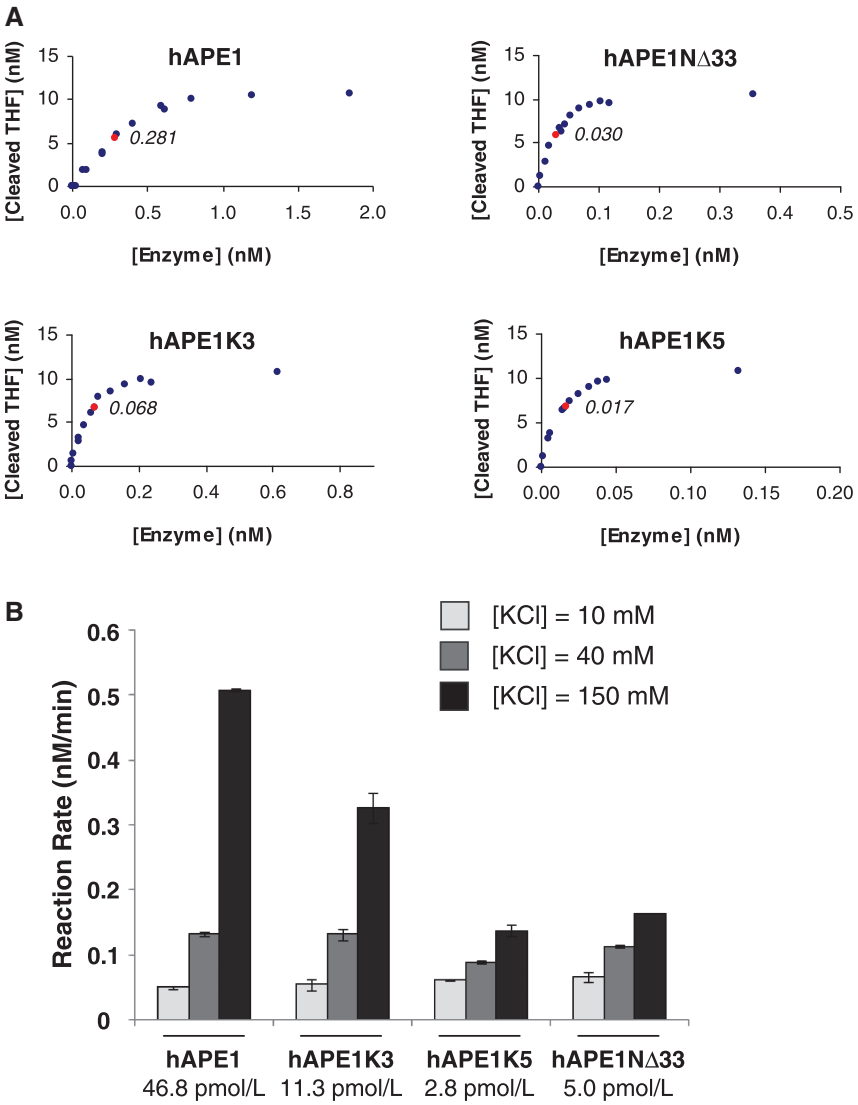

Figure 4. Lysine residues in the 23–33 basic region control the catalytic activity of hAPE1. (**A**) For each different hAPE1 proteins, dose-response experiments were performed in order to evaluate the most suitable concentration of enzyme to be used for the calculation of catalytic constants. An enzyme concentration capable of cleaving ~7 nM 34FDNA over the time of reaction, indicated by the red spots, was selected for time-lapse experiments. (**B**) Effect of K^+ concentration on the endonuclease activities of hAPE1 and of its mutants, on abasic DNA. Reaction rates, obtained as in (A), were quantitatively evaluated and the calculated values are plotted. The resulting histogram shows the average values with SD of three independent experiments performed with different batches of purified proteins.

Such a model would predict a differential sensitivity of the endonuclease activity of the enzyme to ionic strength variation for the different hAPE1 mutants here investigated. We therefore studied the effect of K^+ concentration on the catalytic activity of hAPE1 and of its mutants at concentrations of each enzyme capable of cleaving the same amount of 34FDNA (as performed for the experiments reported in Table 1). Data reported in Figure 4B clearly show that, as expected from previous results (43), the hAPE1 endonuclease activity was at least 4-fold higher in 150 mM KCl than in 40 mM KCl. Consistently with the proposed model, the catalytic activities of both K5 and NΔ33 protein mutants were essentially not affected by the increase in K^+ concentration. As expected, the K3 mutant showed an intermediate behaviour. Accordingly, further enzymatic assays, performed as a function of salt concentration and at the same protein concentrations, show that while hAPE1 displays a maximum activity at 150 mM KCl, hAPE1NΔ33 is more active at 80 mM KCl as the K5 mutant is (Supplementary Figure S3).

Altogether, these data support a model in which the electrostatic charges of lysine residues in the region 23–33 are involved in the modulation of hAPE1 catalytic activity, disclosing new unsuspected details in its functional regulation. They also suggest that molecular interactions, or even PTMs, involving that region, act as strong modulators of hAPE1 endonuclease activity.

Characterization of hAPE1/NPM1 binding and mapping of the amino acids involved in the interaction

We reported earlier that hAPE1 interacts *in vivo* with NPM1 and that this interaction determines the localization of hAPE1 in nucleoli (11). To evaluate the dissociation constant (K_D) of complexes formed between these proteins and their variants, SPR experiments were performed. The overlay of relative sensorgrams is reported in Figure 5A, left. The value of K_D of the hAPE1/NPM1 complex was estimated by plotting RUmax values, from each SPR-binding experiment, as a function of NPM1 concentration (Figure 5A, middle). Data fitting using a non-linear regression analysis provided a K_D value of $3.5 \pm 0.3\,\mu M$. Experiments performed under the same conditions on the immobilized hAPE1NΔ33 gave a poor signal variation (the overlay of the sensorgrams is reported in Figure 5A, right). The degree of mean RU_{max} and the lack of dose-response dependence confirmed the absence of a stable binding between hAPE1NΔ33 and NPM1, as reported earlier (11).

In line with previous findings, SPR experiments also indicated that the N-terminal domain of NPM1 (residues 1–117) is responsible for the binding to hAPE1. In this case, the calculated K_D value was $0.668 \pm 0.005\,\mu M$ (Supplementary Figure S4). Collectively, these data provided support to the hypothesis that the NPM1–hAPE1 interaction is physiologically relevant and that the 1–33 region of hAPE1 is responsible for it (11).

Then, we mapped the hAPE1 amino acids involved in NPM1 interaction through GST pull-down experiments using the hAPE1NΔ9, hAPE1NΔ22 and hAPE1NΔ33 deletion mutants. Similarly to what was observed for RNA binding, hAPE1 amino acids within region 23–33 is critical for the interaction with NPM1 (Figure 5B). Notably, removal of the first 22 amino acids, including the acidic stretch 10–22, increased over 2-fold the hAPE1 affinity to NPM1, thus suggesting that electrostatics plays a role in the interaction between the two proteins. GST pull-down experiments performed with the K to A mutants pointed out that, similarly to what seen for RNA binding, the presence of K31 and K32, in addition to K24, K25 and K27, is required for stable binding of hAPE1 to NPM1 (Figure 5C).

