

Structural Biochemistry of a Type 2 RNase H: RNA Primer Recognition and Removal During DNA Replication

Brian R. Chapados¹, Qing Chai², David J. Hosfield¹, Junzhuan Qiu²
Binghui Shen^{2*} and John A. Tainer^{1*}

¹The Scripps Research Institute
Department of Molecular
Biology, Skaggs Institute for
Chemical Biology MB4, 10550
North Torrey Pines Rd, La Jolla
CA 92037, USA

²Department of Cell and Tumor
Biology, City of Hope National
Medical Center, 1500 East
Duarte Rd, Duarte
CA 91010, USA

DNA replication and cellular survival requires efficient removal of RNA primers during lagging strand DNA synthesis. In eukaryotes, RNA primer removal is initiated by type 2 RNase H, which specifically cleaves the RNA portion of an RNA-DNA/DNA hybrid duplex. This conserved type 2 RNase H family of replicative enzymes shares little sequence similarity with the well-characterized prokaryotic type 1 RNase H enzymes, yet both possess similar enzymatic properties. Crystal structures and structure-based mutational analysis of RNase HIII from *Archaeoglobus fulgidus*, both with and without a bound metal ion, identify the active site for type 2 RNase H enzymes that provides the general nuclease activity necessary for catalysis. The two-domain architecture of type 2 RNase H creates a positively charged binding groove and links the unique C-terminal helix-loop-helix cap domain to the active site catalytic domain. This architectural arrangement apparently couples directional A-form duplex binding, by a hydrogen-bonding Arg-Lys phosphate ruler motif, to substrate-discrimination, by a tyrosine finger motif, thereby providing substrate-specific catalytic activity. Combined kinetic and mutational analyses of structurally implicated substrate binding residues validate this binding mode. These structural and mutational results together suggest a molecular mechanism for type 2 RNase H enzymes for the specific recognition and cleavage of RNA in the RNA-DNA junction within hybrid duplexes, which reconciles the broad substrate binding affinity with the catalytic specificity observed in biochemical assays. In combination with a recent independent structural analysis, these results furthermore identify testable molecular hypotheses for the activity and function of the type 2 RNase H family of enzymes, including structural complementarity, substrate-mediated conformational changes and coordination with subsequent FEN-1 activity.

© 2001 Academic Press

Keywords: RNase HIII structure; type 2 RNase H specificity motifs; RNA primer removal; DNA replication; RNA-DNA hybrid

*Corresponding authors

Introduction

DNA replication in cellular organisms occurs simultaneously on both strands. However, all known DNA polymerases catalyze nucleic acid polymerization exclusively in the 5' to 3' direction. Since DNA polymerases are incapable of initiating replication *de novo*, the replisomes of cellular

organisms contain an RNA polymerase enzyme subunit (primase) that synthesizes short RNA fragments on the lagging strand, which serve as the initial primers for DNA polymerase. As a consequence, lagging strand DNA synthesis results in the production of 100–200 base-pair Okazaki fragments, preceded by a 7–12 nucleotide RNA primer. Persistence of RNA primers during replication is lethal (Qiu *et al.*, 1999a), requiring that the efficient removal of these primers be a critical step during DNA replication. This primer removal activity requires sequence-independent, specific recognition

B.R.C and Q.C. contributed equally to this work.
E-mail addresses of the corresponding authors:
bshen@coh.org; jat@scripps.edu

and hydrolysis of RNA present in an RNA-DNA hybrid duplex. In prokaryotes this function is performed by RNase H plus the exonuclease activity of DNA polymerase I. In eukaryotic systems, DNA polymerases lack an intrinsic nuclease activity, and the primer removal reaction is performed by the coupled activities of an RNase H enzyme plus the 5' → 3' exo/endonuclease activity of 5' flap endonuclease-1 (FEN-1) (Qiu *et al.*, 1999b).

RNase H enzymes are ribonucleotide-specific endonucleases that cleave the RNA portion of RNA-DNA/DNA or RNA/DNA duplexes (Stein & Hausen, 1969). These enzymes, which require divalent metal cations, yield products with 5'-phosphate and 3'-hydroxyl termini (Crouch & Dirksen, 1982; Cathala *et al.*, 1979). RNase H enzymes have been identified in all three kingdoms of life (Cathala *et al.*, 1979; Frank *et al.*, 1998a,b,c, 1999; Ohtani *et al.*, 1999a), and recent phylogenetic sequence analysis has identified two major families of RNase H enzymes: type 1 and type 2. Bacterial RNase HI, yeast RNase H1, mammalian RNase H2 and reverse transcriptase RNase H domains belong to the type 1 family, whereas bacterial RNase HII and HIII, archaeal RNase HII, yeast RNase H2(35) and mammalian RNases H1 are type 2 enzymes (Ohtani *et al.*, 1999b). Several members of the type 1 RNase H family have been well characterized and crystal structures of these enzymes, most notably of *Escherichia coli* RNase HI, reveal the presence of a characteristic fold consisting of a five-stranded central β -sheet flanked on both sides by α -helices (Katayanagi *et al.*, 1990, 1992; Yang *et al.*, 1990a). These structural studies, aided by biochemical and mutational results, have identified the enzyme active site, metal ion requirements, residues essential for substrate binding (Katayanagi *et al.*, 1990, 1992; Kanaya *et al.*, 1990, 1991a; Kashiwagi *et al.*, 1996), and have led to models detailing nucleic acid binding and catalysis (Eder *et al.*, 1993; Crooke, 1995; Kanaya *et al.*, 1994; Haruki *et al.*, 1997; Lima & Crooke, 1997; Keck *et al.*, 1998; Krynetskaia *et al.*, 1999; Wu *et al.*, 1999).

Compared to the type 1 RNase H enzymes, much less is known about the type 2 family of enzymes, but their importance is underscored by both biochemical and genetic studies. Homologs from *E. coli*, *Bacillus subtilis* and *Pyrococcus kodakaraensis* possess ribonucleotide-specific catalytic activity in the presence of RNA/DNA hybrid substrates, necessary for a role in primer removal (Itaya, 1990; Haruki *et al.*, 1998; Ohtani *et al.*, 1999a, 2000). Notably, the type 2 RNase H enzyme from *P. kodakaraensis* specifically cleaves ribonucleotides in a closed, covalent DNA-RNA-DNA oligonucleotide hybridized to complementary DNA (Haruki *et al.*, 1998), which provides an accurate model of an RNA primer flanked by Okazaki fragments. These results demonstrate that type 2 RNase H enzymes possess the catalytic specificity necessary to initiate RNA primer removal. Furthermore, experiments in eukaryotes reveal that type 2 RNase H enzymes are active during replication,

and together with FEN-1, are required for the efficient removal of RNA primers during lagging strand DNA synthesis (Waga & Stillman, 1994; Waga *et al.*, 1994; Turchi *et al.*, 1994; Huang *et al.*, 1994; Murante *et al.*, 1998; Qiu *et al.*, 1999b). Double deletion of the two genes encoding type 2 RNase H homologs in *B. subtilis* leads to cellular lethality, indicating that these enzymes are essential for cellular growth (Itaya *et al.*, 1999). In yeast, cell lines deficient for either FEN-1 or type 2 RNase H exhibit growth defects and combined mutation of these genes leads to severe growth defects or lethality, both of which are rescued by the introduction of a plasmid encoding the yeast type 2 RNase H homolog (Qiu *et al.*, 1999b). In addition, differential sensitivity of *Saccharomyces cerevisiae* mutants lacking type 1, type 2 or both RNase H enzymes to hydroxyurea, ethyl methanesulfonate and caffeine, suggests that type 2 RNase H enzymes recognize a unique class of RNA/DNA hybrid substrates, an activity which is important for maintaining cell cycle integrity (Arudchandran *et al.*, 2000).

Although there is a high percentage of sequence similarity within each family of RNases H (Mian, 1997), sequence comparisons of type 1 and type 2 family members from *E. coli* reveals only ~17% sequence similarity. This low degree of sequence similarity suggests these enzymes have differential functions within the cell, and experiments have shown that type 1 RNase H enzyme may actually function in transcription (Busen *et al.*, 1977; Turchi *et al.*, 1994). The low sequence similarity between the two families of RNase H enzymes, coupled to evident differences in function, suggest that structural comparison of these two enzyme families may provide key insights into the evolutionary functional relationship between the two different RNase H families. Two paramount puzzles for the type 2 RNase H enzymes concern the possible identity of family specific motifs, and the structural chemistry that allows both general A-form duplex nucleic acid binding and specificity for cleaving the RNA portion of an RNA-DNA/DNA hybrid duplex. Here, we address these challenges through integrated structural, mutational, biochemical, and kinetic results for the type 2 RNase H homolog from *Archaeoglobus fulgidus* that delineate the active site and binding motifs, and also provide direct evidence for the involvement of the novel C-terminal domain in substrate binding. Despite the existence of multiple RNase H orthologs in prokaryotic eubacteria, eukaryotic yeast and mammals, only one RNase H, a type 2 RNase H, exists in archaea for which complete genome sequences are known (M. Itaya and S. Kanaya, personal communication). Thus, RNase HII is likely the major enzyme involved in RNA primer removal in these organisms, suggesting that our results are generally representative of type 2 RNase enzymes. The results of type 2 RNase H presented here implicate a specific molecular mechanism for RNA-DNA substrate recognition and catalysis and, furthermore suggest how positioning of the substrate may

promote enzyme-induced, substrate-dependent conformational changes and catalysis.

