2166–2175 *Nucleic Acids Research, 2005, Vol. 33, No. 7*
doi:10.1093/nar/gki510

Eukaryotic RNases H1 act processively by interactions through the duplex RNA-binding domain

Sergei A. Gaidamakov[1], Inna I. Gorshkova[1,2], Peter Schuck[2], Peter J. Steinbach[3], Hirofumi Yamada[1], Robert J. Crouch[1,*] and Susana M. Cerritelli[1]

[1]Laboratory of Molecular Genetics, National Institute of Child Health and Human Development, [2]Protein Biophysics Resource, Division of Bioengineering and Physical Science, Office of Research Services, Office of the Director and [3]Center for Molecular Modeling, Center for Information Technology, National Institutes of Health, Department of Health and Human Services, Bethesda, MD 20892, USA

Received February 17, 2005; Revised and Accepted March 25, 2005

ABSTRACT

Ribonucleases H have mostly been implicated in eliminating short RNA primers used for initiation of lagging strand DNA synthesis. *Escherichia coli* RNase HI cleaves these RNA–DNA hybrids in a distributive manner. We report here that eukaryotic RNases H1 have evolved to be processive enzymes by attaching a duplex RNA-binding domain to the RNase H region. Highly conserved amino acids of the duplex RNA-binding domain are required for processivity and nucleic acid binding, which leads to dimerization of the protein. The need for a processive enzyme underscores the importance in eukaryotic cells of processing long hybrids, most of which remain to be identified. However, long RNA–DNA hybrids formed during immunoglobulin class-switch recombination are potential targets for RNase H1 in the nucleus. In mitochondria, where RNase H1 is essential for DNA formation during embryogenesis, long hybrids may be involved in DNA replication.

INTRODUCTION

Ribonucleases H (RNases H) cleave RNA of RNA–DNA hybrids in a sequence non-specific manner. They have been divided into two classes (type 1/I and 2/II) based on amino acid sequence and biochemical properties (1). At least one type of RNase H is present in all organisms, although they are not essential in most prokaryotes (2) or in single-cell eukaryotes such as *Saccharomyces cerevisiae* (3) or *Crithidia fasciculata* (4). However, in more complex organisms, such as *Drosophila melanogaster* (5) or mouse (6), RNase H1 is required during development. In mouse, RNase H1 is necessary for mitochondrial DNA (mtDNA) formation during embryogenesis. We found that *Rnaseh1*-null embryos stop developing shortly after implantation and undergo apoptotic cell death as a result of their failure to produce new mtDNA. RNase H is also a component of the reverse transcriptase (RT) of HIV-1 and other retroviruses, where it is necessary for the production of infectious virus particles (7). The structures of *Escherichia coli* RNase HI (8,9) and the RNase H domain of HIV-1 RT (10) have been solved and shown to be very similar. Eukaryotic RNases H1 are related by sequence to *E.coli* RNase HI and HIV-1 RNase H, and presumably fold into a very similar 3D structure, an important consideration when developing drugs to target the retroviral enzyme due to the potential deadly side effects of inhibiting the essential cellular enzyme. Therefore, the characterization of the mammalian RNase H1 is key to the progress in developing effective drugs against the RNase H of HIV-1, as well as advancing our understanding of mtDNA biogenesis.

Like the HIV-1 RNase H, which is part of a multi-functional protein, eukaryotic RNases H1 are multi-domain enzymes, and more complex than their prokaryotic counterparts. At their N-terminus they have a double-stranded RNA, RNA–DNA Hybrid binding domain (dsRHbd), a highly conserved sequence of ~50 amino acids (11). A connection region that varies from species to species in length and amino acid composition joins the dsRHbd to the C-terminal RNase H domain. In higher eukaryotes, we have shown by transient transfection experiments that RNase H1 localizes both to the nucleus and mitochondria (6). There are two in-frame AUG codons to produce two protein isoforms with different subcellular localizations (Figure 1a). Translation from the first AUG includes a mitochondrial localization sequence (MLS), preceding the dsRHbd, which directs the protein to the mitochondria. The

*To whom correspondence should be addressed. Tel: +1 301 496 4082; Fax: +1 301 496 0243; Email: robert_crouch@nih.gov

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

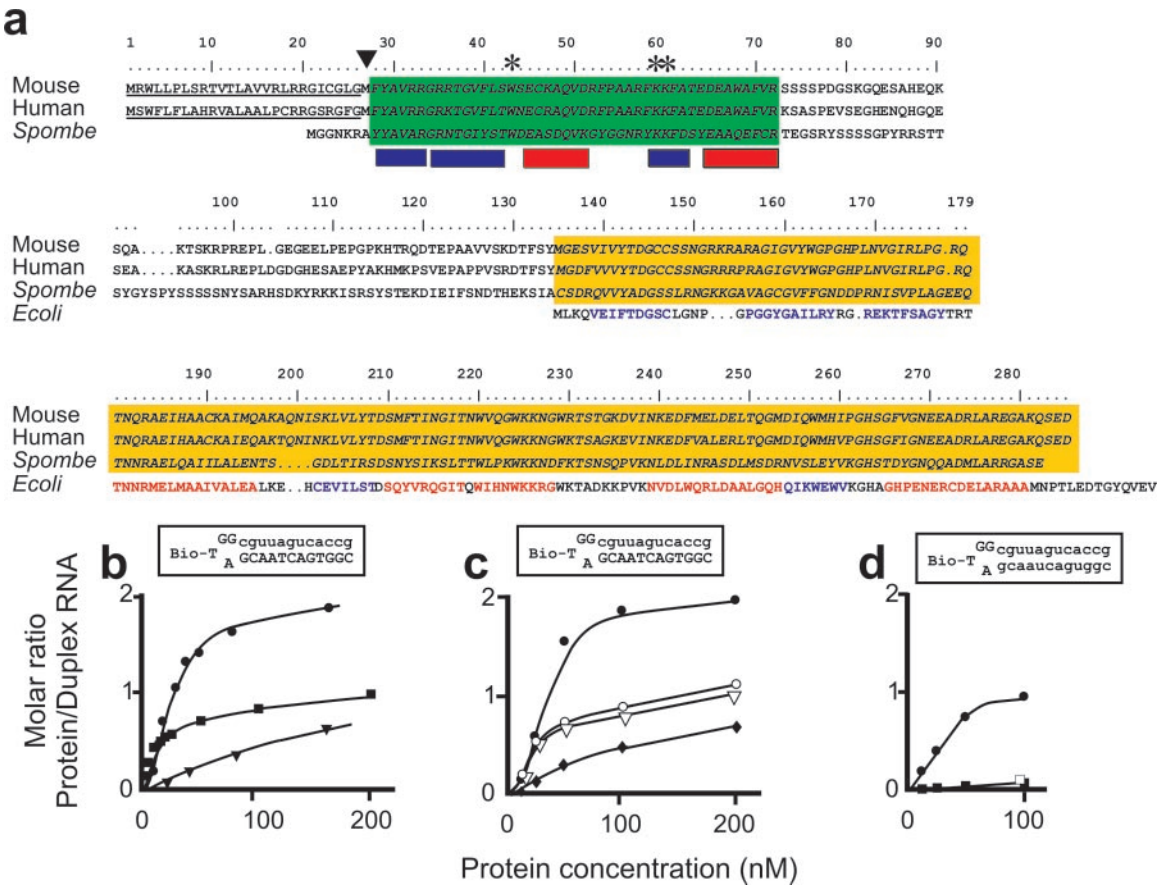


Figure 1. RNases H1 surface plasmon resonance analysis (**a**) RNases H1 of mouse, human, *S.pombe* and *E.coli* sequence alignment are shown. Numbering refers to the mouse protein. Mouse and human RNases H1 mitochondrial localization signals are underlined. The methionine at the start of the protein used in these studies is marked by an inverted solid triangle and amino acids W43, K59 and K60, marked by an asterisk, are the amino acids that were changed to A for RNase $H1^{W}$, RNase $H1^{KK}$ and RNase $H1^{WKK}$. RNase H^{HD} (RNase H domain, gold background) starts at amino acid 137 and ends at amino acid 285. The dsRHbd sequences are on a green background with the positions of the α helices (red boxes) and β strands (blue boxes) of *S.cerevisiae* RNase H1 first dsRHbd noted below the dsRHbd sequences. For the *E.coli* sequence, amino acids in α helices are red and β strands are blue. Increase in protein bound to duplex RNAs as a function of protein concentration. Some of the sensograms from which these data were obtained are presented in Supplementary Figure S1. Wild-type mouse RNase H1 data are represented by a solid black circle, *E.coli* RNase HI by a solid black square, RNase $H1^{HD}$ by an inverted solid black triangle, RNase $H1^{W}$ by an open circle, RNase $H1^{KK}$ by an inverted open triangle and RNase $H1^{WKK}$ by a solid black diamond. (**b**) and (**c**) are from data collected with the 12 bp RNA–DNA hairpin hybrid while the surface in (**d**) is a 12 bp RNA–RNA hairpin duplex RNA of the same sequence as the 12 bp RNA–DNA hybrid. The lack of binding of mouse RNase H1 to the dsRNA at 10 mM $MgCl_2$ is indicated in (d) by the open square. Nucleic acid sequences of the hairpin duplexes are shown above in (b–d) with RNA in lower case and DNA in upper case letters. Bio-T indicates the biotin modified dT to which the nucleic acid was attached to the streptavidin on the chip surface.