To confirm our *in vitro* experiments in an *in vivo* transfection assay, the two hAPE1K3 and hAPE1K5 mutants were expressed as Flag-tagged proteins in HeLa cells. hAPE1NΔ33 was used as a control. Notably, the two mutant proteins (i.e. K3 and K5), particularly the K5, resulted expressed at a slightly lower levels than the wild-type, suggesting that this mutation may alter protein stability. The hAPE1K5 mutant showed a behaviour similar to that of hAPE1NΔ33 (11), with a predominant cytoplasmic localization and no colocalization with NPM1 (Figure 5D, left). The hAPE1K3 mutant, on the contrary, showed an intermediate phenotype (Figure 5D, left). Co-immunoprecipitation analysis (Figure 5D, right) of transiently transfected HeLa cells confirmed the co-immunofluorescence data showing no interaction between hAPE1K5 and NPM1 as that of hAPE1NΔ33. We can therefore infer that the NLS sequence is not sufficient to guarantee for a complete nuclear localization of hAPE1 in the absence of charged residues at position K24, K25, K27, K31 and K32, thus suggesting that the RNA- and/or the NPM1-binding abilities may play a significant role in hAPE1 nuclear maintenance, as hypothesized earlier (11).

Taken together, these data provide the molecular basis for explaining the inhibitory effect of NPM1 on hAPE1 RNA binding observed previously (11). To better substantiate this, we performed competition experiments with NPM1 and hAPE1 in RNA-binding assays by GST pull-down analysis. Figure 5E shows that hAPE1 directly binds rRNA molecules, but the addition of soluble recombinant NPM1 in a 50% molar excess to the GST–hAPE1 bait inhibits hAPE1 binding to rRNA. This effect seems due to protein–protein interaction between hAPE1 and NPM1, since the rRNA-binding activity of GST-NPM1 for the 28S rRNA was lower than that of GST–hAPE1 alone. These data again support the notion that protein–protein interactions account for the inhibitory effect of NPM1 over hAPE1 binding to RNA, as suggested earlier (11).

As NPM1 stimulates hAPE1 endonuclease activity on 34FDNA substrates (11), we investigated whether these critical lysine residues may play a role in enzyme activation. Endonuclease assays under standard conditions at 40 mM KCl, either with or without NPM1, showed that this protein stimulates only the activity of the wild-type hAPE1 (Figure 5F) in a way similar to that described earlier (11). We then checked whether this stimulatory effect could be observed at higher K^+ concentration (150 mM KCl), where binding of hAPE1 to NPM1 still

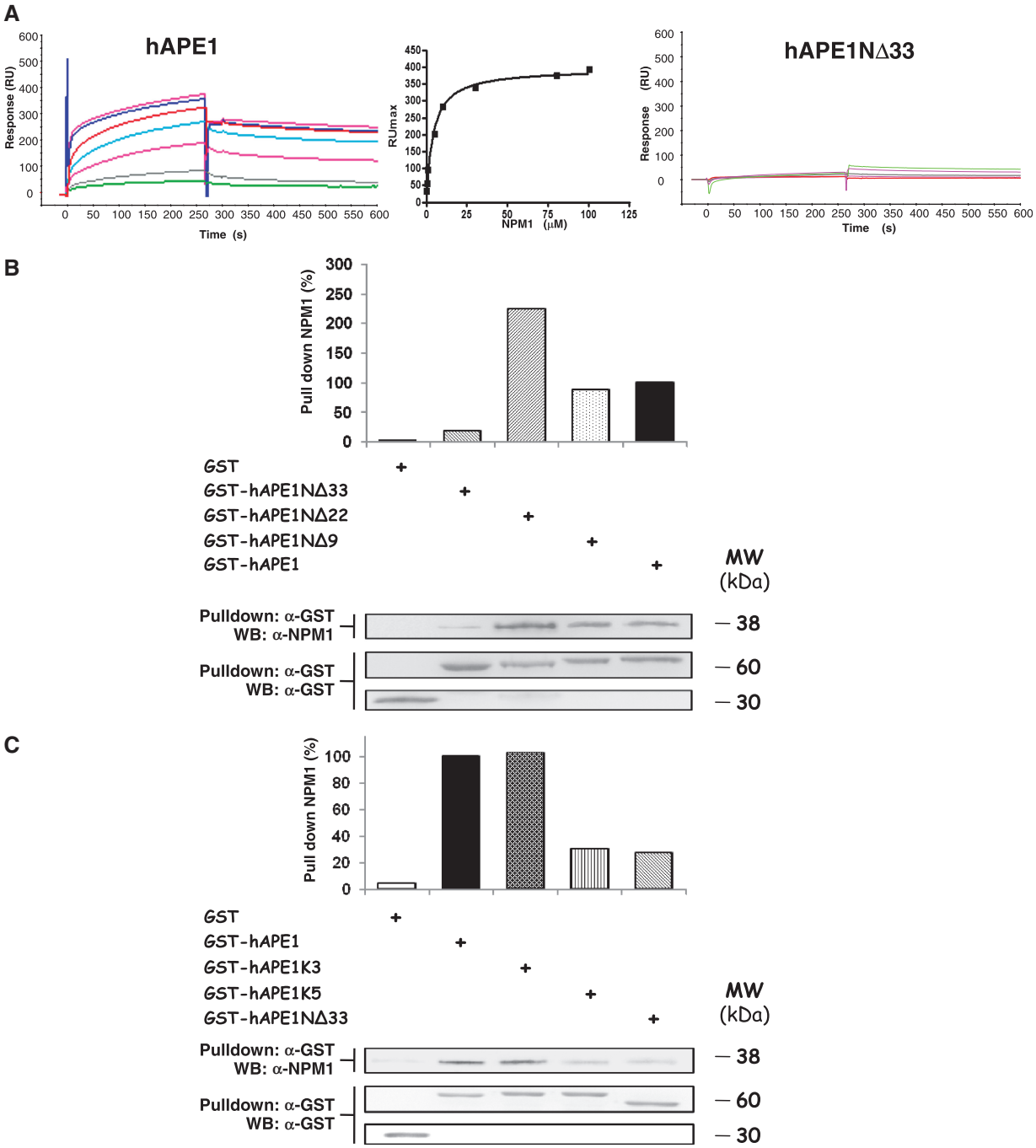

Figure 5. Mapping of the amino acids involved in hAPE1 binding to NPM1. (**A**, left) Overlay of SPR sensorgrams relative to the binding of NPM1 to immobilized hAPE1 (see 'Materials and Methods' section for experimental details). (A, middle) Plot of RU_{max} from each binding versus NPM1 concentrations (micro molar); data were fitted by non-linear regression analysis. (A, right) Overlay of SPR sensorgrams relative to the same experiment carried out on immobilized hAPE1NΔ33. (**B** and **C**) GST pull-down experiments with different hAPE1 deletion (B) or with K to A (C) mutants as baits and a recombinant NPM1 as prey. Pull-down samples underwent western blot analysis using either an anti-GST antibody or an anti-NPM1 antibody. The NPM1 to GST signal intensity ratio for each sample is plotted in the histogram. (**D**) K24, K25, K27, K31 and K32 residues of hAPE1 are necessary for protein nucleolar localization and binding to NPM1. HeLa cells were transfected with pcDNA5.1 expression vector containing the cDNA of Flag-tagged hAPE1 or that of hAPE1NΔ33, hAPE1K3 or hAPE1K5 mutants. Confocal microscopy (left) of HeLa cells fixed and stained with antibodies against NPM1 (red) and ectopic Flag-tagged hAPE1 (green). Total cell extracts from transiently transfected HeLa cells with empty vector, hAPE1, hAPE1NΔ33, hAPE1K3 or hAPE1K5 mutants were co-immunoprecipitated. Co-immunoprecipitated material was separated onto 10% SDS–PAGE and analysed by western blotting to evaluate the levels of NPM1 binding by using the specific antibody. (**E**) hAPE1 binding to 28S rRNA is inhibited by NPM1. GST pull-down experiments with purified hAPE1 were carried out using 100 pmol of each bait protein incubated with the indicated amount of total RNA from HeLa cells. Pull-down samples were analysed as described in Figure 2B, The amounts of 28S rRNA recovered after pull-downs were then normalized versus the amount of bait proteins. The resulting histogram shows the average values with SD of three independent experiments. (**F**) Effect of K^+ concentration on the stimulation of the endonuclease activities of hAPE1 and its mutants induced by NPM1. For each recombinant hAPE1 protein, 20 pM of each enzyme were incubated with or without 0.4 nM NPM1 and reacted, in either 40 or 150 mM KCl, with the 34FDNA as described in 'Materials and Methods' section. Then, reaction mixtures were separated in denaturing gels, which were then scanned for quantification. Histograms, representing the fold of increase in the reaction rates of each hAPE1 recombinant proteins, are the mean of two independent experiments.