Results and Discussion

aRNase H activity

To initiate functional analysis of the type 2 RNase H family, we identified, cloned, over-expressed, and purified RNase HII from *A. fulgidus* (aRNase HII). The purified, recombinant enzyme (Figure 1a) likely represents the minimum functional machinery common to the type 2 RNase H family, including the major mammalian RNase H. To assess the endonuclease activity of recombinant aRNase HII, we used ^{32}P -labeled model Okazaki fragments with varying lengths of ribonucleotides. aRNase HII preferentially cleaves an 18-mer and 37-mer RNA-DNA/DNA duplex at the 5' end of the last ribonucleotide (Figure 1(b) and (c)). When a longer substrate (37-mer), with 21 ribonucleotides is used, aRNase HII additionally cleaves the substrate with much lower efficiency at positions 8 and 11 ribonucleotides from the 5' end (Figure 1(b), bottom, and (c)). As the typical RNA primer length *in vivo* varies from 7-12 ribonucleotides, the primary reaction product is consistent with a role in RNA primer removal. The additional weak clea-

vage sites observed in this assay may result from general ribonucleotide-specific activity in the presence of an RNA/DNA substrate. In agreement with its classification as a structure-specific endonuclease, aRNase HII shows no activity towards a single-stranded RNA-DNA chimeric molecule. These data are consistent with observed eukaryotic type 2 RNase H activity (Turchi *et al.*, 1994; Qiu *et al.*, 1999b). Studies reveal an RNase HII from *P. kodakaraensis*, which shows 42% sequence identity with aRNase HII, cleaves at the 5' end of the last ribonucleotide of the RNA-DNA junction in a DNA-RNA-DNA/DNA substrate. This substrate differs from those used in our experiments, in that the RNA region is flanked both upstream and downstream by DNA. However, the fact that both enzymes specifically recognize a comparable region of the substrates, and produce similar products, implies that key structural features necessary for activity are conserved among the eukaryotic and archaeobacterial enzymes.

aRNase HII and the RNase H nuclease family

We determined the 1.95 Å resolution crystal structure of aRNase HII by multi-wavelength anomalous dispersion using a selenomethionine-substituted protein (see Materials and Methods).

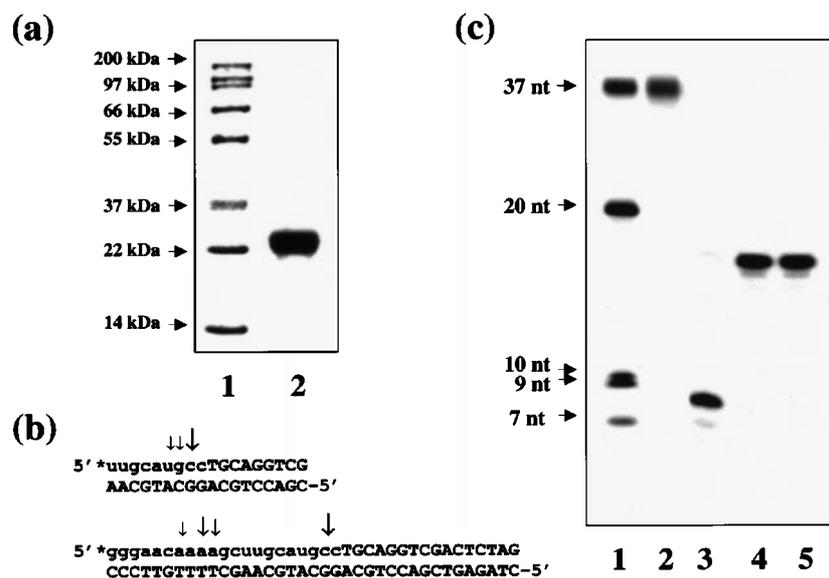


Figure 1. Cleavage specificity of purified recombinant aRNase HII. (a) SDS-12% PAGE of purified, recombinant aRNase HII protein (0.2 mg; lane 2) and molecular mass markers (lane 1) stained with Coomassie brilliant blue. (b) The substrates used for aRNase HII activity. Cleavage sites by aRNase HII on the 18-mer RNA-DNA/DNA (top) and 37-mer RNA-DNA/DNA hybrid (bottom) substrates are denoted with arrows. The sizes of the arrows reflect the relative cleavage product intensities. Deoxyribonucleotides are denoted with uppercase letters, while ribonucleotides are denoted with lowercase letters. The asterisk indicates the radioactively labeled site. (c) Denaturing polyacrylamide gel analysis of aRNase HII hydrolysis products using 37-mer (lane 1) and 18-mer RNA-DNA/DNA (lane 3) substrates. Reactions were carried out at 30°C for ten minutes, and the hydrolysates were separated on a 15% polyacrylamide gel containing 7 M urea. Size markers are shown as numbers of nucleotides (nt) on the left. Structure-specific endonuclease activity was verified by incubating aRNase HII with the 18-mer single-stranded RNA-DNA chimera (lane 5). Stability of 37-mer (lane 2) and 18-mer (lane 4) hybrid duplexes was assayed using untreated RNA-DNA/DNA chimeric duplexes.

The structure of aRNase HII is refined to a crystallographic R value of 0.244 ($R_{\text{free}} = 0.269$) for all data to 1.95 Å resolution (residues 1-200; Table 1), with no residues possessing unfavorable backbone geometry. The aRNase HII structure exhibits a two-domain α/β fold, with overall dimensions of 60 Å × 40 Å × 30 Å (Figure 2(a)). The N-terminal domain (residues 1-159) is comprised of a five-stranded mixed β -sheet, which is flanked by $\alpha 1$, $\alpha 2$ and $\alpha 6$ on one side, and $\alpha 3$, $\alpha 5$ and $\alpha 6$ on the other. The C-terminal region of aRNase HII (residues 160-200) forms a small domain consisting of two short α -helices connected by an ordered loop, which caps the end of the β -sheet. The interface between the two domains is defined by a conserved hydrophobic core centered around the loop between $\beta 1$ and $\beta 2$, which is stabilized by hydrogen bond and van der Waals interactions between main-chain and side-chain atoms of conserved residues in $\alpha 3$, $\alpha 7$ and $\alpha 8$. This boundary, in conjunction with the central β -sheet and surrounding helices, defines a distinct 13 Å deep groove across the center of the enzyme that we propose to be important for substrate binding (see below).

Despite low sequence similarity (17%), the central β -sheet of aRNase HII is superimposable with the β -sheet of *E. coli* RNase HI such that the C^α atoms are within a root-mean-square (rms) deviation of 2 Å (for 54 residues), as determined from a structural alignment with the program SEQUOIA (Figure 3(b)) (Bruns, 1999). Additional structural comparisons reveal the presence of a topologically

identical, mixed, five-stranded β -sheet, surrounded by at least two α -helices in HIV integrase (Dyda *et al.*, 1994), the Holliday junction-specific endonuclease RuvC (Ariyoshi *et al.*, 1994), and the recently determined Tn5 transposase structure (Yang & Steitz, 1995; Davies *et al.*, 2000) (Figure 3(a)). This folding motif, commonly referred to as the RNase H-fold, due to its initial characterization in *E. coli* RNase HI (Yang & Steitz, 1995; Rice *et al.*, 1996), is characteristic of enzymes possessing nuclease and polynucleotidyl transferase activities. The conserved architecture and active site motifs evidently serve as a stable framework onto which variable loops and structural elements are added to modulate substrate binding specificity.