second AUG, after the MLS, initiates a protein that localizes exclusively to the nucleus. In our transient transfections we observed an unequal distribution of the two isoforms, with most of the protein present in the nucleus. The full-length form targeted to mitochondria is probably only transiently present *in vivo*, as most N-terminal MLS sequences are cleaved after they traverse the inner membrane to produce the mature protein (12).

The RNase H domain of eukaryotic RNases H1, when expressed alone, is able to hydrolyze RNA of RNA–DNA hybrids just as the single-domain homologous protein from *E.coli* does, i.e. without the dsRHbd (data presented here and ref. 13). Similarly, the N-terminal region of *S.cerevisiae* RNase H1 binds well to nucleic acids, when produced without the RNase H domain (11). Together, each domain influences the activity of the other in a magnesium-dependent manner. High Mg^{2+} ion concentrations inhibit the dsRHbd while stimulating the RNase H activity of *S.cerevisiae* RNase H1 (14). It is not known how this regulation occurs or how the two regions interact physically, even though the structure of the first dsRHbd of the *S.cerevisiae* RNase H1 has been determined (15), and a rather reasonable model of the RNase H domain is easy to predict (16).

Employing a variety of experimental approaches, we have studied substrate binding and RNase H activity of mouse RNase H1 to determine what novel properties the dsRHbd bestows on the activity of the enzyme. Using a protein that starts at the second AUG codon (Figure 1a), representing most of the enzyme made *in vivo*, and several protein derivatives defective in or devoid of the dsRHbd, we found that the dsRHbd induces dimerization of the protein upon binding to an RNA–DNA hybrid substrate. Dimerization confers processivity to the RNase H activity of the enzyme, a property that indicates the need to process long RNA–DNA hybrids in eukaryotic cells. This work constitutes an important advancement in our understanding of the biological roles of this essential class of enzymes and their possible function in mtDNA biogenesis.

MATERIALS AND METHODS

Protein expression, purification and mutagenesis

E.coli RNase HI was expressed and purified as described previously (17). The cDNAs for mouse and human RNases H1 starting at M27 were cloned in pET-15b (Novagen) (18). cDNAs for proteins starting at M135 (mouse) and M136 (human) were cloned in the pET-15b to produce proteins with only the RNase H domain. The cDNA for *S.pombe* RNase H1 starting at M1 was also cloned in pET-15b, as was the cDNA for a protein starting at C120, which was changed to M to make a protein corresponding to the RNase H domain of *S.pombe*. The QuickChange kit (Stratagene) was used to generate all site-directed mutants using custom-designed primers to change W43A (TGG to GCG), K59A (AAG to GCG) and K60A (AAA to GCA). All protein purifications were similar starting from *E.coli* cells induced to express the corresponding protein by addition of solid lactose to a final concentration of 0.4%. About 20 grams of frozen cells were broken by grinding with alumina in a mortar placed on ice. The extract was suspended in 100 ml of 40 mM sodium phosphate pH 7.0, 300 mM NaCl, 5 mM imidazole, 5% glycerol, 3 mM 2-mercaptoethanol, 20 μg/ml PMSF (buffer A) and treated with DNase I. Cell debris and ribosomes were removed by centrifugation, and proteins were purified using Talon resin column (Clontech, USA), by elution with buffer A containing 200 mM imidazole. Protein-containing fractions were combined, and after adding EDTA to 5 mM, precipitated with an equal volume of saturated $(NH4)_2SO_4$. After centrifugation, the pellet was dissolved in 50 mM sodium phosphate pH 7.0, 50 mM NaCl, 10% glycerol, 3 mM 2-mercaptoethanol, 1 mM EDTA (buffer B) and applied to a 7.5 × 75 mm TSK-GEL HPLC column SP-5PW (TosoHaas, Japan). Elution was performed using a linear gradient from buffer B to buffer B containing 1M NaCl. Tubes with the highest RNase H concentrations were combined and applied to a 10 × 1200 Sephacryl S-200 high resolution column (Amersham Pharmacia Biotech) equilibrated in 50 mM sodium phosphate pH 7.0, 1 M $(NH_4)_2SO_4$, 10% glycerol, 3 mM 2-mercaptoethanol and 1 mM EDTA. For surface plasman resonance (SPR) experiments, the Sephacryl column was run in STB buffer (10 mM Tris–HCl pH 7.5, 50 mM NaCl, 1mM EDTA, 1 mM 2-mercaptoethanol and 0.005% P20), and the protein was used within a few days and if found necessary concentrated on Microcon YM-10 Centrifugal Filter Device (Millipore, Bedford, MA). The latter resulted in a significant loss of protein. The purity of the proteins was >95% as demonstrated by SDS–PAGE analysis with Coomassie blue staining.

Biosensor analysis

Experiments were performed in the BIACORE 1000 or BIACORE 2000 instruments at 25°C. Sensor chips F1 with a carboxymethylated dextran matrix were obtained from BIACORE AB (Piscataway NJ, USA). Running buffers were HBS (10 mM HEPES pH 7.4, 150 mM NaCl, 1 mM EDTA and 0.005% P20) or STB as indicated. To generate a sparsely populated surface with low negative charge that would significantly reduce non-specific binding of RNases H, the chip surface was pre-neutralized. Using HBS buffer, amino coupling reagent (NHS/EDC) was employed to activate the surface followed by deactivation with 1 M ethanolamine pH 8.5. NHS/EDC treatment was repeated. Immobilization of low levels of streptavidin (~200 RU) and capture of biotinylated nucleic acids (Figure 1) at 2 nM were essentially as described previously (19). After nucleic acid capturing on the streptavidin coated chip surface, the surface was washed with 2 M NaCl, mimicking the regeneration step and removing the final traces of non-covalently bound streptavidin.

Proteins in STB buffer were injected at flow rates indicated in the Supplementary Figure S1 legend. After each round of injection bound protein was completely removed by two passages of 10 μl of 2 M NaCl followed by 10 μl of STB at flow rate 20 μl/min. All experiments were performed in duplicate and were repeated on at least two different chips.

Data transformation of the primary sensograms and overlay plots were prepared with BIAevaluation 3.2 software (BIACORE AB, Uppsala, Sweden). The response from the reference surface was subtracted from that of each of the experimental surfaces to correct for refractive index changes, matrix effects, nonspecific binding, injection noise and baseline drift (20).

Molar ratio (υ) (the number of moles of RNase H bound per mole of RNA–DNA hybrid) was estimated using the equation

$$\upsilon = \left(\Delta R_{\text{RNase H}}/\Delta R_{\text{hybrid}} \times 0.8\right) \times \left(\text{MW}_{\text{hybrid}}/\text{MW}_{\text{RNase H}}\right),$$

where ΔR_{hybrid} is the increase in surface concentration expressed in response units (RU) upon binding of the hybrid on streptavidin surface, $\Delta R_{\text{RNase H}}$ is the increase in RU upon binding of the corresponding RNase H to the hybrid, and $\text{MW}_{\text{hybrid}}$ and $\text{MW}_{\text{RNase H}}$ are the molecular weights of the hybrid and RNase H, respectively. An increase of 1000 RUs corresponds to the binding of 1 ng protein/mm^2 and for nucleic acids 1 ng/mm^2 gives 800 RUs (21).