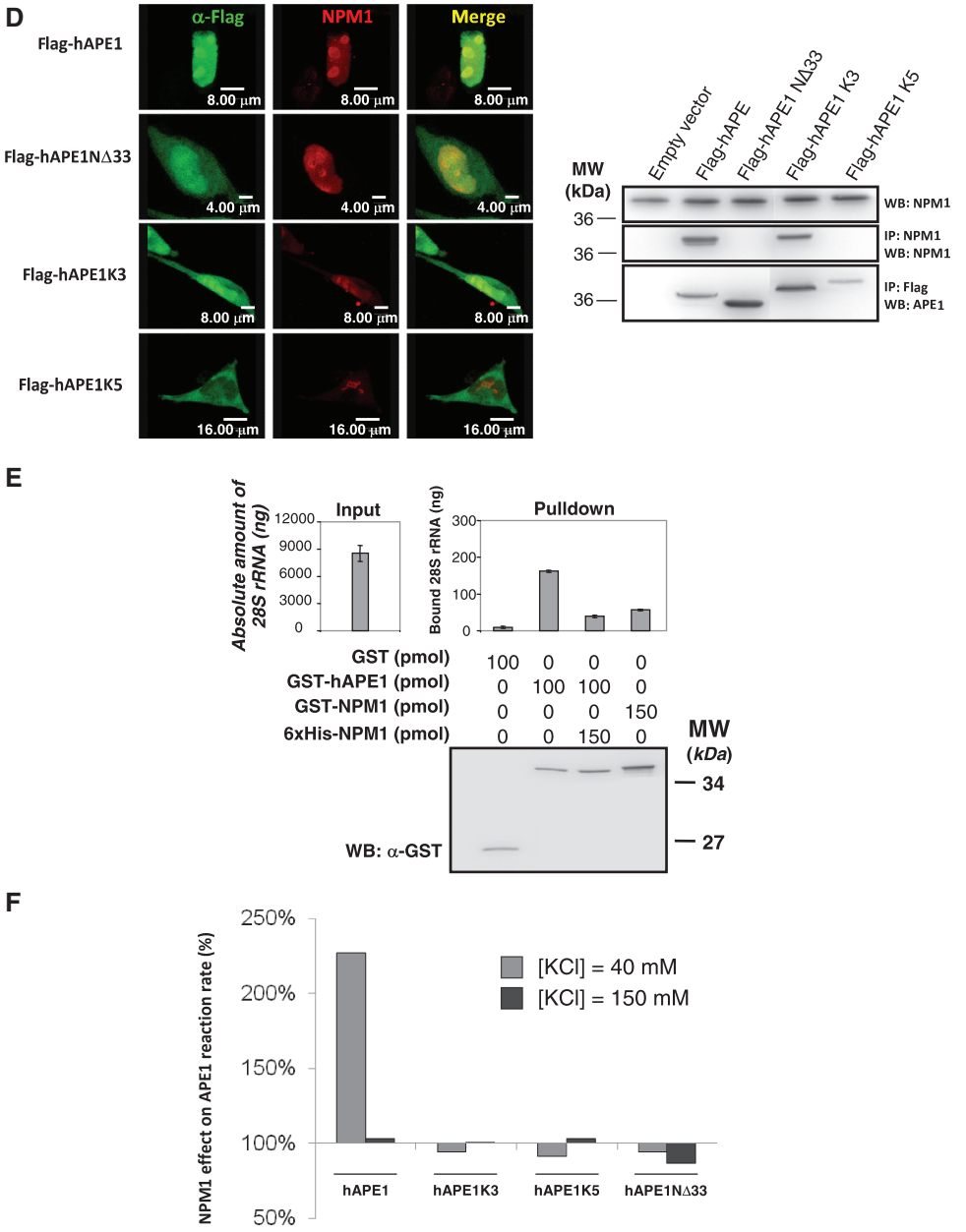

Figure 5. Continued.

occurs (11). As expected, no effect was observed for any of the enzymes tested. Thus, these results support the hypothesis that the hAPE1 region 23–33, through electrostatics, behaves as a site of allosteric regulation of the enzyme endonuclease activity.

hAPE1 *in vivo* acetylation at K27, K31, K32 and K35

K to A mutation mimicks constitutive acetylation *in vivo*. To probe whether the critical lysines we studied in the present work are acetylated *in vivo*, Flag-tagged hAPE1 from stably expressing HeLa cells (11) was immunopurified, excised from SDS–PAGE, S-alkylated, digested with endoprotease AspN, and subjected to massive nanoLC–ESI–LIT–MS/MS and MALDI–TOF–MS experiments. The protease was chosen to generate peptides having suitable dimension/properties for optimal PTMs mapping analysis of the 15–35 amino acid region.

In addition to a number of non-modified peptides, nanoLC–ESI–LIT–MS/MS analysis revealed the occurrence of specific acetylated species bearing a variable modification extent. In particular, peptides (15–33)Ac, (15–33)Ac2, (15–33)Ac3, (15–35)Ac, (15–35)Ac3 and (15–35)Ac4 at m/z 2186.4, 2228.3, 2270.4, 2429.1, 2513.3 and 2555.5 (theoretical value: 2186.2, 2228.2, 2270.2, 2429.3, 2513.3 and 2555.3), respectively, were detected in the total ion current MS trace profile of the hAPE1 digest.

Strong signals corresponding to the non-acetylated peptide counterparts were also observed (data not shown). Confirmation of the peptide assignment and direct localization of the acetylation sites was obtained from MS/MS data. As representative examples, the fragmentation spectra of peptides (15–33)Ac3 and (15–35)Ac4 are presented in Figure 6. Accordingly, nanoLC–ESI–LIT–MS/MS experiments demonstrated that hAPE1 is acetylated at K27, K31, K32 and K35. Based on the data from MALDI–TOF–MS analysis that, combined to nanoLC–ESI–LIT–MS/MS, allowed coverage of 68% of the protein sequence, no additional PTMs were observed within this peptide region. Furthermore, the reliability of our approach was substantiated by the analysis of the hAPE1 tryptic digest, which confirmed acetylation occurring at K6 and K7 (data not shown), as demonstrated earlier (8).

Evolutionary analysis of APE1N-terminal sequence

While being highly conserved in mammals, the APE1 33N-term is poorly conserved in other organisms including some vertebrates (5,17). We sought divergences in APE1 33N-terms within vertebrates in order to correlate them with protein binding to RNA and endonuclease activities. An evolutionary tree depicting the relationship between the vertebrate sequences is shown in Figure 7A. The sequences from amphibians and fishes represent a separate branch from mammalian ones. As reported earlier for the Cys content analysis, it turned out that, based on the N-terminal portion of the protein, zebrafish APE1 (zAPE1) is evolutionarily the most distant sequence from hAPE1 within vertebrate sequences (17). Alignment analysis (Figure 7B) suggests that during phylogenesis, the first three lysine residues (i.e. K24, K25 and K27) were conserved, while the other two (i.e. K31 and K32), occurring in mammals and *Xenopus*, might come from an insertion. Thus, fishes' APE1, which lacks these lysine residues, should act as a loss-of-function protein with regard to its RNA-binding. To test this hypothesis, we analysed the RNA-binding activity of zAPE1 in comparison with that of hAPE1. EMSA (Figure 7C, left), performed with purified recombinant proteins (Supplementary Figure S5), clearly demonstrated that zAPE1 binds to RNA at a significantly lower extent than hAPE1. We consequently evaluated the endonuclease activity of zAPE1 on abasic dsDNA in comparison to that of hAPE1 and of the different K to A mutants in a set of dose-response experiments. As it is evident in Figure 7C, right, zAPE1 endonuclease activity was higher than that of hAPE1 and similar to that of the K to A and of the NΔ33 deletion mutants. Therefore, both the RNA-binding and the endonuclease activities of zAPE1 are different from that of hAPE1 and resemble those of the mutants we tested, strongly supporting the evolutionary-based hypothesis that the K31 and K32 evolutive insertion may represent a phylogenetic acquisition of a further point for functional regulation of the protein through PTMs. Domain swapping experiments are actually in progress to further substantiate this hypothesis.