The C-terminal cap domain, which is unique to the type 2 RNase H family members (Figures 2(a) and 3(b), (c)), likely confers substrate-specific binding capacity to the core nuclease architecture, thus accounting for the endonucleolytic specificity of this class of enzymes. As expected, the central nuclease domain is highly conserved between the aRNase HII and mRNase HII structures, while alignment of the cap domain and the lower lip of the central groove show greater structural deviation (Figure 3(c)). Interestingly, structural superposition of aRNase HII and the recently determined type 2 RNase H from *M. jannaschii* (mRNase HII) (Lai *et al.*, 2000) reveals a shift of 1.3 Å in the C^α atoms of the cap domain relative to the position of the central β -sheets (Figure 3(c)). This difference demonstrates the potential for the

Table 1. Crystallographic data collection and refinement statistics

	aRNase HII ^a	aRNase HII-Co(III) ^a	Se-aRNase HII ^{b,c}		
A. X-ray diffraction data					
Wavelength (Å)	0.98	0.863	0.97956	0.97970	0.96111
Unit cell (Å)					
<i>a, b</i>	74.84	74.56	75.74	75.74	75.74
<i>c</i>	142.65	139.68	107.31	107.31	107.31
Resolution (Å)	1.95	2.15	2.70	2.70	2.70
Observations	129,952	82,747	220,010	221,301	223,741
Unique reflections	18,024	13,210	9034	9089	9158
I/σ	18.7	14.3	36.2	44.5	41.5
Completeness (%)	96.7	98.2	99.0	99.0	98.9
Final shell	95.9	93.3	100.0	100.0	100.0
R_{sym}^d	0.094	0.079	0.053	0.046	0.047
Final shell	0.413	0.414	0.299	0.394	0.456
B. Model refinement					
Resolution (Å)	20.0 - 1.95	20.0 - 2.15	20.0 - 2.70		
R_{cryst}^e	0.244	0.223	0.262		
R_{free}^f	0.269	0.272	0.303		
Residues	1-200	1-200	1-195		
Protein atoms	1582	1582	1515		
Solvent atoms	145	119	8		
rmsd. bonds (Å)	0.005	0.005	0.009		
rmsd. angles (deg.)	1.12	1.17	1.31		

^a Space group = $P6_322$.

^b Space group = $P4_32_12$.

^c Initial figure of merit = 0.55; refined, density modified = 0.82.

^d $R_{\text{sym}} = \sum_{hkl} |I(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} I(hkl)$.

^e $R_{\text{cryst}} = \sum_{hkl} |F_{\text{obs}}(hkl) - (F_{\text{calc}}(hkl))| / \sum_{hkl} (F_{\text{obs}}(hkl))$.

^f $R_{\text{free}} = R_{\text{cryst}}$ calculated using the 5% of reflections against which the model was not refined.

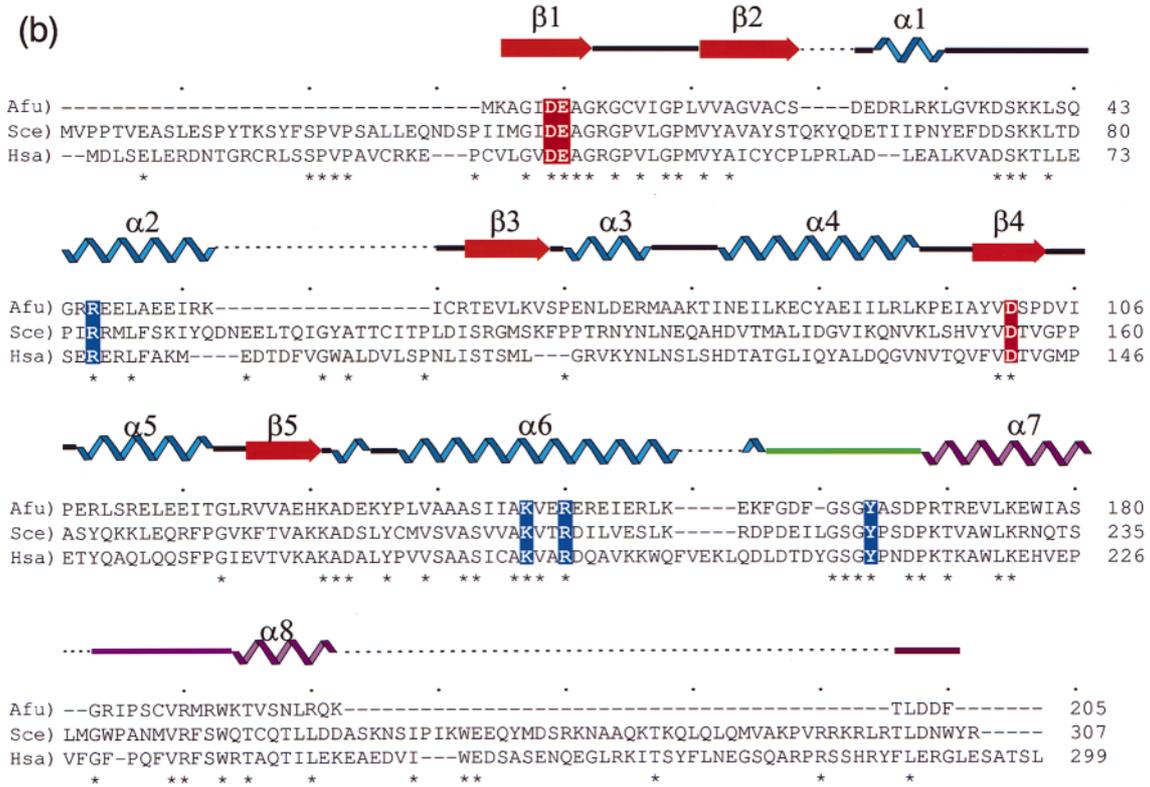
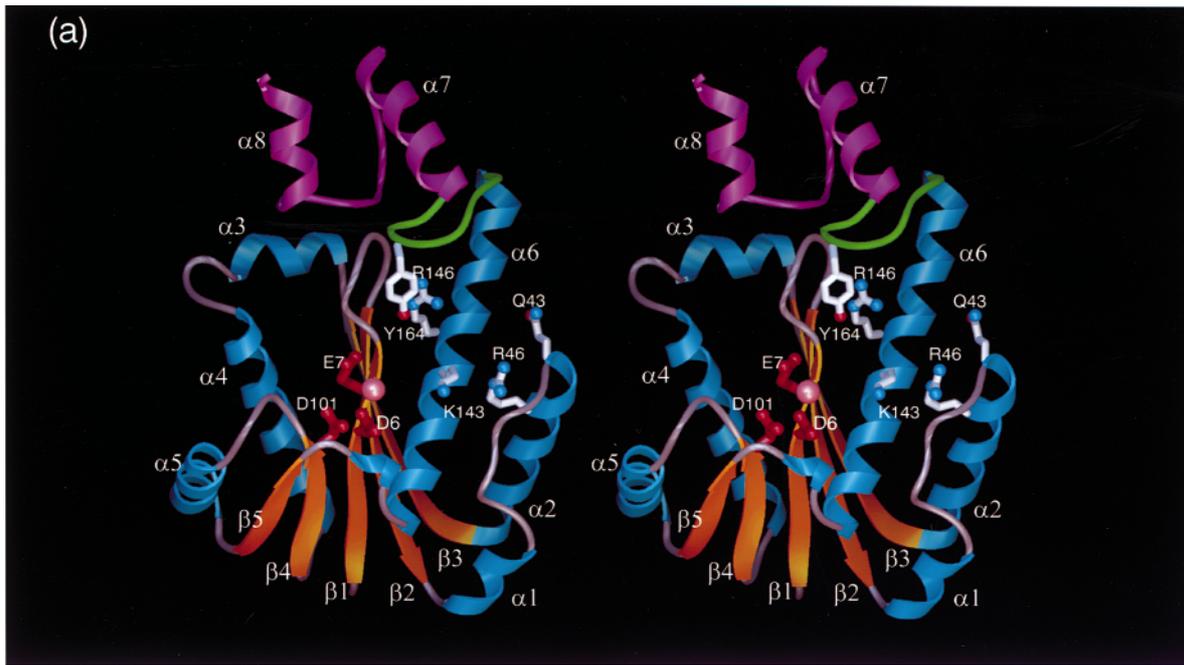


Figure 2. aRNase HIII structure and structure-based type 2 RNase H secondary structure, sequence alignment, and residue function. (a) Stereo view of aRNase HIII overall fold and topology with α -helices (blue coils) and β -strands (gold) labeled sequentially. The single metal ion (pink sphere), the C-terminal “cap” domain (magenta), and the tyrosine finger (green) along with active site (red) and putative substrate binding (white) residues are shown looking straight into the binding groove. (b) Amino acid sequence alignment of aRNase HIII (Afu) with homologs from *S. cerevisiae* (Sce) and *Homo sapiens* (Hsa), indicating the proposed active site residues (red) and residues that affect substrate binding (blue). Secondary structure is indicated above the sequence, denoting β -sheets (arrows), α -helices (ribbons) and highlighting the cap domain (magenta) and the tyrosine finger (green).