Substrates

All substrates, except poly(rA)/poly(dT), were purchased from Integrated DNA Technologies, Coralville, IA. Labeling of the 18 and 36 mer RNAs used standard protocols with (γ-^{32}P-ATP and polynucleotide kinase. Labeled RNA was purified by PAGE and extracted from gel slices by the crush and soak method, and 100–200:1 of labeled oligonucleotide solution was used for annealing with an ~5-fold molar excess of complementary DNA oligonucleotide, incubated for 10 min at 65°C, then slow cooled to room temperature (15 min). Poly(rA)/poly(dT) was prepared as described previously (22).

Gel mobility shift assays (GMSAs)

GMSAs were performed in a final volume of 8 μl containing 10 nM solution of 5′-^{32}P-labeled RNA–DNA hybrid in buffer 50 mM Tris–HCl pH 7.9, 50 mM NaCl, 10% glycerol, 1 mM DTT, 0.5 mM EDTA, 0.04% bromophenol blue (BB), 0.04% xylene cyanol (XC). Various amounts (Figure 2) of enzyme were added and samples were incubated at 4°C for 10 min, after which 7 μl of each reaction mixture were loaded on 8% TBE (10 mM Tris–borate pH 7.5 and 1 mM EDTA) polyacrylamide gel (0.04 × 30 × 40 cm) that had been pre-run for 1.5–2 h. Electrophoresis was carried out at 4°C at 500 V until BB migrated 10 cm (~1.5 h). After electrophoresis, the gel was dried and exposed to phosphor screens (STORM). Data were analyzed with ImageQuant 5.2.

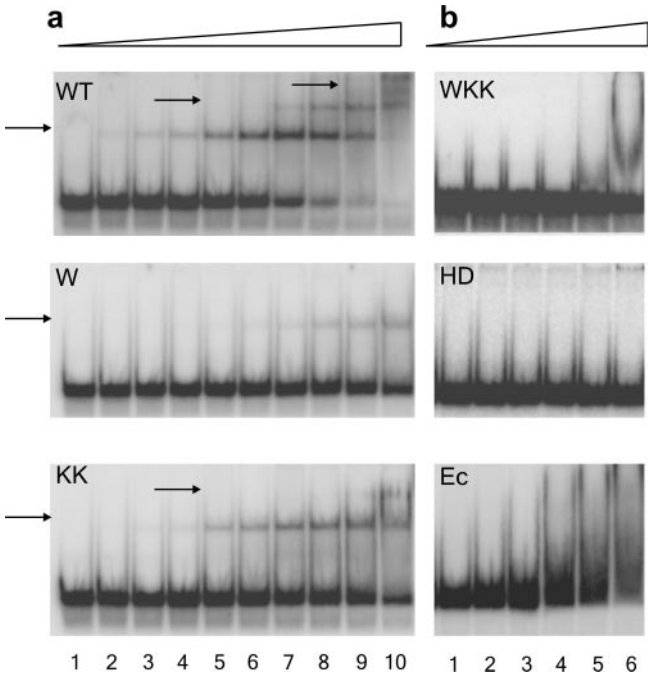


Figure 2. GMSA of 36 bp RNA–DNA hybrid and RNases H. RNA–DNA hybrid was incubated with increasing (left to right) concentrations of RNases H and electrophoresed as described in the Materials and Methods. RNases H1 analyzed were wild-type mouse RNase H1 (WT), RNase $H1^{W}$(W), RNase $H1^{KK}$ (KK), RNase $H1^{WKK}$ (WKK), RNase $H1^{HD}$ (HD) and *E.coli* RNase HI (Ec). Arrows mark the positions of the shifted bands. In (**a**) concentrations were 0, 5, 10, 20, 39, 78, 156, 312, 625 and 1250 nM for lanes 1–10, respectively. In (**b**) for the top RNase H^{WKK} (WKK) the concentrations were: 39, 78, 156, 312, 625 and 1250 nM for lanes 1–6, respectively. For the middle panel RNase H^{HD} (HD) the concentrations were: 60, 120, 240, 470, 940 and 1880 nM for lanes 1–6, respectively. In the bottom *E.coli* RNase HI (Ec), the concentrations were: 47, 94, 187, 375, 750 and 1500 nM for lanes 1–6, respectively.

RNase H activity assays

RNase H activity was determined as described previously (22) by measuring the conversion of 1–10 pmol of ^{32}P-labeled poly(rA)/poly(dT) into an acid-soluble form by incubation with the different enzymes for 30 min at 37°C in 50 mM Tris–HCl pH 7.9, 50 mM NaCl, 1 mM DTT and $MgCl_2$ ranging from 0.5 to 100 mM.

Digestion and analysis of RNA–DNA hybrids

Assays were performed in a final volume of 100 µl containing 10 nM solution of 5′-^{32}P-labeled RNA–DNA in 50 mM Tris–HCl pH 7.9, 50 mM NaCl, 5 mM $MgCl_2$, 1 mM DTT, 20 µg/ml BSA and 4% glycerol. After digestion for the times indicated in Figure 3, 10 µl aliquots were taken and mixed with equal volumes of stop buffer (98% formamide, 20 mM EDTA, 0.04% BB and 0.04% XC) followed by heating to 65°C for 5 min and loading of 7 µl of sample onto 12% TBE polyacrylamide gel (0.04 × 30 × 40 cm) and electrophoresed at 1500 V until BB migrated 25 cm (~90 min). Gels were dried and exposed to phosphor screens and analyzed using ImageQuant 5.2.

Analysis of poly(rA)/poly(dT) digestion

Various enzymes (as described in the legend of Figure 4) were tested for digestion of poly(rA)/poly(dT). Assays were performed in a final volume of 100 µl of reaction mixture containing 400 pmol ^{32}P-labeled poly(rA)/poly(dT) in 50 mM Tris–HCl pH 7.9, 50 mM NaCl, 5 mM $MgCl_2$, 1 mM DTT, 20 µg/ml BSA and 4% glycerol. After digestion for the times indicated in the legend to Figure 4, 10 µl aliquots were taken and mixed with equal volumes of stop buffer followed by heating to 65°C for 5 min and loading of 7 µl onto a 12% TBE polyacrylamide gel (0.04 × 30 × 40 cm) and electrophoresed at 1500 V until BB migrated 25 cm (~90 min). After electrophoresis, the gel was dried and exposed to phosphor screens (STORM). Data were analyzed with ImageQuant 5.2.

Modeling of *E.coli* RNase HI and HIV-1 PPT

The crystal structure of HIV-1 RT complexed with the RNA–DNA double helix (pdb code 1hys) was superimposed with the crystal structure of RNase HI from *E.coli* (pdb code 1rnh). Using the program Genemine (23), 12 alpha carbon atoms were best-fit to orient the structures. T470 to N474, E492, V493 and T497 to Y501 of the RT were superimposed on T40 to N44, E64, V65 and T69 to Y73 of RNase HI, respectively. The root mean square displacement (rmsd) for these atoms was 1.22 Å. When H539 of RT and H124 of RNase H were also included in the orientation, the rmsd increased to 1.67 Å. For comparison, the HIV-1 RT–DNA–RNA complex was also oriented to the structure of uncomplexed HIV-1 RT (pdb code 3hvt). Orientation based on the 12 residues specified above produced an rmsd of 0.53 Å. Upon including H539 in the orientation, the rmsd was 0.95 Å. Figure 5 was generated using PyMOL Molecular Graphics System (http://www.pymol.org).

RESULTS

Complexes between RNases H and RNA duplexes

Ribonucleases H cleave RNA of RNA–DNA hybrids in a sequence non-specific manner (24), presumably by variably positioning the RNA of the hybrid in the catalytic site through interactions with the DNA strand. For *E.coli* RNase HI the basic protrusion provides some of the contacts to DNA (24), while in HIV-1 RT the RNase H primer grip supplies this function (25). Eukaryotic RNases H1 have a region corresponding to the basic protrusion that could serve as the substrate-binding site for hydrolysis. However, the dsRHbd of the eukaryotic enzymes is also involved in nucleic acid binding (11,14). To determine the contribution of the dsRHbd to substrate interactions, we have used SPR to compare the binding of mouse RNase H1 and *E.coli* RNase HI to various RNA–DNA hybrids.