DISCUSSION

Although some data suggest that hAPE1N-terminal region is important for recruiting proteins involved in BER processes (20) or in RNA metabolism (11), current opinion confines its value to the presence of the NLS sequence and of a pair of lysine residues (K6 and K7) whose acetylation controls the expression of some genes (8,32). Based on the experimental evidences that: (i) the truncated form of hAPE1 and its K to A mutants in the N-terminal domain, have higher endonuclease activity than the wild-type protein on abasic dsDNA (this study); (ii) hAPE1NΔ33 does not cleave 34FRNA (11); (iii) hAPE1NΔ33, as well as hAPE1 K5 mutant shows a significantly reduced ability to bind RNA molecules and NPM1 with respect to hAPE1 [(11) and this study]; (iv) the behaviour of K3 mutant on both binding to RNA and NPM1 and endonucleolytic activity is intermediate between that of the wild type and that of hAPE1NΔ33 (this study), we propose that the hAPE1 first 33 amino acids might have an important role in regulating its activities and the regulatory function of the K residues within the region 23–33 may require differential and non-mutually exclusive modulation.

hAPE1 full-length is endowed with distinct domains capable of binding oligonucleotides through electrostatic interactions (15). In addition to the positively charged residues within the catalytic domain that drive the binding of abasic nucleotides to the active site, the recognition of substrates and/or reaction product also involves residues within the 33N-term. Although positively charged residues are abundant in the 33N-term—eight lysines and two arginines conferring to the whole protein a net positive charge—results presented in this study clearly indicate that among them, a specific subset of lysine residues, located in the portion 23–33, plays an active role in hAPE1 catalysis.

The experimental data reported here on the role of the 33N-term in hAPE1 catalysis may be tentatively interpreted in structural terms by assuming that the binding of apurinic/apyrimidinic dsDNA is essentially driven by the residues of the catalytic site. On the other hand, the reduced affinity of the hAPE1 catalytic site for other DNA structures, such as reaction product or single stranded RNA/DNA molecules, makes the role of the 33N-term relevant. Its involvement in oligonucleotide recognition is not surprising, as Gly41 is not far (<8 A) from a 15-bp DNA duplex containing an abasic site (Supplementary Figure S6), when this nucleotide is complexed with a truncated (NΔ40) form of APE1 (15). Within this framework, the decreased catalytic activity of hAPE1 on dsDNA compared to that exhibited by hAPE1NΔ33 should be ascribed to the contribution of the 33N-term to the electrostatic binding of the reaction product, being consistent with the increased activity shown by hAPE1K5 and with the effect of KCl concentration on the hAPE1 activity and not on that of the mutants hAPE1K5 and hAPE1NΔ33. The hypothesis is supported by other studies, which show that the catalytic activity of hAPE1 on abasic RNA/ssDNA (where the substrate binding is markedly dependent on the presence of

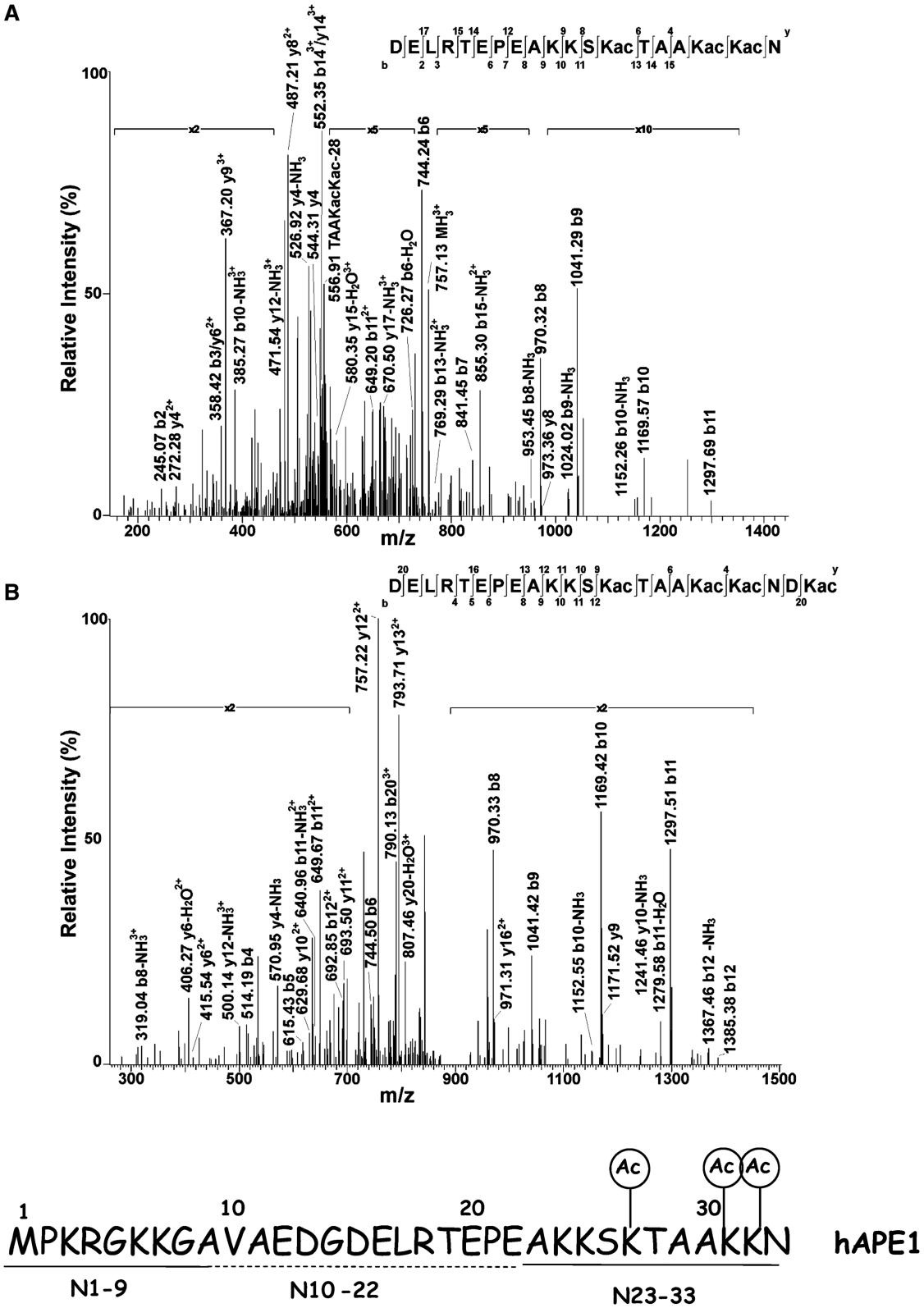

Figure 6. Lysine residues in the 23–33 basic region are acetylated *in vivo*. (**A**) Tandem mass spectrometry of some acetylated peptides present in the endoprotease AspN digest of hAPE1 purified from HeLa cells (see 'Materials and Methods' section for details). The fragmentation mass spectrum of the quadruply charged ion at m/z 568.5 associated with the peptide (15–33)Ac3 is shown. (**B**) The fragmentation mass spectrum of the triply charged ion at m/z 852.5 associated with the peptide (15–35)Ac4 in shown. Spectral assignment demonstrated that hAPE1 is acetylated at K27, K31, K32 and K35, as highlighted in the peptide sequence associated to each panel.