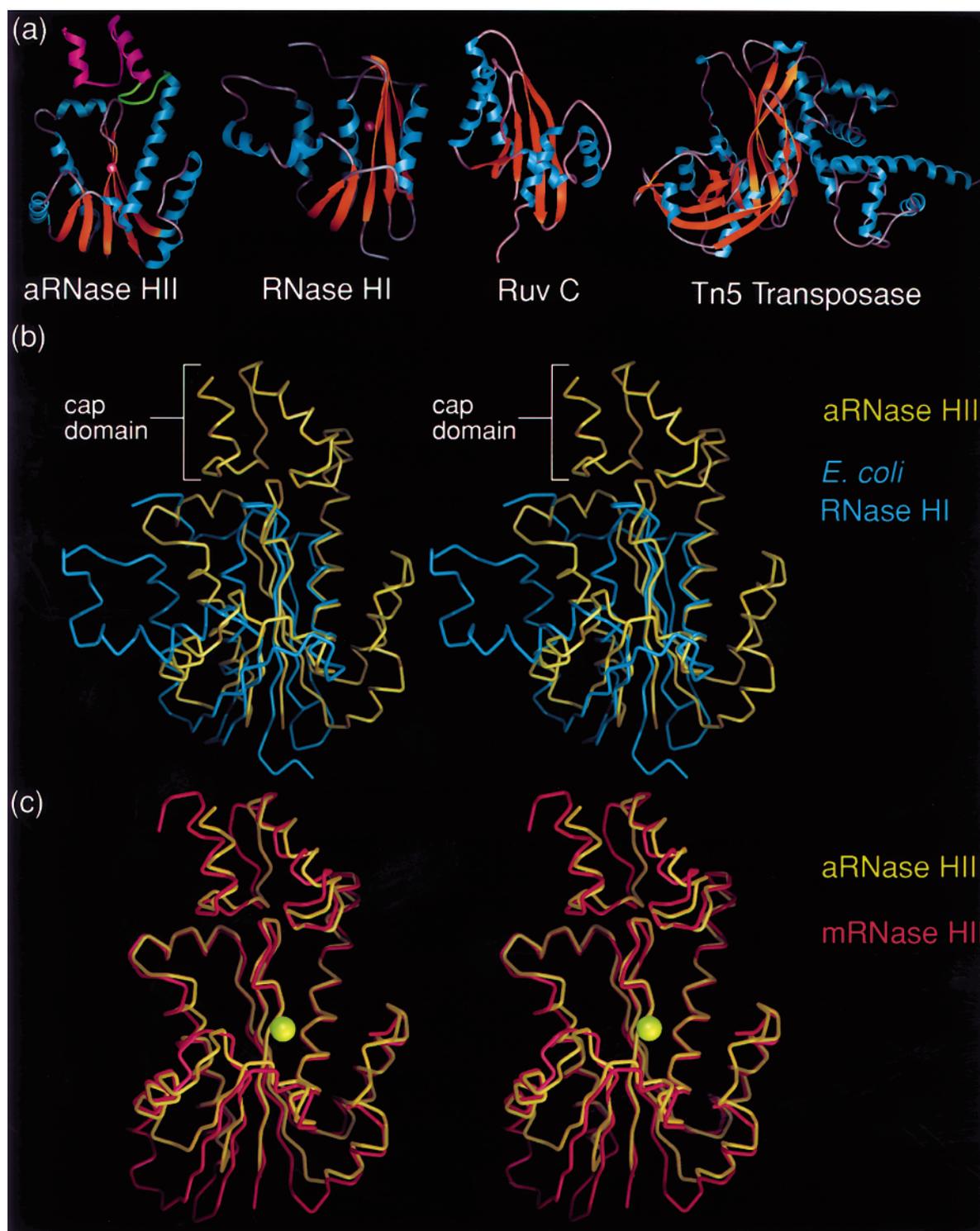


Figure 3. The RNase H family fold. (a) Structural alignments of the β -sheet fold for aRNase HII, *E. coli* RNase HI (PDB: 1RDD), *E. coli* RuvC (PDB: 1HJR) and Tn5 transposase DNA complex (PDB: 1F3I), representing the fold common to members of the RNase H nuclease family. (b) Structural superposition of C α atoms from aRNase HII (gold) and *E. coli* RNase HI (blue) calculated with the program SEQUOIA. The unique C-terminal cap domain, present only in type 2 RNase H enzymes, confers substrate specific binding capacity to the conserved nuclease core domain. (c) Structural superposition of C α atoms from aRNase HII (gold) and mRNase HII (red) calculated with the program SEQUOIA. The central core nuclease domains are nearly identical, while the cap domain, the helical lip of the central groove, and the outside edge of the β -sheet show greater structural deviations.

movement of a conserved Gly-Ser-Gly-Tyr motif (Figure 2(a), green), located in the $\alpha 6$ - $\alpha 7$ loop of the cap domain, towards the conserved active site motif, which we propose to act in substrate discrimination.

Active site and implicated metal coordination

The active site of aRNase HIII is identified by the presence of a conserved, acidic catalytic triad located near the floor of the groove created by $\beta 1$, $\beta 4$ and $\alpha 6$. This active site motif is common among enzymes exhibiting the RNase H-fold, particularly the type 1 RNase H family, in which these residues are clustered in a small pocket near the middle of the central strand in the β -sheet and evidently serve as ligands to a single divalent metal ion (Katayanagi *et al.*, 1990, 1993; Yang *et al.*, 1990b). Structural alignment of *E. coli* RNase HI and aRNase HIII (Figure 3(b)) results in a surprisingly near overlap of two metal ligands, Asp6 and Asp101, whose C α atom positions are within 1 Å of conserved aspartate residues in *E. coli* RNase HI. Interestingly, the C α position of Glu7 is 5 Å away from the cognate residue in *E. coli* RNase HI. Despite this discrepancy, the similarity of the active site environments between type 1 and type 2 RNase H enzymes, combined with comparable catalytic functions suggest an analogous catalytic mechanism. However, the mechanism by which these enzymes utilize metals in catalysis is unclear,

and both one and two metal ion mechanisms have been proposed to explain the metal dependency of RNase H enzymes. The one-metal ion mechanism, as proposed for *E. coli* RNase HI, is thought to involve a general base carboxylate-hydroxyl relay mechanism, in which coordination of a Mg $^{2+}$ by one carboxyl group mediates interactions with substrate nucleic acid, while the other two conserved acidic residues position and activate a water molecule for attack at the scissile phosphate (Kayana & Ikehara, 1995). Alternatively, the two-metal ion mechanism is based on that proposed for the Klenow fragment 3' \rightarrow 5' exonuclease domain, in which one metal ion interacts with the substrate, stabilizing a reaction intermediate, while the other is involved in positioning and activating a water molecule for attack at the scissile bond (Beese & Steitz, 1991; Steitz & Steitz, 1993).

To experimentally examine the divalent metal ion coordination of type 2 RNase H enzymes, we soaked the apoenzyme in a solution containing 10 mM cobalt(III) hexammine chloride. An electron density map calculated from Co(III) containing crystals shows difference density for the Co(III) ion (Figure 4). As Co(III) ligands are non-exchangeable with solvent, the metal ion in this structure is coordinated through hydrogen bonding of amine groups to carbonyl oxygen atoms of Asp6, Asp101 and the backbone carbonyl oxygen atom of Asp6. Preliminary biochemical evidence indicates that cobalt(III) hexammine chloride inhibits aRNase HIII