Previously, we had employed SPR to measure binding of *E.coli* RNase HI and HIV-1 RT to 36 bp RNA–DNA and observed one mole of RT (19) and three moles of *E.coli* RNase HI (26) per 36 bp hybrid, consistent with the fact that the *E.coli* enzyme is much smaller, and occupies less space on the substrate. We started this work by using SPR analysis to compare the binding of mouse RNase H1 and *E.coli* RNase HI to the 36 bp RNA–DNA hybrid substrate. We observed about six moles of the mouse protein per mole of the hybrid (Supplementary Figure S1a), twice the value we obtained for *E.coli* RNase HI, an unexpected result because the

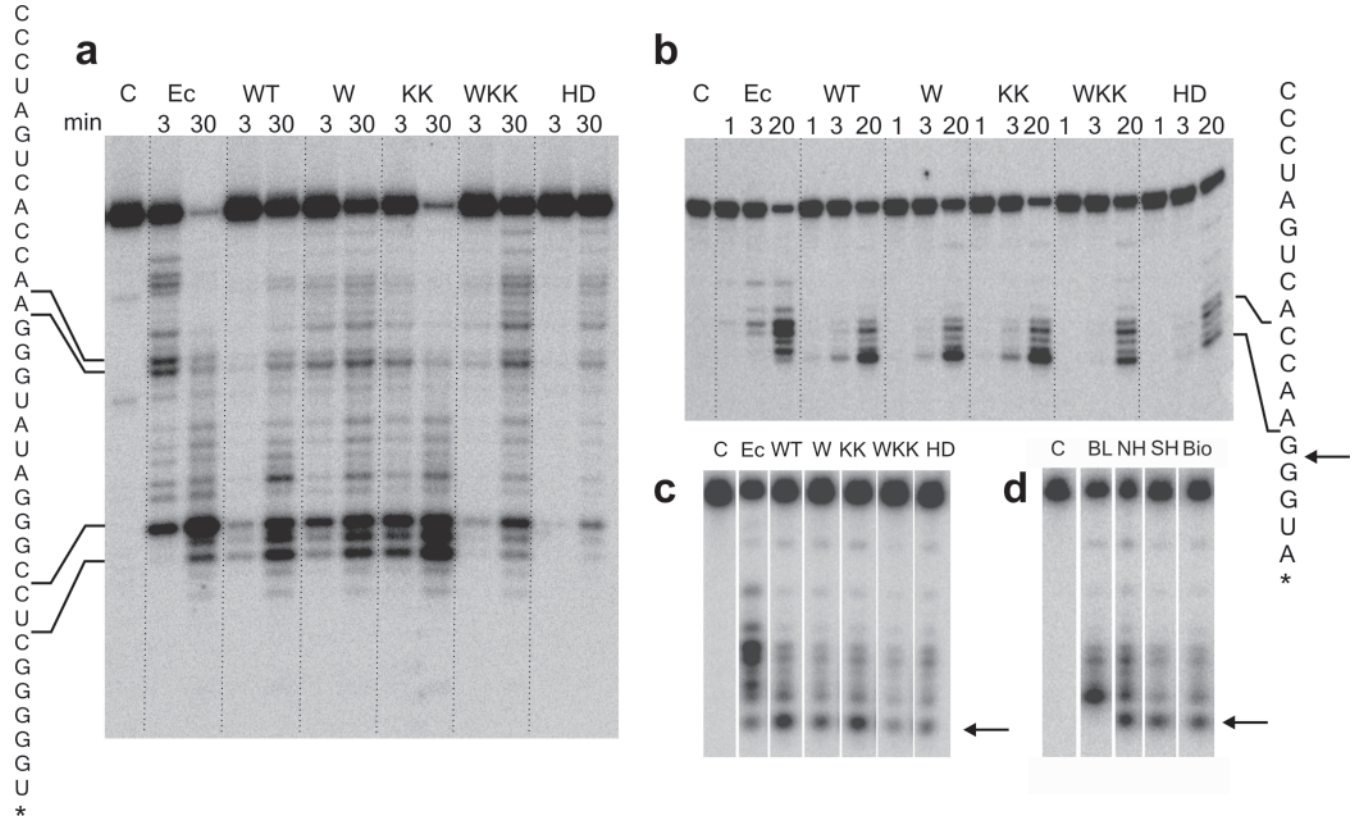


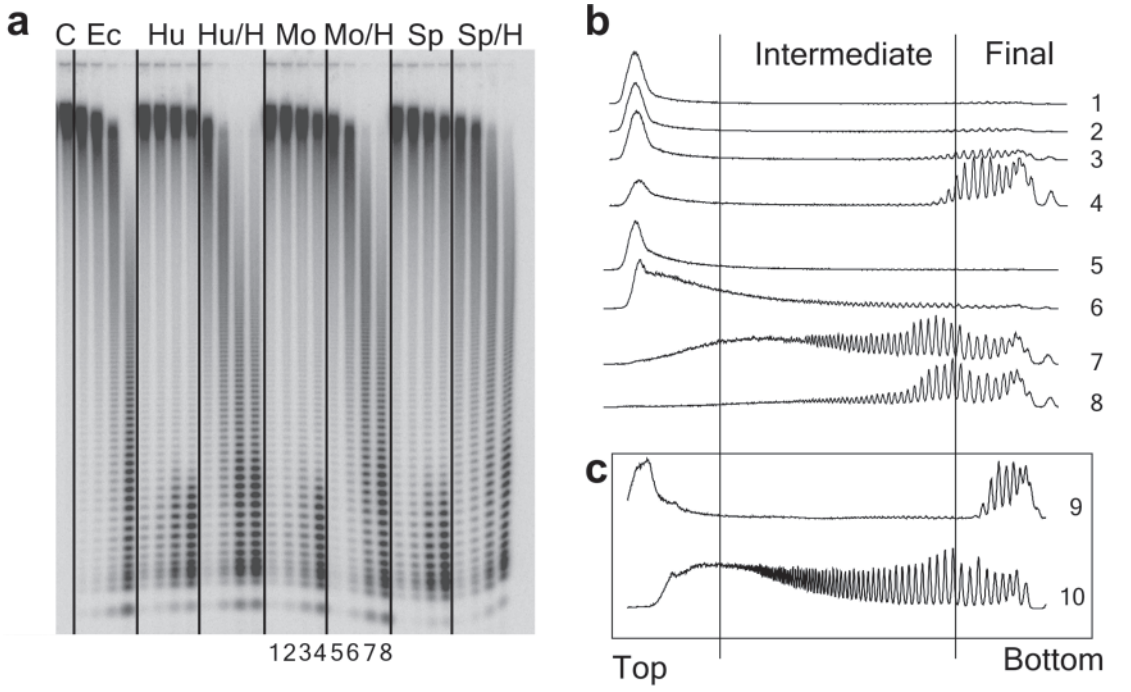
Figure 3. Cleavage patterns of RNases H. RNases H digestion products were analyzed (see Materials and Methods) by PAGE. The sequences of the RNAs are shown with an asterisk marking the label at the 5′-end of the RNA. Lines indicate the positions of the cleavage products on the sequence. Proteins are represented as in Figure 2. C corresponds to the no enzyme control. For the 36 bp RNA–DNA hybrid (**a**) and the 18 bp RNA–DNA hybrid (**b**) the times at which the digestion products were taken is shown above the lanes. In (**c**) the DNA of the 18 bp hybrid had an additional unpaired T residue at its 3′-terminus. (**d**) The DNA of the 18 bp hybrid (BL) was modified at the 3′-terminus with amino (NH), sulfhydryl (SH) and biotinylated (Bio) linkers. The amino linker was C3, the sulfhydryl linker was C6 and the biotin linker was a 15 carbon mixed polarity spacer (3′ Biotin-TEG IDT Inc.). The arrows on the right of (b, c and d) indicate the additional site of cleavage resulting from the modification of the DNA. Based on their relative specific activities, the quantities of enzyme used were adjusted such that approximately the same level of degradation was observed.

mouse protein is larger and would be predicted to cover more, not less of the substrate. The binding of multiple proteins to a nucleic acid is a complicated process to study by SPR. To simplify these analyses, we designed a 12 bp hairpin hybrid that was stable in solution, and could be attached to the chip surface. We checked the interaction of *E.coli* RNase HI with the 12 bp hairpin hybrid, and observed saturation at a ratio of one mole of protein per one mole of hybrid, as expected (Figure 1b). For the mouse protein we found two moles of enzyme bound to the 12 bp hairpin hybrid (Figure 1b), again double that of the *E.coli* enzyme and consistent with the data obtained for the 36 bp RNA–DNA.