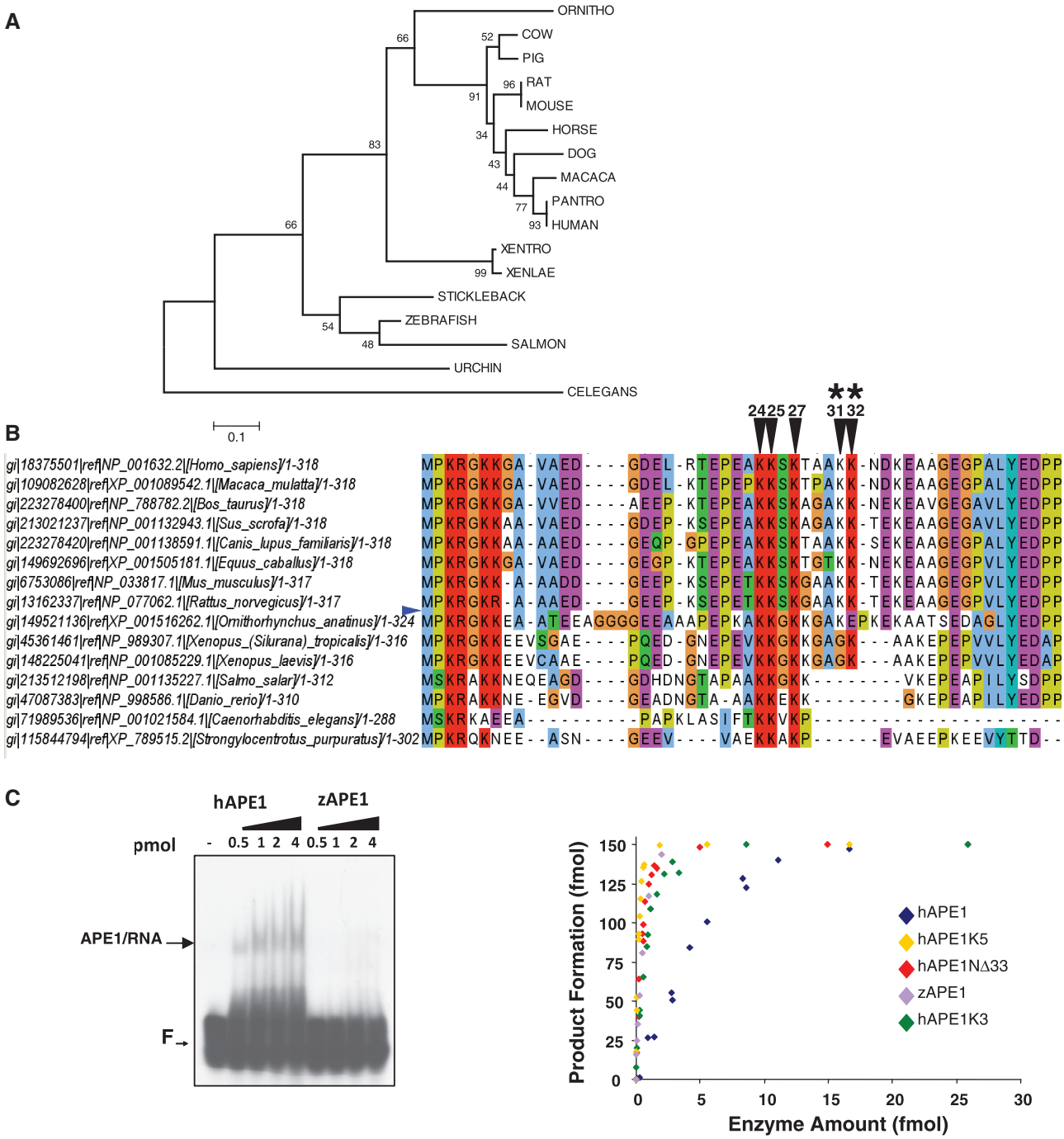

Figure 7. Evolutionary analysis of APE1N-terminal sequence. (**A**) Evolutionary tree for vertebrate APE1 sequences based on the analysis of amino acid sequences aligned with the first 60 amino acids in hAPE1. Numbers indicate the bootstrap values obtained during the construction of the tree. Mammalian and non-mammalian vertebrate APE1 sequences are found in distinct branches with *C. elegans* representing the sequence most evolutionarily distant from the human one. HUMAN, *H. sapiens*; PANTRO, *P. troglodytes*; MACACA, *M. mulatta*; HORSE, *E. caballus*; PIG, *S. scrofa*; COW, *B. taurus*; DOG, *C. lupus familiaris*; ORNITHO, *O. anatinus*; CELEGANS, *C. elegans*; URCHIN, *S. purpuratus*; MOUSE, *M. musculus*; RAT, *R. norvegicus*; XENLAE, *X. laevis*; XENTRO, *X. tropicalis*; SALMON, *S. salar*; ZEBRAFISH, *D. rerio*; STICKLEBACK, *G. aculeatus*. (**B**) Multiple sequence alignment of N-term APE1 sequences. Multiple alignment was performed using the software MAFFT (v6.531b) (35), available at the EMBL-EBI server (www.ebi.ac.uk/Tools/sequence.html). Arrows indicate K24, K25, K27, K31 and K32, respectively. (**C**, left) zAPE1 has lower affinity for RNA than hAPE1, as determined by EMSA. Reactions were performed as described in 'Materials and Methods' section with homogeneously purified recombinant proteins. The mixtures were separated on a native (non-denaturing) 8% (w/v) polyacrylamide gel at 120 V for 2 h. In each reaction, 2.5 pmol of ssFRNA were used. F indicates the free 34FRNA oligonucleotide ^{32}P-labelled probe. A representative EMSA gel image of three independent experiments, using the abasic ssFRNA sequence, is shown. (C, right) zAPE1 is catalytically more active than hAPE1 over abasic DNA. AP endonuclease activity of different mutant APE1 enzymes was tested by performing a dose-response experiment, as described in the 'Materials and Methods' section.

the 33N-term) is strongly reduced >50 mM KCl, whereas that on dsDNA is high even at 200 mM (43–45).

Substrate recognition, cleavage and product release are separable molecular steps. The stimulation of DNA Pol β over hAPE1 activity, through elimination of the incision product, strongly suggests that release of the catalysis product is the limiting step in the removal of AP-sites by hAPE1 (15,46). Interestingly, Marenstein *et al.* (47) showed no product inhibition effect on the rate of hydrolysis of abasic ssDNA, while such an inhibition was observed for the corresponding dsDNA, thus suggesting that dsDNA product can make the difference in being more strongly bound to the N-terminal disordered arm. Altogether, our results point to the fact that acetylation of the critical lysine residues located in the 23–33 protein sequence and involved in inhibition of both NPM1 and RNA interaction, is also very important in modulating hAPE1 endonuclease activity, presumably by preventing the enzyme N-terminal region to assume an electrostatically stabilized conformation associated to a reduced endonuclease activity. Therefore, we can speculate that the N-terminal tail of hAPE1 may regulate the overall enzyme catalytic turnover, reducing the rate of product release without appreciably affecting the AP-site recognition specificity of the enzyme. The possibility that the product release step may be regulated *in vivo* by PTMs or interactions with other proteins, such as NPM1 (11), CSB (21) or XRCC1 (20), could be viewed in the light of the necessity to coordinate the BER activities on different substrates in order to protect the abasic DNA repair intermediate and to ensure its coordinated processing.