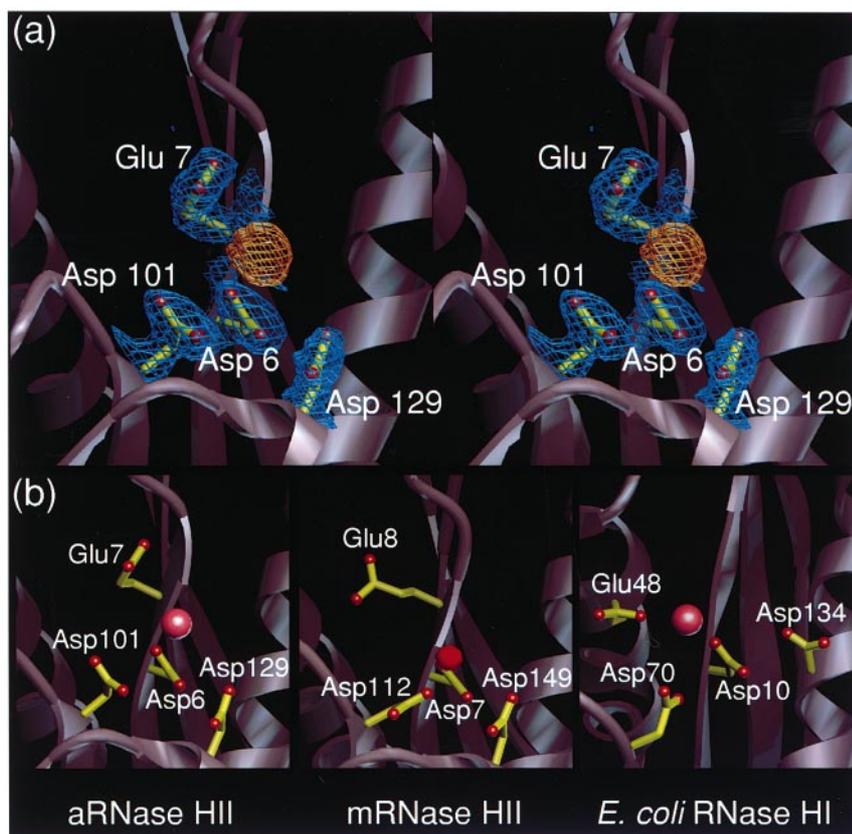


Figure 4. aRNase HIII active site defined by the presence of Co(III) inhibitor. (a) aRNase HIII active site in the presence of cobalt(III) hexammine chloride, coordinated by Asp6, Asp101 and Glu7, and Asp129 (yellow). Electron density map is shown for the acidic triad Asp6, Asp101 and Glu7, and Asp129 ($2F_{\text{obs}} - F_{\text{calc}}$ contoured at 1.2σ , blue) as well as the Co $^{3+}$ ($F_{\text{metal}} - F_{\text{nat}}$ difference density contoured at 8σ , gold). (b) Comparison of the proposed active site residues in aRNase HIII, mRNase HIII and *E. coli* RNase HI enzymes. All three structures are shown in the same orientation (as in (a)) after structural alignment with the program SEQUOIA. Metal ions are shown as pink spheres (aRNase HIII and *E. coli* RNase HI; PDB code: 1RDD). The red icosahedron represents a water molecule located near the active site in the mRNase HIII structure (PDB code: 1EKE), which was determined without a bound metal ion.

(unpublished results), possibly resulting from the absence of exchangeable water ligands capable of activation for hydrolytic cleavage of the RNA, and implicating the divalent metal in type 2 RNase H enzymes in activating the catalytic water molecule. Alternatively, cobalt(III) hexamine chloride could sterically block access to active site residues, preventing direct ligation of a divalent cation. While the presence of a single cobalt(III) hexamine chloride in the active site agrees with previous structural observations, it does not show that catalysis necessarily proceeds through a single-metal ion mechanism, as the presence of a nucleic acid substrate may provide the phosphate ligands necessary for a second metal site (Keck *et al.*, 1998). However, these aRNase HII experimental results are consistent with the structurally identified acidic triad, and suggest further mutational and biochemical experiments necessary to elucidate the nature of metal ion-dependent activity of this class of enzymes.

Structural comparison of the independently determined aRNase HII and mRNase HII type 2 enzymes supports the metal-binding role for two herein identified, conserved aspartate residues (Lai *et al.*, 2000). The third metal ligand in mRNase HII, however, was proposed to be a conserved aspartate corresponding to Asp129 in aRNase HII (Lai *et al.*, 2000) (Figure 4(b)). Conservative mutation, and subsequent activity analysis of structurally identified, catalytically important residues Asp6, Asp101 and Glu7 indicate that endonucleolytic activity is severely inhibited by mutation of these residues, whereas mutation of Asp129 shows an

approximate 50% reduction in enzymatic activity (Figure 5). The absence of catalytic activity in the D6N, D101N and E7Q mutants agrees with the results of identical mutations of corresponding residues in *E. coli* RNase HI (Asp10, Asp70, Glu48; Figure 4(b)). The moderate decrease in catalytic activity observed in the D129N mutant is also consistent with conservative mutation of the cognate residue in *E. coli* RNase HI (Asp134; Figure 4(b)). Based on extensive mutational analysis of this residue in *E. coli* RNase HI, it was proposed that the presence of a polar atom at this position, capable of participating in hydrogen-bonding interactions is essential for catalysis (Haruki *et al.*, 1994; Kashiwagi *et al.*, 1996). Furthermore, based on these data, this residue was proposed to activate a water molecule for attack at the scissile bond, while the acidic triad was primarily involved in metal coordination (Kashiwagi *et al.*, 1996), which we suggest acts to appropriately position the 5' phosphate for catalytic attack by the activated water molecule. Thus, this aRNase HII structure, coupled with these mutational and biochemical results, supports the catalytic importance of the acidic triad Asp6, Asp101 and Glu7, and identifies potential metal ligands that are distinct from those proposed for mRNase HII (Figure 4(b)). The existence of two experimentally testable definitions of the conserved catalytic triad, derived from independently determined structures, coupled with preliminary mutational analysis of Asp129, suggests possible molecular mechanisms; however, future biochemical and mutational experiments

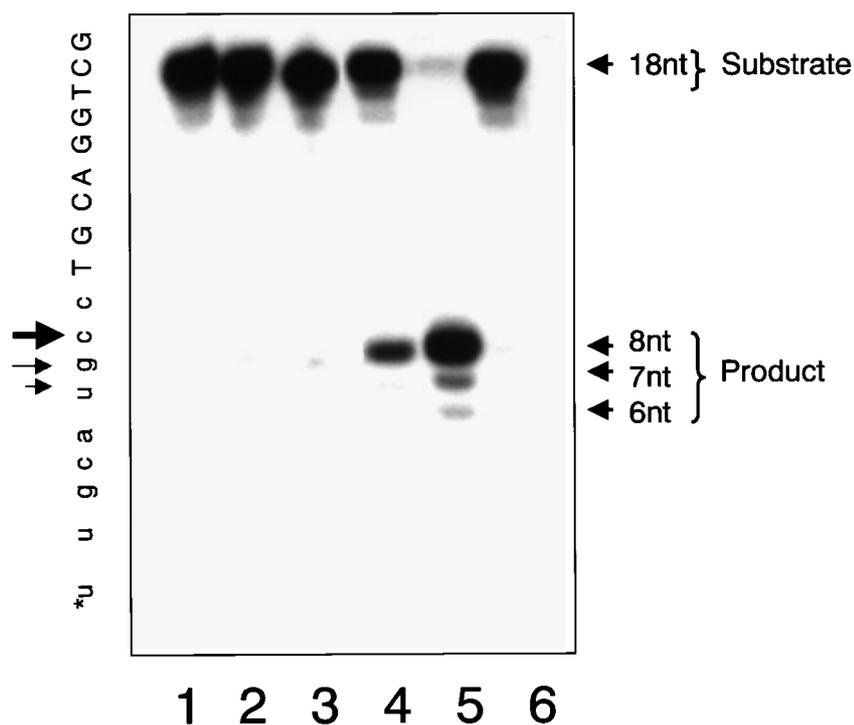


Figure 5. Catalytic activity of mutant aRNase HII with 18-mer RNA-DNA/DNA duplex substrate. Enzymatic activity of each protein was determined at 30°C for ten minutes in 10 mM Tris-HCl (pH 8.0) buffer containing 50 mM NaCl, 1.5 mM bovine serum albumin, 5 mM β -mercaptoethanol, with 25 nM protein and the 18-mer RNA-DNA/DNA duplex substrate. For the control, the labeled substrate was added to a reaction omitting the enzyme. A base-hydrolyzed RNA ladder was prepared as described (see Materials and Methods), and the RNA sequence is shown on the left. Arrows indicate the site of cleavage. Key: 1, aRNase H mutant D6N; 2, mutant E7Q; 3, mutant D101N; 4, mutant D129N; 5, wild-type aRNase H; and 6, control.

will be important for elucidating the mechanism of the type 2 RNase H enzymes.