To test whether the differences in stoichiometry between the two proteins were due to the extra sequences present in the mouse enzyme, we made a construct consisting only of the RNase H domain (RNase $H1^{HD}$) of mouse RNase H1 and examined its binding to the 12 bp hairpin hybrid. Compared with the full-length protein, the binding curve looked very different. At the protein concentrations used in these assays RNase $H1^{HD}$ could not reach saturation. Even at the highest concentrations, the protein could only approach a one to one molar ratio to the hybrid substrate (Figure 1b), more similar to *E.coli* RNase HI than to the full-length mouse enzyme. These data show a strong influence of the N-terminal region of mouse RNase H1 on the manner in which the full-length protein interacts with hybrid substrates.

Binding of 2 mol of mouse RNase H1 to 12 mer hybrid substrate requires a functional dsRHbd

The dsRHbd is highly conserved among eukaryotic RNases H1 and is involved in duplex RNA-binding (11,14,18). We mutated amino acids important for binding to check their effect in protein-hybrid interactions by SPR. Single mutant W43A (RNase $H1^{W}$), double mutant K59A, K60A (RNase $H1^{KK}$), and triple mutant W43A, K59A, K60A (RNase $H1^{WKK}$) protein derivatives of mouse RNase H1 (Figure la) were impaired in binding in Northwestern assays (data not shown), and in SPR analysis they showed different stoichiometries than the wild-type enzyme (Figure 1c). At 100 nM concentration RNase H1 was already approaching saturation by binding 2 mol of protein per mole of 12 bp hybrid. In contrast RNase $H1^{W}$ and RNase $H1^{KK}$ were binding at <1:1 molar ratio, similar to *E.coli* RNase HI. At the 100 nM concentration RNase $H1^{WKK}$ only reached ~0.5 mol of protein per mole of RNA–DNA hybrid, following a saturation pattern very similar to that of the RNase $H1^{HD}$ (Figure 1b and c), and demonstrating a correlation between the presence of an active



d Ratio of final to intermediate products

Protein	WT	W	KK	WKK	HD	Ec
Ratio	1.27±0.01	0.23±0.08	0.36± 0.02	0.11±0.04	0.09±0.04	0.14±0.03

Figure 4. Degradation of poly(rA)/poly(dT) by RNases H. Polyacrylamide gel separation of the products of digestion of ^{32}P-labeled poly(rA)/poly(dT) is shown in (**a**) with *E.coli* RNase HI (Ec), human RNase H1 (Hu), the RNase H domain of human RNase H1 (Hu/H), mouse RNase H1 (Mo), the RNase H domain of mouse RNase H1 (Mo/H), *S.pombe* RNase H1 (Sp) and the RNase H domain of *S.pombe* RNase H1 (Sp/H). C represents the no enzyme control. Lanes 1–4 are from samples taken after 1, 3, 10 and 30 min of incubation with mouse RNase H1 and lanes 5–8 correspond to the same time points for the RNase H domain of mouse RNase H1. Similar loading was done for the RNases H from other species. Graphical representations of the products in lanes 1–8 are shown in (**b**). In (**c**) graphical data from a separate set of experiments are shown for mouse RNase H1 (lane 9) and *E.coli* RNase HI (lane 10). Products were divided into 'final' and 'intermediate' indicated by the vertical lines in the graphical representation (b and c). Data from at least three independent experiments describing the ratio of 'final' to 'intermediate' products are shown in (**d**) for wild-type and mutant mouse RNases H1 and for *E.coli* RNase HI.

duplex RNA-binding domain and a two-to-one protein to substrate interaction.

Purified mouse RNase H1 is monomeric

SPR results indicated different stoichiometric binding of mouse and *E.coli* RNase H1/I to RNA–DNA hybrids. The *E.coli* enzyme is known to be a monomer in solution (27), but the mouse protein could be a dimer through interactions involving the dsRHbd. We checked the state of association of mouse RNase H1 by analytical ultracentrifugation (AUC) using both sedimentation velocity and sedimentation equilibrium and found that the enzyme was predominantly a monomer in solution, with >95% of the protein in monomeric state (Supplementary Figure S2). Sedimentation equilibrium experiments of mouse RNase H1 in the presence of the 12 bp RNA–DNA hybrid revealed an average stoichiometry of the protein-hybrid complex of 2.03:1 (data not shown) in good agreement with the SPR data. These results are consistent with the formation of a dimeric enzyme induced by binding to RNA–DNA hybrids.

One mole of mouse RNase H1 binds to 12 mer RNA–RNA duplexes

Because the dsRHbd of eukaryotic RNases H1 is capable of binding RNA–RNA duplexes as well as RNA–DNA hybrids (11,14), we examined the interactions of mouse and *E.coli* RNases H1/I with a 12 bp RNA–RNA hairpin by SPR. *E.coli* RNase HI failed to bind to the RNA–RNA duplex demonstrating a strong preference for RNA–DNA. In contrast, mouse RNase H1 bound well, although with a molar ratio of one mole of protein per mole of RNA–RNA, a clear distinction from the binding to RNA–DNA (Figure 1d), and an indication that the protein interacts in a different way with these two similar duplex RNAs. An interesting feature of the binding of mouse RNase H1 to dsRNA is that it is completely inhibited in the presence of 10 mM Mg^{2+} (Figure 1d), in agreement with previous reports describing the interaction of *S.cerevisiae* RNase H1 and dsRNA by northwestern assays (14). Because RNase H1 cleaves an RNA–DNA hybrid in the presence of Mg^{2+} ion, it is not possible to perform binding assays in the presence of Mg^{2+} ion with the full-length protein and a hybrid substrate.

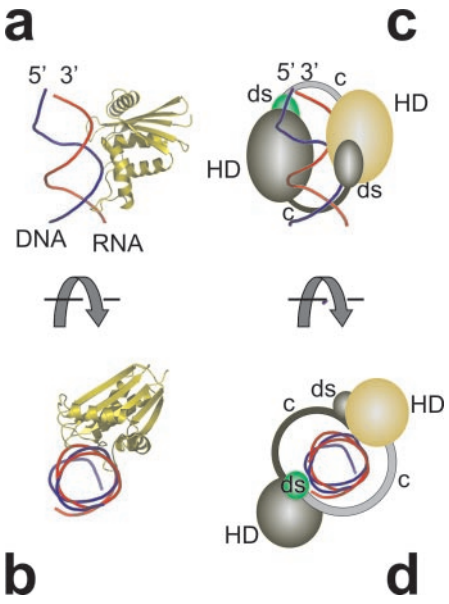


Figure 5. Model of interaction of RNase H1 with RNA–DNA hybrid. Interaction of *E.coli* RNase HI with HIV-1 PPT RNA–DNA hybrid is shown (**a** and **b**). The model was constructed as described in Materials and Methods. The RNA strand is in red and the DNA strand in blue. 5′-DNA and 3′-RNA mark the position of the ends of the hybrid. The right side of the figure (**c** and **d**) is the model for two molecules of mouse RNase H1 interacting with ~12 bp of the PPT. The yellow RNase H domain (HD) is in the same orientation as the *E.coli* RNase HI protein on the left side of the figure and is connected to the green dsRHbd (ds) by the gray connection domain c. The second RNase H1 molecule is in dark green and is in the opposite orientation, making it unable to cleave the hybrid. The lower portion of the figure (b and d) are an end-on views of structures (a and c) with the hybrid in (d) surrounded by the two RNase H1 molecules.

GMSAs indicate dimer formation

Binding of six molecules of mouse RNase H1 to a 36 bp RNA–DNA hybrids could occur in different ways. We used GMSA to identify the nature of these interactions. Six molecules binding independently would produce six shifted bands, while under cooperative interactions, the binding of the first protein would rapidly attract the rest resulting in a single band of a complex containing six moles of protein per mole of hybrid. A third possibility is dimer formation upon substrate binding, which would generate three shifted bands consisting of one, two or three dimers per 36 bp RNA–DNA hybrid. The latter result was observed in the experiments shown in Figure 2a (upper panel). As the protein concentration increased, the wild-type enzyme shifted the hybrid into three bands (arrows in Figure 2a) indicative of three independent dimers.