Notably, some of the lysine residues (i.e. K27, K31 and K32) involved in the interaction with NPM1 and with RNA and responsible for modulating the catalytic activity of the enzyme, were found acetylated *in vivo* (Figure 6), thus providing further biological relevance to our findings. Based on the evidence that mutations inserted to mimic constitutive acetylation (the K to A mutants) progressively increase the catalytic activity of hAPE1 on abasic dsDNA, the occurrence of such specific PTMs may well explain the observation made by different laboratories that immunopurified hAPE1 forms from mammalian cells retains significantly higher enzymatic activity than that of recombinant proteins expressed in *E. coli*. From our data we can also infer that simultaneous multiple acetylation is required to modulate the different hAPE1 known activities.

Taken together and in light of previous works (11,27,48), our current data allow to propose the molecular model of APE1 functions reported in Figure 8. After APE1 scans nucleic acids in search of its substrate and catalyzes the cleavage of the abasic site, acetylation of the critical lysine residues (i.e. K27, K31 and K32) should stimulate the overall hAPE1 endonuclease activity through increasing the product release rate. Specific acetylation may also modulate the interaction of hAPE1 with NPM1 (and possibly with other unidentified protein partners) and thus could redirect a functionally active form of hAPE1 from buffering storage sites, such as nucleoli, toward the active sites of DNA repair. Moreover, acetylation of the critical lysine residues may also inhibit hAPE1 action in RNA metabolism by preventing its RNA-binding activity. Notably, a recent paper provided the first evidence for ubiquitination occurring on K24, K25 and K27 (31). Interestingly, an ubiquitin-fused hAPE1 recombinant protein, in which the ubiquitin sequence is sandwiched between A23 and K24 (mimicking constitutively ubiquitinated hAPE1) was clearly excluded from nuclei (31). Thus, it can be speculated that post-translational modifications occurring on or near these critical lysine residues may play a significant role in controlling hAPE1 subcellular distribution, its functional activation and thus its biological function in mammalian cells.

From an evolutionary perspective, hAPE1 activity was made tunable (through post-translational modifications or the interaction with other proteins) or expanded toward other substrates with the acquisition of a positively charged N-terminus, without major modifications of its catalytic site, which retained the 'canonical' function towards abasic dsDNA. The existence of such an N-terminal extension for mammals only (Figure 7), may suggest its evolutionary significance in facing increased functional complexity. The intrinsic lack of structure of this domain can confer functional advantages to hAPE1, including the ability to bind to several different targets (i.e. NPM1, XRCC1, CSB, RNA, etc.), thus allowing precise control over the thermodynamics of the binding process (49). Since the 33N-term is required for both stabilizing the interaction to RNA and controlling the overall hAPE1 enzymatic activity through decreasing the rate of product release, this double role provides a 'relay' for molecular regulation with extremely useful biological significance and could also be envisioned in the light of preventing formation of harmful unprotected DNA strand breaks. Interestingly, through multiple interactions with different protein partners, the enzymatic activity of hAPE1 could be finely tuned, in association with the biological circumstances, in a time-dependent manner. While this work was in progress, a article by Yu *et al.* demonstrated that the K residues within the 24–35 region, which have a differential protection pattern between free and DNA-bound complexes of APE1, are involved in a conformational change concomitant with DNA binding and catalysis and with interaction with Pol β (50). This work and our current data strongly support the notion that these critical residues may indeed constitute a structural switch for hAPE1 functional regulation.

Regulation of the interaction of hAPE1 with NPM1 (11) may have profound biological consequences during cell response to stressor signals. NPM1 physically interacts with p53 and is believed to regulate this tumor suppressor protein (51). p53-signaling network, and its involvement of ubiquitin in particular, plays an important role in APE1 functional regulation, as recently demonstrated (31). During DNA damage stress response, ARF dissociates from NPM1 within nucleoli and relocalizes to the nucleoplasm where it sequesters the p53-ubiquitinating enzyme MDM2, thus leading to p53 up-regulation and functional activation (52). These data support the hypothesis that nucleolus may represent a major cellular stress sensor transmitting signals to the

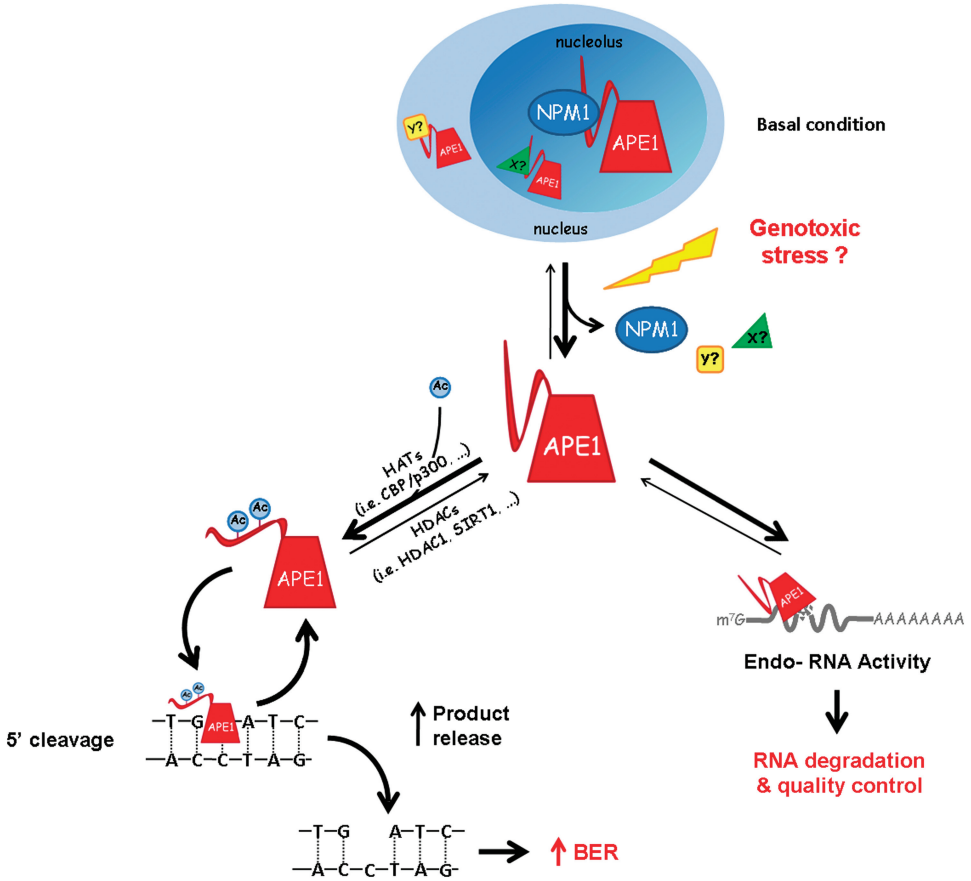

Figure 8. Proposed mechanism of regulation of hAPE1 functions and compartmentalization through acetylation of critic lysine residues. Within cells, APE1 is present with different degrees of acetylation on residues 27, 31, 32, 35 (Ac-APE1) and may also be involved in interaction with NPM1 and/or different protein partners (indicated here as 'X' and 'Y'). Genotoxic stress may shift the equilibrium between the non-acetylated and the Ac-APE1 forms toward these latter and more active form, possibly by its reduced affinity for the incised DNA product. Acetylation, in turn, may inhibit binding to RNA substrates thus, possibly, this PTM may change the equilibrium of the hAPE1 pool from the RNA-bound to the DNA-bound substrates. Moreover, acetylation may also control compartmentalization of hAPE1 within specific subnuclear structures (such as nucleoli), where the enzyme could be stored once bound to interacting partners as NPM1. Whether NPM1 may specifically participate in DNA repair inside of nucleoli, or even outside of the subnuclear condensed area, is currently unknown. In the first case, the genome containing rRNA gene may be a favoured substrate for hAPE1 *in vivo*. On the other hand, if the latter is the case, stimuli triggering NPM1 to get out of nucleoli may affect the cell response in terms of hAPE1 impact on genome stability. Further experiments are currently underway to address these issues.

system for regulation of p53 activity (53). Thus, a careful elucidation of the p53-signalling networks and their involvement in the DNA repair pathway coordinated by hAPE1 is an important subject for molecular carcinogenesis and deserves further study. To this aim, work is now in progress in our laboratory to study the modulation of the hAPE1 interactome network during genotoxic damage and to evaluate its impact on the dynamics of nucleolar hAPE1 distribution. In conclusion, our work underscores the functional importance of the neglected unstructured hAPE1N-terminal domain and points to the need of structural details on this protein region for understanding the fine-tuning mechanisms regulating this multi-functional protein.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

AKNOWLEDGEMENTS

Authors thank Raphaël Guérois for his help with the interpretation of protein multiple sequence alignment, Dr Mark R. Kelley for zAPE1 expressing vector and Dr David Wilson III for critically reading the article and helpful suggestions.