Substrate binding groove

In the type 2 RNase H structure, the central five-stranded β -sheet is flanked by α -helices on both sides, and is capped on top by the C-terminal subdomain, defining a distinct groove (29 Å long \times 18 Å wide \times 13 Å deep) across the center of one face of the enzyme. Overall, the groove is slightly bent, resembling an open claw, with the conserved Gly-Ser-Gly-Tyr motif (Figure 2(b)) on the $\alpha 6$ - $\alpha 7$ loop jutting inward, poking a "tyrosine-finger" towards the active site (Figure 2(a)). The inner walls of the groove are positively charged for interaction with either the negatively charged phosphate backbone, or the 2'-OH present in the minor groove of an RNA-DNA/DNA hybrid duplex (Figure 6(a)). The active site lies approximately 10 Å below and just inside of the bottom lip of the groove formed by $\beta 4$, the $\beta 4$ - $\alpha 5$ loop, the bottom of $\alpha 6$, and the $\alpha 5$ - $\alpha 6$ loop (Figure 2(a)). Outside the central catalytic cavity, $\alpha 6$, $\beta 2$, $\beta 3$, and $\alpha 4$ make up the walls of the groove, while the top rim is anchored to the central β -sheet by hydrophobic interactions between loop regions at the bottom of the cap domain, and the $\beta 1$ - $\beta 2$ loop (Figure 2).

The prominent active site binding groove houses the majority of conserved basic and polar residues within the type 2 RNase H family (Figure 6(b)). These residues are asymmetrically distributed within the groove, with the majority of conserved residues concentrated on one face of the binding groove. As the arrangement of α -helices and loops lining the groove is different on each side of the active site, the cluster of conserved residues provides a directionality to the groove that is likely important for substrate alignment. Thus, the RNA/DNA hybrid duplex substrate likely binds across the enzyme surface, with the helical axis perpendicular to the central β -sheet, similar to the duplex DNA binding mode observed for uracil-DNA glycosylase (Slupphaug *et al.*, 1996; Parikh *et al.*, 1998, 2000).

These type 2 RNase H structures support proposals that RNase H enzymes discriminate RNA/DNA duplexes from normal DNA by specifically recognizing double-stranded nucleic acids that adopt non-standard structures, as RNA-DNA/DNA hybrid duplexes adopt a conformation that is intermediate between *A*-form and *B*-form DNA (Fedoroff *et al.*, 1993; Gyi *et al.*, 1996, 1998; Conn *et al.*, 1999). Three positively charged or polar residues, Lys143, Arg46, and Gln43 and the N-terminal helical dipole of $\alpha 2$ (Figure 2(a)) penetrate into the substrate-binding groove from $\alpha 6$ and the flanking $\alpha 1$ - $\alpha 2$ loop region, and are spaced an average of 6 Å apart from each other. The precise spacing of these residues evidently defines a "phosphate ruler" that could confer selective binding specificity *via* charge complementarity, due to the shorter inter-phosphate distances in *A*-form

nucleic acid, relative to *B*-form (5.9 Å, compared to 7.0 Å) (Saenger, 1984). Of these three residues, both Lys143 and Arg46 are absolutely conserved within all type 2 RNase H enzymes, and the farthest residue from the active site, Gln43, is adjacent to the positive dipole of $\alpha 2$ and is substituted in several homologs by a polar or positively charged residue (Figure 2(b)). Additional conserved residues including Arg146, located above Lys143, likely aid in the positioning of an RNA/DNA duplex by interacting with negatively charged phosphate groups. In addition to this phosphate recognition role, Arg146 may stabilize the $\beta 1$ - $\beta 2$ loop and thereby couple substrate binding to movement of the cap domain. We propose that cap domain motions are furthermore likely to be coupled to substrate binding through the positioning of a tyrosine finger motif containing the sequence Gly-Ser-Gly-Tyr. The key residue of this motif, Tyr164, is ~ 7 Å above the active site and is evidently poised to interact with substrate nucleic acid bases through hydrophobic stacking interactions. Thus, the existence of this multivalent, structure-specific binding surface, which could allow specific recognition of an *A*-form nucleic acid conformation *via* the Tyr finger and Arg-Lys phosphate ruler motifs, is likely important for substrate targeting in the primer removal pathway of higher eukaryotes.

The structure of mRNase HIII contained a single molecule of 2-(*N*-morpholino) ethanesulfonic acid (Mes) (Lai *et al.*, 2000) bound in the upper left corner of the central groove, as oriented in Figure 3(c), suggesting a different substrate binding face from that proposed here for aRNase HIII. However, structural and biochemical results for type 2 RNase H enzymes suggest it is unlikely that specific, substrate discriminating protein-nucleic acid contacts occur in the region where Mes binds. Instead, the positively charged residues located near the Mes binding site, and outside of $\alpha 9$ and $\alpha 4$, likely interact weakly with the phosphate backbone of a bound substrate, providing electrostatic complementarity with the duplex nucleic acid.

To experimentally test these proposed binding modes, we assayed the substrate binding and catalytic properties of specific aRNase HIII mutants. Individual alanine mutations of residues Arg46, Lys143, Arg146, and Tyr164, result in significantly decreased affinities, with small changes in k_{cat} (Table 2). The most dramatic effects were observed with the R46A mutant, which displayed a 60-fold increase in K_{m} . Coupled with the over 30-fold decrease in substrate binding affinity observed for the K143A mutant, these data support the role of these residues in generic *A*-form substrate recognition, as key constituents of the proposed phosphate ruler motif. A significant decrease in binding affinity was also observed with the Y164A mutant, which displayed a 44-fold increase in K_{m} . This dramatic effect of the Tyr164 mutant suggests to us that this residue may intercalate into the duplex to stabilize a bent conformation that may be import-

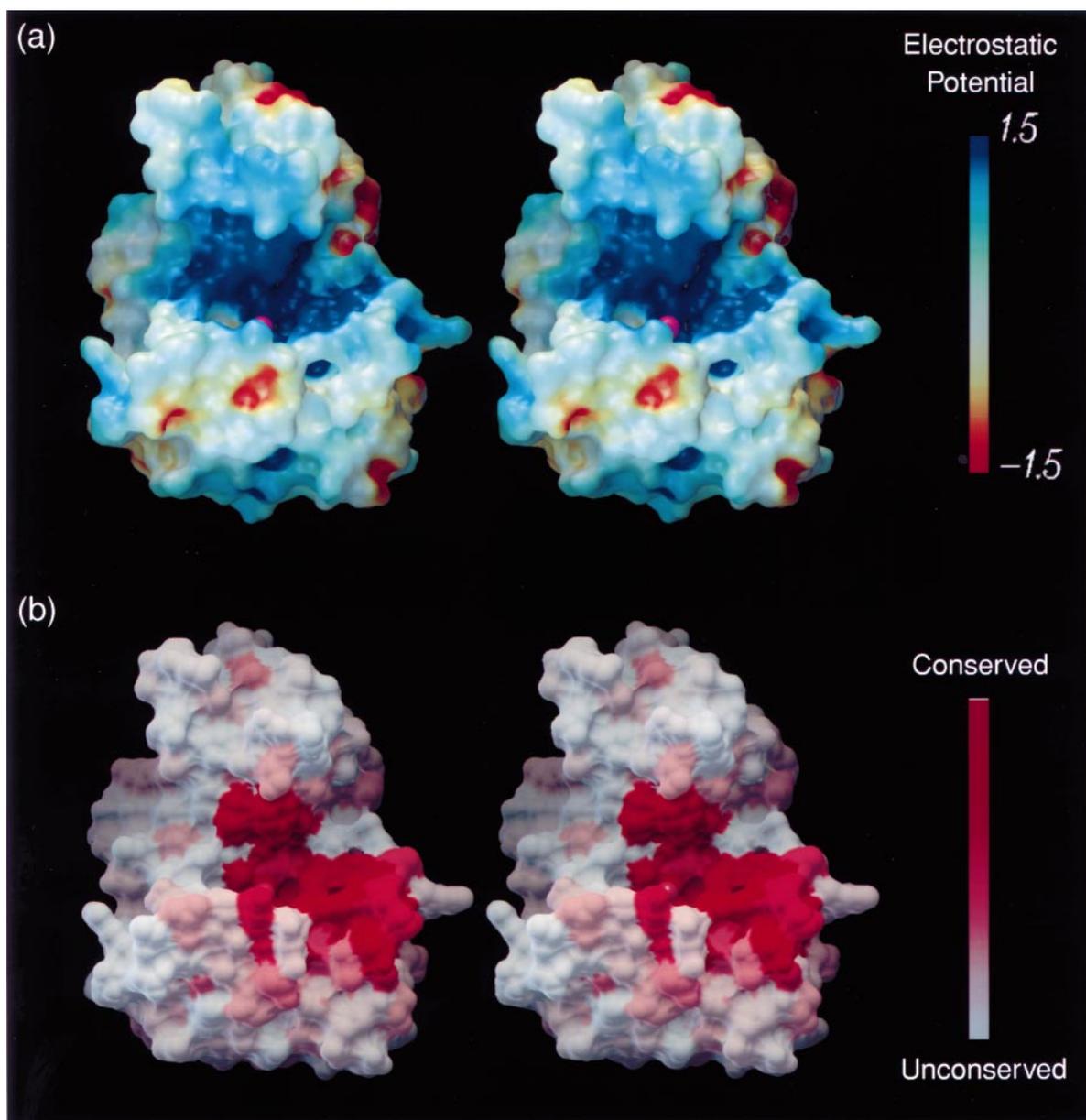


Figure 6. aRNase HIII substrate binding implicated by electrostatic and conserved surfaces. (a) Electrostatic surface of aRNase HIII, indicating the electrostatic complementarity important for multivalent interaction with substrate nucleic acid. The electrostatic surface potential was calculated using UHBD (Davis *et al.*, 1991), mapped onto the solvent-accessible surface calculated using a probe radius of 1.4 Å and displayed using AVS (Upson *et al.*, 1989). The surface is colored by electrostatic potential (+1.5 kT/e to $-1.5 kT/e$) with electropositive regions colored blue and electronegative regions colored red. (b) Solvent-accessible surface of aRNase HIII colored by residue conservation, based on sequence alignment of 19 type 2 RNase H sequences. Highly conserved regions are red, while unconserved regions are white. The position of the metal ion is represented by a pink sphere. The highly conserved regions of the central groove house the active site/metal ligands and key residues shown to effect substrate binding.