We also checked mouse RNase H1 mutant proteins altered in their dsRHbd and *E.coli* RNase HI by GMSA (Figure 2). *E.coli* RNase HI and the RNase H1^{HD} of mouse did not produce discrete bands, only a smear in the case of the *E.coli* enzyme, and no change of mobility for the mouse H domain protein (Figure 2b middle and lower panel). The triple mouse mutant protein (RNase H1^{WKK}) bound the substrate at very high protein concentrations, but instead of distinct bands, it produced aggregates (Figure 2b upper panel). The single (RNase H1^{W}) and the double (RNase H1^{KK}) mutant proteins formed bands migrating with the same mobility as the bands observed for the wild-type RNase H1 (Figure 2a), suggesting that these proteins can also form stable dimers, although they required significantly higher protein concentrations, especially RNase H1^{W}. Moreover, while RNase H1 produced three shifted-bands with the 36 bp RNA–DNA hybrid, RNase H1^{KK} formed two, and RNase H1^{W} only one. These differences revealed the importance of W43, K59 and K60 in dimer formation. W43 was the most critical, and interestingly is the most highly conserved amino acid of the dsRHbd.

All of the enzymes were able to cleave the 36 bp RNA–DNA (see the next section) but using GMSA some of the proteins interacted only transiently with the substrate and did not form stable complexes or dissociated during the electrophoresis producing a smear rather than discrete bands. Thus, stable complexes were formed only through the dsRHbd domain that is defective or missing in *E.coli* RNase HI, mouse RNase 1^{WKK} and RNase H1^{HD}.

Table 1. RNase H specific activity

	5 mM	10 mM	50 mM	100 mM
RNase H1	33.3	40	100	28.3
RNase H1^{W}	30	33.3	0.8	0.6
RNase H1^{KK}	30	38.3	8.3	1
RNase H1^{WKK}	10	8	0.11	0.1
RNase H1^{HD}	1.6	1.2	0.03	0.03
E.coli HI	11.8	6.6	1.6	0.8

Specific activities of various enzymes were determined at the indicated Mg^{2+} ion concentrations as described in Materials and Methods. All values are relative to RNase H1 at 50 mM, which was set at 100%.

RNase H activity of *E.coli* and mouse RNase HI/1

We also examined the effect of the dsRHbd on the RNase H catalytic activity of the mouse wild-type and mutated RNases H1. We checked the degradation of poly(rA)/poly(dT) by the different proteins in the presence of various concentrations of Mg^{2+} ion. *E.coli* RNase HI had a Mg^{2+} ion optimum of ~5–10 mM, as reported previously (24). However, the mouse wild-type protein had its highest activity at ~50 mM Mg^{2+} ion (Table 1), an unusually high value, and still retained ~30% of its activity at 100 mM Mg^{2+}. The dsRHbd seemed to confer resistance to the inhibition of the RNase H activity by high Mg^{2+} ion concentration, because the proteins defective in or devoid of the dsRHbd have Mg^{2+} ion dependence similar to the *E.coli* enzyme. RNase H1^{W} and RNase H1^{KK}, both only partially affected in their binding ability by GMSA, had Mg^{2+} ion optima ~10 mM. RNase H1^{WKK} and the RNase H1^{HD} had even lower Mg^{2+} ion optima of ~5 mM.

A fully active dsRHbd also enhanced the RNase H activity of the protein. Mouse wild-type RNase H1 had the highest specific activity of all the proteins tested, especially at its optimum of 50 mM Mg^{2+} ion. Decreased activity appeared to correlate with decreased functionality of the dsRHbd. Single and double mutant proteins had activities similar to wild type at 5–10 mM Mg^{2+}, although much lower at 50 mM Mg^{2+} ion. The triple mutant protein had an activity similar to *E.coli* RNase HI, and RNase H1^{HD} had the lowest enzymatic activity.

Similarity in cleavage products between *E.coli* and mouse RNase HI/1

We checked the influence of the dsRHbd on the degradation pattern of hybrid substrates using several RNA–DNA duplexes ranging from 12 to 36 bp. On all of these substrates both mouse and *E.coli* RNases H cleaved at numerous sites but not at every phosphodiester bond (Figure 3, and data not shown). In general, wild-type and mutant mouse proteins showed a similar degradation pattern, with only the relative frequencies of the hydrolytic products varying slightly. For the 36 bp RNA–DNA hybrid, RNA oligonucleotides of 7, 8 and 9 bases in length were prominent products (Figure 3a). For the 18 bp RNA–DNA substrate, mouse proteins generated a series of oligomers of 5–9 nt (Figure 3b). These same products were also observed when using *E.coli* RNase HI, but unlike the mouse proteins, the shorter oligonucleotides were only produced after significant levels of degradation had occurred (compare the 3 and 30 min time points in Figure 3a and b). It seems that the dsRHbd facilitated cleavage closer to the end of the RNA–DNA hybrid (e.g. compare the relative amounts seven and nine oligomer products of wild-type RNase H1 and RNase $H1^{WKK}$ in Figure 3a). The different mutant mouse proteins showed a distribution of hydrolytic product in between the pattern of wild-type and *E.coli* RNases H related to the strength of the dsRHbd. As with other properties the more altered the dsRHbd the more similar the product formation was to the one shown by *E.coli* RNases HI.

The data reported in Figure 3a and b were obtained with blunt end hybrids as substrates. When an additional unpaired T residue is added to the 3′-end of the DNA, we observed a new RNA cleavage product one nucleotide shorter than without the extra-unpaired T (marked by arrow in Figure 3c). Interestingly, adding any of several different substituents, which could be useful for cross-linking studies, to the 3′-end of the DNA also permitted the formation of the shorter tetrameric RNA oligonucleotide regardless of the enzyme used (Figure 3d). All of the proteins, including *E.coli* RNase HI were able to cleave at this new site (Figure 3c), suggesting that the enzymes recognized these substrates in a very similar manner.

Digestion of poly(rA)/poly(dT) is affected by dsRHbd

Digestion products of short RNA–DNA hybrids were very similar for *E.coli* and mouse RNase HI/1. However, when we examined products of hydrolysis of longer hybrids by these two enzymes, and the mutant protein derivatives, we found a very different pattern of degradation related to the presence of an active dsRHbd. *E.coli* RNase HI cleaved poly(rA)/poly(dT) in a distributive manner producing a continuum of oligonucleotides of various length ranging in size from that of the starting substrate to two nucleotides (Figure 4). In contrast, mouse RNase H1, even at the earliest time points, had very few degradation products of intermediate size, releasing mainly small oligonucleotides indicative of complete digestion of the substrate (Figure 4). Mouse RNase $H1^{HD}$ produced oligonucleotides in a broad size range, similar to the *E.coli* enzyme. A difference in digestion pattern between full-length and the RNase H domain protein derivative was also observed for other eukaryotic RNases H1. Human and *Schizosaccharomyces pombe* enzymes degraded poly(rA)/poly(dT) similarly to the mouse protein and relied on their dsRHbd to produce mainly final products (Figure 4a).

We considered the presence of final products without abundant intermediate product formation as an indication of processive cleavage. To estimate the level of processivity for the different enzymes, we classified products <13 nt in length as 'final' and those falling between 13 and ~180 (marked by the vertical line in Figure 4b and c) as intermediate products. The ratio of final to intermediate products was calculated from several experiments choosing samples with ⩾85% of the starting poly(rA)/poly(dT) still remaining. An ~10-fold difference was found when comparing wild-type mouse and *E.coli* RNases H (Figure 4d). RNase $H1^{HD}$ and RNase $H1^{WKK}$ both gave values even lower than *E.coli* RNase HI. RNase $H1^{W}$ and RNase $H1^{KK}$ had ratios higher than *E.coli* RNase HI but much lower than the wild-type enzyme (Figure 4d), indicating an intermediate level of processivity. From our results, we concluded that eukaryotic RNases H1 were acting processively and that their processivity was dependent on the dsRHbd.

DISCUSSION

Mouse RNase H1 forms dimers on RNA–DNA hybrids

Experiments using two different techniques, SPR and AUC, demonstrated that mouse RNase H1 was a monomer in solution, but formed a two to one complex on a 12 bp hairpin hybrid and about a six to one complex on a 36 bp RNA–DNA. GMSA experiments were consistent with the two-to-one complex resulting from dimer formation on the 36 bp RNA–DNA hybrid rather than sequential binding. We have modeled these interactions by wrapping two molecules of RNase H1 around the RNA–DNA hybrid (Figure 5). To generate this model, we first replaced the RNase H of HIV-1 RT, complexed with the polypurine tract (PPT) (25), by the known structure of the *E.coli* RNase HI (8) (Figure 5a and b). Because the RNase H domain of mouse RNase H1 is similar in sequence and presumably in structure to the *E.coli* and HIV-RT proteins, the model presented in Figure 5a could also be valid for the RNase H domain of the mouse protein. However, a model of a complex of the full-length protein would have to include the additional N-terminal region of eukaryotic RNase H1 and its interactions with the hybrid substrate. There is little room to accommodate more than one RNase H1 molecule on the ~12 bp of the PPT RNA–DNA shown in Figure 5a in a manner that would place the catalytic site of a second enzyme in register for cleavage. One way to fit two protein molecules is by placing them on opposite sides of the 12 bp hybrid (Figure 5c,d), either in the same or opposite orientations. With the proteins in opposite orientation, only one molecule would be able to cleave the substrate because of the asymmetric nature of the RNA–DNA structure.