FUNDING

MIUR (FIRB RBRN07BMCT_008, MERIT and PRIN 2008CCPKRP_002/003, to G.T. and A.S.); Galileo Project 2008 (to G.T. and J.P.R.); Association pour la Recherche sur le Cancer (ARC1118, to J.P.R.); EU-US bilateral project from Italian Ministry of Foreign Affairs (to G.T.). Funding for open access charge: MIUR (FIRB RBRN07BMCT_008 and PRIN 2008CCPKRP_002/003 (to G.T.).

Conflict of interest statement. None declared.

REFERENCES

1. Fung,H. and Demple,B. (2005) A vital role for Ape1/Ref1 protein in repairing spontaneous DNA damage in human cells. *Mol. Cell*, **17**, 463–470.
2. Izumi,T., Brown,D.B., Naidu,C.V., Bhakat,K.K., Macinnes,M.A., Saito,H., Chen,D.J. and Mitra,S. (2005) Two essential but distinct functions of the mammalian abasic endonuclease. *Proc. Natl Acad. Sci. USA*, **102**, 5739–5743.
3. Vascotto,C., Cesaratto,L., Zeef,L.A., Deganuto,M., D'Ambrosio,C., Scaloni,A., Romanello,M., Damante,G., Taglialatela,G., Delneri,D. *et al.* (2009) Genome-wide analysis and proteomic studies reveal APE1/Ref-1 multifunctional role in mammalian cells. *Proteomics*, **9**, 1058–1074.
4. Tell,G., Damante,G., Caldwell,D. and Kelley,M.R. (2005) The intracellular localization of APE1/Ref-1: more than a passive phenomenon? *Antioxid. Redox Signal.*, **7**, 367–384.
5. Tell,G., Quadrifoglio,F., Tiribelli,C. and Kelley,M.R. (2009) The many functions of APE1/Ref-1: not only a DNA-repair enzyme. *Antioxid. Redox Signal.*, **11**, 601–620.
6. Xanthoudakis,S., Miao,G.G. and Curran,T. (1994) The redox and DNA-repair activities of Ref-1 are encoded by nonoverlapping domains. *Proc. Natl Acad. Sci. USA*, **91**, 23–27.
7. Izumi,T. and Mitra,S. (1998) Deletion analysis of human AP-endonuclease: minimum sequence required for the endonuclease activity. *Carcinogenesis*, **19**, 525–527.
8. Bhakat,K.K., Izumi,T., Yang,S.H., Hazra,T.K. and Mitra,S. (2003) Role of acetylated human AP-endonuclease (APE1/Ref-1) in regulation of the parathyroid hormone gene. *EMBO J.*, **1**, 6299–6309.
9. Chung,U., Igarashi,T., Nishishita,H., Iwanari,H., Iwamatsu,A., Suwa,T., Mimori,T., Hata,K., Ebisu,S., Ogata,E. *et al.* (1996) The interaction between Ku antigen and REF1 protein mediates negative gene regulation by extracellular calcium. *J. Biol. Chem.*, **271**, 8593–8598.
10. Kuninger,D.T., Izumi,T., Papaconstantinou,J. and Mitra,S. (2002) Human AP-endonuclease 1 and hnRNP-L interact with a nCaRE-like repressor element in the AP-endonuclease 1 promoter. *Nucleic Acids Res.*, **30**, 823–829.
11. Vascotto,C., Fantini,D., Romanello,M., Cesaratto,L., Deganuto,M., Leonardi,A., Radicella,J.P., Kelley,M.R., D'Ambrosio,C., Scaloni,A. *et al.* (2009) APE1/Ref-1 interacts with NPM1 within nucleoli and plays a role in the rRNA quality control process. *Mol. Cell. Biol.*, **29**, 1834–1854.
12. Tell,G., Wilson,D.M. 3rd and Lee,C.H. (2010) Intrusion of a DNA repair protein in the RNome world: is this the beginning of a new era? *Mol. Cell. Biol.*, **30**, 366–371.
13. Barnes,T., Kim,W.C., Mantha,A.K., Kim,S.E., Izumi,T., Mitra,S. and Lee,C.H. (2009) Identification of Apurinic/apyrimidinic endonuclease 1 (APE1) as the endoribonuclease that cleaves c-myc mRNA. *Nucleic Acids Res.*, **37**, 3946–3958.
14. Strauss,P.R. and Holt,C.M. (1998) Domain mapping of human apurinic/apyrimidinic endonuclease. Structural and functional evidence for a disordered amino terminus and a tight globular carboxyl domain. *J. Biol. Chem.*, **273**, 14435–14441.
15. Mol,C.D., Izumi,T., Mitra,S. and Tainer,J.A. (2000) DNA-bound structures and mutants reveal abasic DNA binding by APE1 and DNA repair coordination. *Nature*, **403**, 451–456.
16. Gorman,M.A., Morera,S., Rothwell,D.G., de La Fortelle,E., Mol,C.D., Tainer,J.A., Hickson,I.D. and Freemont,P.S. (1997) The crystal structure of the human DNA repair endonuclease HAP1 suggests the recognition of extra-helical deoxyribose at DNA abasic sites. *EMBO J.*, **16**, 6548–6558.
17. Georgiadis,M.M., Luo,M., Gaur,R.K., Delaplane,S., Li,X. and Kelley,M.R. (2008) Evolution of the redox function in mammalian apurinic/apyrimidinic endonuclease. *Mutat. Res.*, **643**, 54–63.
18. Reardon,B.J., Lombardo,C.R. and Sander,M. (1998) Drosophila Rrp1 domain structure as defined by limited proteolysis and biophysical analyses. *J. Biol. Chem.*, **273**, 33991–33999.
19. Takeuchi,R., Ruike,T., Nakamura,R., Shimanouchi,K., Kanai,Y., Abe,Y., Ihara,A. and Sakaguchi,K. (2006) Drosophila DNA polymerase zeta interacts with recombination repair protein 1, the Drosophila homologue of human abasic endonuclease 1. *J. Biol. Chem.*, **281**, 11577–11585.
20. Vidal,A.E., Boiteux,S., Hickson,I.D. and Radicella,J.P. (2001) XRCC1 coordinates the initial and late stages of DNA abasic site repair through protein-protein interactions. *EMBO J.*, **20**, 6530–6539.
21. Wong,H.K., Muftuoglu,M., Beck,G., Imam,S.Z., Bohr,V.A. and Wilson,D.M. 3rd (2007) Cockayne syndrome B protein stimulates apurinic endonuclease 1 activity and protects against agents that introduce base excision repair intermediates. *Nucleic Acids Res.*, **35**, 4103–4113.
22. Fan,Z., Beresford,P.J., Zhang,D., Xu,Z., Novina,C.D., Yoshida,A., Pommier,Y. and Lieberman,J. (2003) Cleaving the oxidative repair protein Ape1 enhances cell death mediated by granzyme A. *Nat. Immunol.*, **4**, 145–153.
23. Guo,Y., Chen,J., Zhao,T. and Fan,Z. (2008) Granzyme K degrades the redox/DNA repair enzyme Ape1 to trigger oxidative stress of target cells leading to cytotoxicity. *Mol. Immunol.*, **45**, 2225–2235.
24. Jackson,E.B., Theriot,C.A., Chattopadhyay,R., Mitra,S. and Izumi,T. (2005) Analysis of nuclear transport signals in the human apurinic/apyrimidinic endonuclease (APE1/Ref1). *Nucleic Acids Res.*, **33**, 3303–3312.
25. Chattopadhyay,R., Wiederhold,L., Szczesny,B., Boldogh,I., Hazra,T.K., Izumi,T. and Mitra,S. (2006) Identification and characterization of mitochondrial abasic (AP)-endonuclease in mammalian cells. *Nucleic Acids Res.*, **34**, 2067–2076.
26. Mitra,S., Izumi,T., Boldogh,I., Bhakat,K.K., Chattopadhyay,R. and Szczesny,B. (2007) Intracellular trafficking and regulation of mammalian AP-endonuclease 1 (APE1), an essential DNA repair protein. *DNA Repair*, **6**, 461–469.
27. Tell,G., Crivellato,E., Pines,A., Paron,I., Pucillo,C., Manzini,G., Bandiera,A., Kelley,M.R., Di Loreto,C. and Damante,G. (2001) Mitochondrial localization of APE/Ref-1 in thyroid cells. *Mutat. Res.*, **485**, 143–152.
28. He,T., Weintraub,N.L., Goswami,P.C., Chatterjee,P., Flaherty,D.M., Domann,F.E. and Oberley,L.W. (2003) Redox factor-1 contributes to the regulation of progression from G0/G1 to S by PDGF in vascular smooth muscle cells. *Am. J. Physiol. Heart Circ. Physiol.*, **285**, 804–812.
29. Szczesny,B. and Mitra,S. (2005) Effect of aging on intracellular distribution of abasic (AP) endonuclease 1 in the mouse liver. *Mech. Ageing Dev.*, **126**, 1071–1078.
30. Grillo,C., D'Ambrosio,C., Scaloni,A., Maceroni,M., Merluzzi,S., Turano,C. and Altieri,F. (2006) Cooperative activity of Ref-1/APE and ERp57 in reductive activation of transcription factors. *Free Radic. Biol. Med.*, **41**, 1113–1123.
31. Busso,C.S., Iwakuma,T. and Izumi,T. (2009) Ubiquitination of mammalian AP endonuclease (APE1) regulated by the p53-MDM2 signaling pathway. *Oncogene*, **28**, 1616–1625.
32. Fantini,D., Vascotto,C., Deganuto,M., Bivi,N., Gustincich,S., Marcon,G., Quadrifoglio,F., Damante,G., Bhakat,K.K., Mitra,S. *et al.* (2008) APE1/Ref-1 regulates PTEN expression mediated by Egr-1. *Free Radic. Res.*, **42**, 20–29.
33. Parlanti,E., Locatelli,G., Maga,G. and Dogliotti,E. (2007) Human base excision repair complex is physically associated to DNA replication and cell cycle regulatory proteins. *Nucleic Acids Res.*, **35**, 1569–1577.
34. Altschul,S.F., Gish,W., Miller,W., Myers,E.W. and Lipman,D.J. (1990) Basic local alignment search tool. *J. Mol. Biol.*, **215**, 403–410.
35. Katoh,K. and Toh,H. (2008) Recent developments in the MAFFT multiple sequence alignment program. *Brief. Bioinform.*, **9**, 286–298.
36. Tamura,K., Dudley,J., Nei,M. and Kumar,S. (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol. Biol. Evol.*, **24**, 1596–1599.
37. Saitou,N. and Nei,M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.*, **4**, 406–425.
38. Fields,G.B. and Noble,R.L. (1990) Solid phase peptide synthesis utilizing 9-fluorenyl-methoxycarbonyl amino acids. *Int. J. Pept. Protein Res.*, **35**, 161–214.