ant for substrate recognition and catalysis, similarly to that seen in the structure of endonuclease IV specifically binding abasic site-containing DNA (Hosfield *et al.*, 1999). Consistent with the importance of these Tyr finger and phosphate ruler motif residues in substrate binding, combined mutations caused the enzymatic activity of aRNase HIII to approach zero (data not shown), and thus support a model in which substrate binding specificity is tightly coupled to enzymatic activity.

Implications for substrate recognition and catalysis

The aRNase HIII structures and mutagenesis data presented here suggest a specific, testable molecular mechanism for substrate binding and positioning by the type 2 family of RNase H enzymes. The overall size and shape of the substrate binding groove suggests that the substrate nucleic acid is bound such that the helical axis is perpendicular to

Table 2. Kinetic parameters of wild-type and mutant aRNases HII

aRNase HII	K_m (μM)	k_{cat} (min^{-1})	k_{cat}/K_m ($\text{min}^{-1} \mu\text{M}^{-1}$)
Wild-type	0.06 ± 0.01	8.0 ± 0.23	133.3
R46A	3.60 ± 0.02	12.0 ± 0.30	3.3
K143A	1.88 ± 0.27	12.0 ± 0.36	6.4
R146A	1.60 ± 0.12	11.5 ± 0.75	7.2
Y164A	2.68 ± 0.27	12.5 ± 0.92	4.7

the central β -sheet. Importantly, an RNA-DNA hybrid duplex can bind with the 3' end of the RNA portion located near the active site metal, while ribonucleotides further upstream make minor groove 2'-OH or phosphate backbone contacts to the phosphate ruler motif formed by positively charged residues lining the floor of the binding groove (Figure 7). This implicated nucleic acid binding orientation positions the substrate such that the tip of $\alpha 7$, along with the tyrosine finger motif from the cap domain, are partially wedged into the major groove preferentially positioned to interact with the phosphate backbone, with Tyr164 positioned near the scissile bond.

The substrate binding mode proposed here for aRNase HII is consistent with earlier studies on

E. coli RNase HI, in which the substrate position was proposed to be similar with respect to the central β -sheet (Kanaya *et al.*, 1991b; Nakamura *et al.*, 1991; Katayanagi *et al.*, 1992). However, unlike the proposed *E. coli* RNase HI binding model, in which substrate binding and positioning requires roughly 1.5 turns of RNA/DNA hybrid substrate, the aRNase HII groove is only long enough to accommodate a single turn of an A-form double helix. Furthermore, specific minor groove contacts appear to extend only four to five nucleotides upstream of the scissile bond. Thus, the binding mode proposed here may explain the ability of type 2 RNase H enzymes to specifically cleave the phosphodiester bond of a closed, covalent RNA-DNA hybrid duplex containing four ribonucleotides flanked by DNA, leaving a single ribonucleotide on the 5' end of the substrate (Ohtani *et al.*, 1999a). During lagging strand DNA synthesis, this activity, which is a hallmark of type 2 RNase H enzymes, creates a potentially lethal intermediate, should the ribonucleotide be incorporated into the newly synthesized DNA strand.

Biochemical experiments, which demonstrate the ability of FEN-1 to complement RNase H activity by removing the final ribonucleotide, along with these aRNase HII structural results, suggest how the cell may be protected from the product of type 2 RNase H activity. The underlying molecular mechanism that couples these activities is unknown; however, our structural and biochemical results indicate that binding specificity is mediated by the cap domain and supported by multivalent interactions with the substrate, suggesting that aRNase HII may be able to bind tightly to its reaction product and then be displaced by FEN-1 binding to its resulting substrate. This type of coordinated hand-off to couple pathway steps resembles that proposed for enzymes of the DNA base-excision repair pathway, in which structure-specific recognition is coupled to general endonuclease activity (Parikh *et al.*, 1998; Mol *et al.*, 2000; Tainer & Friedberg, 2000). Given the implicated importance of the cap domain in specific substrate binding, the toxicity of the free double-stranded nucleotide product, and the existence of functional FEN-1 in archaeobacteria, this novel cap domain in the type 2 RNase H family may also play a key role in mediating pathway coordination.

Our aRNase HII structural and mutational results address a poorly understood aspect of RNase H activity, which comes from the finding that both type 1 and type 2 enzymes bind A-form

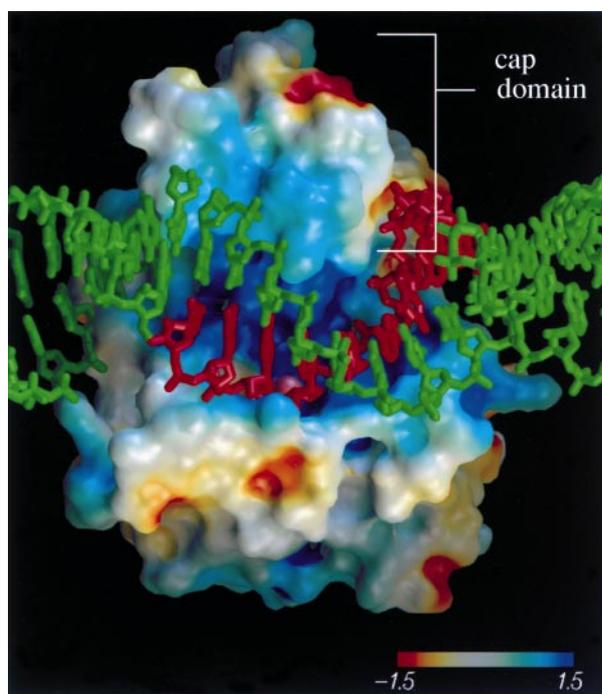


Figure 7. aRNase HII proposed substrate binding mode. Front view of aRNase HII in which an RNA/DNA hybrid (green), generated based on NDB:rh0001, has been docked onto the electrostatic surface. A portion of the RNA strand of the hybrid duplex (red) is shown threading through the substrate binding groove. The substrate was manually positioned by placing negatively charged regions of the minor groove and phosphate backbone near the phosphate ruler formed by residues Gln43, Arg46, Lys143 and Arg146, which are proposed to mediate substrate binding interactions.

duplex RNA, and RNA/DNA hybrid duplexes, but these enzymes only cleave the hybrid duplex (Lima *et al.*, 1997; Wu *et al.*, 1999). These binding studies indicate that while the substrate binding region may appear optimized to bind an A-form double helix, the catalytic mechanism requires a specific feature or flexibility in the substrate duplex. Solution structures of RNA/DNA hybrid duplexes highlight the flexible nature of RNA/DNA hybrid duplexes, compared to canonical A-form RNA duplexes (Fedoroff *et al.*, 1993; Gyi *et al.*, 1996, 1998). Our combined structural and kinetic analyses of aRNase HIII indicate that Tyr164 plays an important role in substrate binding, and this tyrosine side-chain is furthermore structurally poised to couple a hinge movement in the cap domain to substrate binding through the tyrosine-finger motif. Such a functionally coupled conformational change may be mediated through protein-nucleic acid interactions near the floor of the groove (Lys143, Arg46, Gln43), and/or by movement of the β 1- β 2 loop through modified interactions with Arg146 in the presence of substrate. In this scenario, the rigidity of an A-form double helix may limit conformational change, preventing the proper interaction of Tyr164 with the substrate, and thus blocking catalysis. Conversely, the structural plasticity of an RNA/DNA hybrid duplex may accommodate such movement. Thus, the shape of the substrate binding groove, the positioning of positively charged residues including the phosphate ruler motif within the groove, and the conformational flexibility of the tyrosine finger motif may couple general structural complementarity to substrate-mediated conformational change in order to achieve catalytic specificity.