Dimers of mouse RNase H1 are processive nucleases

Completely encircling the substrate is a common manner for a protein to become processive (28). The model presented in Figure 5c and d is consistent with the main conclusion of the paper: eukaryotic RNases H1 are processive enzymes. Processive enzymes usually have large binding surfaces and participate in electrostatic and hydrophobic interactions to

generate a complex between the protein and the polymer with relative high affinity, but no sequence specificity (28). In the case of eukaryotic RNases H1, interactions with an RNA–DNA hybrid occurred through the dsRHbd and the RNase domain, and from our results it appeared that both regions are important, because missing either binding surface prevented both dimer formation and processive cleavage. RNase H1 bound dsRNA only through the dsRHbd, the RNase H domain did not interact with this polymer (Figure 1d and data not shown), and consequently dimers were not formed on dsRNA. Similarly, proteins deficient in (RNase $H1^{KK}$, RNase $H1^{W}$ and RNase $H1^{WKK}$) or lacking the dsRHbd (RNase $H1^{HD}$) exhibit reduced dimerization and processivity levels, depending on the robustness of the binding of the dsRHbd.

The dsRHbd is a highly conserved amino acid sequence with a central core of hydrophobic residues. The surface residues include a conserved tryptophan (W43) and a few positively charged amino acids, including two highly conserved lysines (K59 and K60) (15). For mouse RNase H1 the hydrophobic contacts provided by W43 were more important than the electrostatic interaction of K59 and K60, because RNase $H1^{W}$ was more defective in dimerization and processivity than RNase $H1^{KK}$ (Figures 2 and 4). The mutant protein lacking electrostatic and hydrophobic interactions (RNase $H1^{WKK}$) had a highly impaired dsRHbd and behaved essentially the same as the protein without the dsRHbd (RNase $H1^{HD}$) and *E.coli* RNase HI; they bound only through the RNase H domain, which supports neither dimerization nor processivity.

Dimerization and processivity do not affect hydrolysis of short hybrids

Processivity was most obvious during the degradation of long hybrids (Figure 4) and had little effect on the hydrolysis of shorter substrates (Figure 3). We observed only small differences in the distribution of degradation products between *E.coli* RNase HI, and mouse RNase H1 and the different mutant protein derivatives, indicating that the RNase H domain largely dictated the site of hydrolysis. This is in marked contrast to the RNase H of RTs. A protein derived from MuLV RT containing only the RNase H domain retains activity but loses its specificity (29), while in HIV-1 RT the RNase H domain relies on other portions of the protein for activity (30,31).

Although in mouse the dsRHbd had little effect on specificity, it affected both the specific activity and Mg^{2+} ion optimum of the enzyme. Wild-type RNase H1 had the highest activity, especially at its Mg^{2+} ion optimum of 50 mM. The more impaired the dsRHbd, the lower the RNase H activity, and the higher the inhibition by Mg^{2+} ion. Of all mutant proteins, RNase $H1^{KK}$ was the least defective in the dsRHbd and had the highest activity at 50 mM Mg^{2+} ion, indicating that the dsRHbd was responsible for the unusually high Mg^{2+} ion optimum of mouse RNase H1. The physiological significance of a 50 mM Mg^{2+} ion optimum is not clear, but previously we had suggested a model in which the activities of the two domains of the protein were inversely related in a magnesium-dependent manner (14).

Several of the properties of mouse RNase H1 reported here are quite different from those reported for the human enzyme (13,32–34), including Mg^{2+} ion optimum and the effect of the dsRHbd on cleavage site preference. The protein we used resembles that found *in vivo* and was produced in a soluble form in *E.coli*. The human RNase H1 protein used by the ISIS Pharmaceuticals group includes the MLS, representing the precursor of the mitochondrial isoform of RNase H1 and requires solublization after expression in *E.coli*. Therefore, we attribute discrepancies to the different modes of protein purification, and to the fact that we have used different protein isoforms, not to the biochemical differences between the human and mouse proteins.

In vivo role of RNase H1

A requirement for a processive RNase H in eukaryotes may reflect the presence of RNA–DNA hybrids that are not normally found in bacteria. *E.coli* RNase HI acts in a distributive manner and can readily degrade the short RNA–DNA hybrids resulting from primer formation for DNA replication (35). Transcription of DNA normally does not produce RNA–DNA hybrids, but in *E.coli* uncoupling of transcription and translation leads to the generation of substantial amounts of hybrid that are presumably longer than the primers of DNA synthesis, and under some conditions they accumulate to levels that RNase HI cannot overcome (36). In *S.cerevisiae*, RNA–DNA hybrids form when RNA-binding proteins fail to interact with nascent transcripts and can result in the stimulation of recombination (37). In mouse long RNA–DNA hybrids are found at the sites transcribed in immunoglobulin class switch regions and such long RNA–DNA hybrids may create a requirement for the processivity of RNase H1 (38,39). In mitochondria, there are also long transcripts with the potential to form RNA–DNA hybrids (40). The essential role of mouse RNase H1 in mtDNA biogenesis during development (6), and the newly uncovered processive nature of the enzyme suggest a possible need to process long hybrids during mtDNA replication.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at NAR Online.

ACKNOWLEDGEMENTS

We thank Dr Sergei A. Kazakov for aiding in the design of hairpin hybrids, Dr Sergey G. Lokhov and the Friday Seminar for helpful discussions, and Sarah Shelby for construction of several mutants. Dr Kiyoshi Mizuuchi generously provided use of the BIACORE 2000. Funding to pay the Open Access publication charges for this article was provided by National Institute of Child Health and Human Development, National Institutes of Health, Department of Health and Human Services.

Conflict of interest statement. None declared.

REFERENCES

1. Ohtani,N., Haruki,M., Morikawa,M., Crouch,R.J., Itaya,M. and Kanaya,S. (1999) Identification of the genes encoding Mn2+-dependent RNase I-III and Mg2+-dependent RNase HIII from *Bacillus subtilis*: Classification of RNases H into three families. *Biochemistry*, **38**, 605–618.