39. Aletras,A., Barlos,K., Gatos,D., Koutsogianni,S. and Mamos,P. (1995) Preparation of the very acid-sensitive Fmoc-Lys(Mtt)-OH. Application in the synthesis of side-chain to side-chain cyclic peptides and oligolysine cores suitable for the solid-phase assembly of MAPs and TASPs. *Int. J. Pept. Protein Res.*, **45**, 488–496.
40. Johnsson,B., Löfås,S. and Lindquist,G. (1991) Immobilization of protein to a carboxymethyldextran modified gold surface for biospecific interaction analysis in SPR sensors. *Anal. Biochem.*, **198**, 268–277.
41. Calvanese,L., Saporito,A., Oliva,R., D'Auria,G., Pedone,C., Paolillo,L., Ruvo,M., Marasco,D. and Falcigno,L. (2009) Structural insights into the interaction between the Cripto CFC domain and the ALK4 receptor. *J. Pept. Sci.*, **15**, 175–183.
42. D'Ambrosio,C., Arena,S., Fulcoli,G., Scheinfeld,M.H., Zhou,D., D'Adamio,L. and Scaloni,A. (2006) Hyperphosphorylation of JNK-interacting protein 1, a protein associated with Alzheimer disease. *Mol. Cell. Proteomics*, **5**, 97–113.
43. Wilson,D.M. 3rd (2005) Ape1 abasic endonuclease activity is regulated by magnesium and potassium concentrations and is robust on alternative DNA structures. *J. Mol. Biol.*, **345**, 1003–1014.
44. Fan,J., Matsumoto,Y. and Wilson,D.M. 3rd (2006) Nucleotide sequence and DNA secondary structure, as well as replication protein A, modulate the single-stranded abasic endonuclease activity of APE1. *J. Biol. Chem.*, **281**, 3889–3898.
45. Berquist,B.R., McNeill,D.R. and Wilson,D.M. 3rd (2008) Characterization of abasic endonuclease activity of human Ape1 on alternative substrates, as well as effects of ATP and sequence context on AP site incision. *J. Mol. Biol.*, **379**, 17–27.
46. Masuda,Y., Bennett,R.A. and Demple,B. (1998) Dynamics of the interaction of human apurinic endonuclease (Ape1) with its substrate and product. *J. Biol. Chem.*, **273**, 30352–30359.
47. Marenstein,D.R., Wilson,D.M. 3rd and Teebor,G.W. (2004) Human AP endonuclease (APE1) demonstrates endonucleolytic activity against AP sites in single-stranded DNA. *DNA Repair*, **3**, 527–533.
48. Yamamori,T., DeRicco,J., Naqvi,A., Hoffman,T.A., Mattagajasingh,I., Kasuno,K., Jung,S.B., Kim,C.S. and Irani,K. (2010) SIRT1 deacetylates APE1 and regulates cellular base excision repair. *Nucleic Acids Res.*, **38**, 832–845.
49. Wright,P.E. and Dyson,H.J. (1999) Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J. Mol. Biol.*, **293**, 321–331.
50. Yu,E., Gaucher,S.P. and Hadi,M.Z. (2010) Probing conformational changes in Ape1 during the progression of Base Excision Repair. *Biochemistry*, **49**, 3786–3796.
51. Colombo,E., Marine,J.C., Danovi,D., Falini,B. and Pelicci,P.G. (2002) Nucleophosmin regulates the stability and transcriptional activity of p53. *Nat. Cell. Biol.*, **4**, 529–533.
52. Lee,C., Smith,B.A., Bandyopadhyay,K. and Gjerset,R.A. (2005) DNA damage disrupts the p14ARF-B23(nucleophosmin) interaction and triggers a transient subnuclear redistribution of p14ARF. *Cancer Res.*, **65**, 9834–9842.
53. Rubbi,C.P. and Milner,J. (2003) Disruption of the nucleolus mediates stabilization of p53 in response to DNA damage and other stresses. *EMBO J.*, **22**, 6068–6077.