2. Itaya,M. and Crouch,R.J. (1991) A combination of RNase H (rnh) and recBCD or sbcB mutations in *Escherichia coli* K12 adversely affects growth. *Mol. Gen. Genet*., **227**, 424–432.
3. Arudchandran,A., Cerritelli,S.M., Narimatsu,S.K., Itaya,M., Shin,D.Y., Shimada,Y. and Crouch,R.J. (2000) The absence of ribonuclease H1 or H2 alters the sensitivity of *Saccharomyces cerevisiae* to hydroxyurea, caffeine and ethyl methanesulphonate: implications for roles of RNases H in DNA replication and repair. *Genes Cells*, **5**, 789–802.
4. Ray,D.S. and Hines,J.C. (1995) Disruption of the Crithidia fasciculata RNH1 gene results in the loss of 2 active forms of ribonuclease H. *Nucleic Acids Res*., **23**, 2526–2530.
5. Filippov,V., Filippova,M. and Gill,S.S. (2001) Drosophila RNase H1 is essential for development but not for proliferation. *Mol. Gen. Genomics*, **265**, 771–777.
6. Cerritelli,S.M., Frolova,E.G., Feng,C.G., Grinberg,A., Love,P.E. and Crouch,R.J. (2003) Failure to produce mitochondrial DNA results in embryonic lethality in Rnaseh1 null mice. *Mol. Cell*, **11**, 807–815.
7. Tisdale,M., Schulze,T., Larder,B.A. and Moelling,K. (1991) Mutations within the RNase H domain of human immunodeficiency virus type 1 reverse transcriptase abolish virus infectivity. *J. Gen. Virol*., **72**, 59–66.
8. Yang,W., Hendrickson,W.A., Crouch,R.J. and Satow,Y. (1990) Structure of ribonuclease H phased at 2 A resolution by MAD analysis of the selenomethionyl protein. *Science*, **249**, 1398–1405.
9. Katayanagi,K., Miyagawa,M., Matsushima,M., Ishikawa,M., Kanaya,S., Ikehara,M., Matsuzaki,T. and Morikawa,K. (1990) 3-dimensional structure of ribonuclease H from *Escherichia coli*. *Nature*, **347**, 306–309.
10. Davies,J.F.,II, Hostomska,Z., Hostomsky,Z., Jordan,S.R. and Matthews,D.A. (1991) Crystal structure of the ribonuclease H domain of HIV-1 reverse transcriptase. *Science*, **252**, 88–95.
11. Cerritelli,S.M., Fedoroff,O.Y., Reid,B.R. and Crouch,R.J. (1998) A common 40 amino acid motif in eukaryotic RNases H1 and caulimovirus ORF VI proteins binds to duplex RNAs. *Nucleic Acids Res*., **26**, 1834–1840.
12. Wiedemann,N., Frazier,A.E. and Pfanner,N. (2004) The protein import machinery of mitochondria. *J. Biol. Chem*., **279**, 14473–14476.
13. Wu,H.J., Lima,W.F. and Crooke,S.T. (2001) Investigating the structure of human RNase H1 by site-directed mutagenesis. *J. Biol. Chem*., **276**, 23547–23553.
14. Cerritelli,S.M. and Crouch,R.J. (1995) The non-RNase H domain of *Saccharomyces Cerevisiae* RNase H1 binds double-stranded RNA—magnesium modulates the switch between double-stranded RNA-binding and RNase H activity. *RNA*, **1**, 246–259.
15. Evans,S.P. and Bycroft,M. (1999) NMR structure of the N-terminal domain of *Saccharomyces cerevisiae* RNase HI reveals a fold with a strong resemblance to the N-terminal domain of ribosomal protein L9. *J. Mol. Biol*., **291**, 661–669.
16. Crouch,R.J. and Cerritelli,S.M. (1998) RNases H of *S.cerevisiae*, *S.pombe*, *C.fasciculata*, and *N.crassa*. In Crouch,R.J. and Toulmé,J.J. (eds), *Ribonucleases, H*.. INSERM, Paris, pp. 79–100.
17. Ma,W.P., Hamilton,S.E., Stowell,J.G., Byrn,S.R. and Davisson,V.J. (1994) Sequence specific cleavage of messenger RNA by a modified ribonuclease H. *Bioorg. Med. Chem*., **2**, 169–179.
18. Cerritelli,S.M. and Crouch,R.J. (1998) Cloning, expression, and mapping of ribonucleases H of human and mouse related to bacterial RNase HI. *Genomics*, **53**, 300–307.
19. Gorshkova,I.I., Rausch,J.W., Le Grice,S.F.J. and Crouch,R.J. (2001) HIV-1 reverse transcriptase interaction with model RNA–DNA duplexes. *Anal. Biochem*., **291**, 198–206.
20. Morton,T.A. and Myszka,D.G. (1998) Kinetic analysis of macromolecular interactions using surface plasmon resonance biosensors. In Ackers,G.K. and Johnson,M.L. (eds), *Methods in Enzymology, Energetics of Biological Marcomolecules, Pt B*. Academic Press, Inc., NYVol. 295, pp. 268–294.
21. Fisher,R.J., Fivash,M., Casasfinet,J., Erickson,J.W., Kondoh,A., Bladen,S.V., Fisher,C., Watson,D.K. and Papas,T. (1994) Real-time DNA-binding measurements of the ETS1 recombinant oncoproteins reveal significant kinetic differences between the p42 and p52 isoforms. *Protein Sci*., **3**, 257–266.
22. Dirksen,M.L. and Crouch,R.J. (1981) Selective inhibition of RNase H by dextran. *J. Biol. Chem*., **256**, 11569–11573.
23. Lee,C. and Irizarry,K. (2001) The GeneMine system for genome/proteome annotation and collaborative data mining. *IBM Systems J*., **40**, 592–603.
24. Kanaya,S. (1998) Enzymatic activity and protein stability of *E.coli* ribonuclease HI. In Crouch,R.J. and Toulmé,J.J. (eds), *Ribonucleases H*. INSERM, Paris, pp. 1–38.
25. Sarafianos,S.G., Das,K., Tantillo,C., Clark,A.D., Ding,J., Whitcomb,J.M., Boyer,P.L., Hughes,S.H. and Arnold,E. (2001) Crystal structure of HIV-1 reverse transcriptase in complex with a polypurine tract RNA:DNA. *EMBO J*., **20**, 1449–1461.
26. Haruki,M., Noguchi,E., Kanaya,S. and Crouch,R.J. (1997) Kinetic and stoichiometric analysis for the binding of *Escherichia coli* ribonuclease HI to RNA–DNA hybrids using surface plasmon resonances. *J. Biol. Chem*., **272**, 22015–22022.
27. Kanaya,S., Kohara,A., Miyagawa,M., Matsuzaki,T., Morikawa,K. and Ikehara,M. (1989) Overproduction and preliminary crystallographic study of ribonuclease H from *Escherichia coli*. *J. Biol. Chem*., **264**, 11546–11549.
28. Breyer,W.A. and Matthews,B.W. (2001) A structural basis for processivity. *Protein Sci*., **10**, 1699–1711.
29. Zhan,X.Y. and Crouch,R.J. (1997) The isolated RNase H domain of murine leukemia virus reverse transcriptase—retention of activity with concomitant loss of specificity. *J. Biol. Chem*., **272**, 22023–22029.
30. Smith,J.S., Gritsman,K. and Roth,M.J. (1994) Contributions of DNA polymerase subdomains to the RNase H activity of human immunodeficiency virus type 1 reverse transcriptase. *J. Virol*., **68**, 5721–5729.
31. Zuniga,R., Sengupta,S., Snyder,C., Leon,O. and Roth,M.J. (2004) Expression of the C-terminus of HIV-1 reverse transcriptase p66 and p51 subunits as a single polypeptide with RNase H activity. *Protein Eng. Des. Sel*., **17**, 581–587.
32. Lima,W.F., Nichols,J.G., Wu,H.J., Prakash,T.P., Migawa,M.T., Wyrzykiewicz,T.K., Bhat,B. and Crooke,S.T. (2004) Structural requirements at the catalytic site of the heteroduplex substrate for human RNase H1 catalysis. *J. Biol. Chem*., **279**, 36317–36326.
33. Lima,W.F., Wu,H.J., Nichols,J.G., Prakash,T.P., Ravikumar,V. and Crooke,S.T. (2003) Human RNase H1 uses one tryptophan and two lysines to position the enzyme at the 3′-DNA/5′-RNA terminus of the heteroduplex substrate. *J. Biol. Chem*., **278**, 49860–49867.
34. Wu,H.J., Lima,W.F. and Crooke,S.T. (1999) Properties of cloned and expressed human RNase H1. *J. Biol. Chem*., **274**, 28270–28278.
35. Ogawa,T. and Okazaki,T. (1984) Function of RNase H in DNA replication revealed by RNase H defective mutants of *Escherichia coli*. *Mol. Gen. Genet*., **193**, 231–237.
36. Broccoli,S., Rallu,F., Sanscartier,P., Cerritelli,S.M., Crouch,R.J. and Drolet,M. (2004) Effects of RNA polymerase modifications on transcription-induced negative supercoiling and associated R-loop formation. *Molec. Micro*., **52**, 1769–1779.
37. Huertas,P. and Aguilera,A. (2003) Cotranscriptionally formed DNA–RNA hybrids mediate transcription elongation impairment and transcription-associated recombination. *Mol. Cell*, **12**, 711–721.
38. Yu,K.F., Chedin,F., Hsieh,C.L., Wilson,T.E. and Lieber,M.R. (2003) R-loops at immunoglobulin class switch regions in the chromosomes of stimulated B cells. *Nature Immunol*., **4**, 442–451.
39. Daniels,G.A. and Lieber,M.R. (1995) RNA–DNA complex formation upon transcription of immunoglobulin switch regions: Implications for the mechanism and regulation of class switch recombination. *Nucleic Acids Res*., **23**, 5006–5011.
40. Shadel,G.S. (2004) Coupling the mitochondrial transcription machinery to human disease. *Trends Genet*., **20**, 513–519.