These aRNase HIII structural and mutational results, together with the crystal structure of mRNase HIII (Lai *et al.*, 2000), provide important experimental grounds for comparisons and identify testable hypotheses to study the activity and function of type 2 RNase H enzymes, including the major mammalian RNase H. Importantly, the structurally and biochemically defined aRNase HIII active site and binding groove suggest substrate interactions that differ from those implicated from mRNase HIII alone. In particular, kinetic analysis of key substrate binding residues in aRNase HIII suggests that important substrate contacts are made within the binding groove, distinct from the Mes binding site seen in the mRNase HIII structure. Our aRNase HIII results also implicate key acidic residues important for catalysis, identifying somewhat different roles for the active site residues than those proposed from the mRNase HIII structure alone. These differences thus suggest two independent testable hypotheses for both substrate binding and active site catalysis, and highlight the importance and significance of using different experimental systems to examine the function of the major mammalian RNase H family of enzymes. Moreover, conserved type 2 RNase H structural features identified here, including the two-domain architec-

ture comprised of an RNase H-fold plus the novel helix-loop-helix C-terminal cap domain stabilized by a hydrophobic interface, verify the presence of structurally conserved motifs that distinguish this type 2 RNase H class of enzymes from other RNase H enzymes within the context of a common nuclease family.

Materials and Methods

Cloning of aRNase HIII

An open reading frame in the *A. fulgidus* genome corresponding to a type 2 RNase H was identified in the NCBI database (<http://www.ncbi.nlm.nih.gov/>), and amplified by the polymerase chain reaction (PCR) using *A. fulgidus* genomic DNA purchased from ATCC (Rockville, MD). PCR primers were designed to incorporate an *Nde*I restriction site at the translation start site, and a *Bam*HI restriction site in the 3'-untranslated region. The PCR product was digested with *Nde*I and *Bam*HI, ligated into pET-28b (Novagen, Madison, WI), digested with the same enzymes, and a representative positive clone was sequenced for verification and designated pET28b-RH2.

aRNase HIII overexpression and purification

pET28b-RH2 was used to transform *E. coli* BL21(DE3), and a single colony was inoculated in 3 ml of Luria broth (LB) with 30 μ g/ml kanamycin and cultured overnight at 37°C. This culture was used to inoculate nine liters of LB kanamycin, and induced with 0.4 mM isopropyl β -D-thiogalactopyranoside when the cell density reached an A_{600} of 0.6. Following induction, the cells were grown for four to five hours, pelleted, and resuspended in buffer B (10 mM Tris (pH 7.5), 150 mM NaCl, 10 mM imidazole) by brief sonication. Cells were lysed by heating the suspension at 75°C for 45 minutes and then immediately cooled to 0°C in an ice-water bath. This protocol, which takes advantage of the archaeobacterial protein's thermostability, not only lysed the cells but also precipitated the majority of contaminating mesophilic proteins. The resulting solution was centrifuged at 25,000 g, and the supernatant was loaded onto a 25 ml Ni-NTA-Sepharose column (Pharmacia Biotech, Piscataway, NJ) pre equilibrated with buffer B. After washing with 100 ml of buffer B, followed by 100 ml of buffer B containing 20 mM imidazole, aRNase HIII was eluted with buffer B containing 500 mM imidazole and dialyzed against buffer C (50 mM Mes (pH 6.5), 200 mM NaCl). At this point, aRNase HIII was 90% pure as judged by SDS-PAGE analysis. To remove traces of contaminants, the protein solution was loaded onto a 20 ml Porous HQ column (Pharmacia, Piscataway, NJ) pre-equilibrated with buffer C. Pure aRNase HIII eluted immediately upon washing with 50 ml of buffer C, free of nucleic acid and protein contaminants.

Selenomethionine-labeled protein was produced by overexpression of aRNase HIII from the pET28b-RH2 plasmid in *E. coli* B834 cells (Novagen, Madison, WI). These cells are methionine-auxotroph BL21(DE3) derivatives, which incorporate selenium-labeled methionine into expressed proteins when the growth medium is supplemented with seleno-L-methionine (Sigma, St. Louis, MO). The over-expression conditions and purification strategy was identical with that used for wild-type aRNase HIII, except buffer B contained 10 mM

β -mercaptoethanol (Sigma, St. Louis, MO) and buffer C contained 1 mM dithiothreitol (DTT).

Crystallization, data collection, and processing

Crystals of wild-type aRNase HIII were grown by hanging drop vapor diffusion by mixing equal volumes of protein solution (40 mg/ml) with the precipitant solution (10% polyethylene glycol monomethylether (MPEG) 5000, 10% butanol, 20 mM sodium citrate, 100 mM Tris (pH 8.5)). Single, diffraction quality crystals grew over a period of three days to a size of 0.2 mm \times 0.2 mm \times 0.4 mm. The crystals belong to the hexagonal space group $P6_522$ ($a, b = 74 \text{ \AA}$, $c = 140 \text{ \AA}$) and contain one molecule of aRNase HIII in the crystallographic asymmetric unit. A cobalt(III) hexammine chloride complex was obtained by soaking crystals in mother liquor fortified with 10 mM cobalt(III) hexammine chloride for one hour. All data were collected at -180°C from crystals flash-cooled in mother liquor containing 10% glycerol at the Stanford Synchrotron Radiation Laboratory (SSRL) on beam-line 9-1, and 9-2 for wild-type and cobalt derivatives, respectively.

Selenomethionine-substituted protein was crystallized in the tetragonal space group $P4_32_12$ ($a, b = 74 \text{ \AA}$, $c = 105 \text{ \AA}$), by using the hanging drop vapor diffusion method by mixing equal volumes of protein solution (20 mg/ml) with precipitant (12% MPEG 5000, 10% butanol, 1 mM DTT, 20 mM sodium citrate, 100 mM Tris (pH 8.5)). Experimental phases were obtained by multiwavelength anomalous dispersion (MAD) from a single crystal of SeMet-aRNase HIII using four wavelengths near the selenium absorption edge. Data were collected at -180°C from crystals flash-cooled in mother liquor containing 10% glycerol at beamline 5.0.2 at the Advanced Light Source (ALS). All data were indexed and integrated with the programs DENZO and SCALE-PAK (Otwinowski, 1993), and data reduction was performed using the CCP4 suite of programs (Collaborative Computational Project Number 4, 1994).

Structure determination

Three of the expected selenium positions were located using the program SOLVE (Terwilliger & Berendzen, 1999), and experimental phases to 2.7 \AA resolution were calculated using SHARP (de La Fortelle & Bricogne, 1997), and improved by solvent flattening with SOLOMON (Collaborative Computational Project Number 4, 1994). The experimental maps were of sufficient quality to allow 192 of 205 residues to be manually modeled using the program XFIT (McRee, 1999). Positional and grouped temperature-factor refinement, followed by simulated annealing, and manual model building using the CNS program suite (Brünger *et al.*, 1998) gave an intermediate model that was used to phase the 1.95 \AA data set by molecular replacement using AMoRe (Navaza, 1994). Further refinement and manual rebuilding into bias-reduced simulated-annealing omit maps gave a final model with a crystallographic R value of 0.244 ($R_{\text{free}} = 0.269$) for all data with no σ cutoff. After refinement, structural comparison of the 1.95 \AA model from the hexagonal crystals to the crystallographically independent 2.7 \AA model from the tetragonal crystals showed a root-mean-square deviation of 0.9 \AA , as determined with SEQUOIA. The majority of these slight deviations are located in the cap domain, consistent with the idea of a domain movement during catalysis suggested

by the flexible link joining the domains. Electrostatic surface calculations were performed with UHBD (Davis *et al.*, 1991), structure superpositions and alignments were done with SEQUOIA (<http://www.scripps.edu/~bruns>), and Figures were generated using the Advanced Visualization Systems software (Upson *et al.*, 1989).

RNase H activity assays

The RNA/DNA substrates used in this study were assembled from the hybrid oligonucleotides: RD18:5'-(UUGCAUGCC)d(TGCACCTCG) and RD37: 5'-(GGGAACAAAAGCUUGCAUGCC)d(TGCAGGTCC ACTCTAG) and complementary DNA oligonucleotides. The oligoribonucleotides were 5'-end-labeled using [γ - ^{32}P]ATP and phage T4 polynucleotide kinase and hybridized in reaction buffer (20 mM Tris, 10 mM MgCl_2 , 25 mM NaCl) containing 100 nM labeled oligoribonucleotide and 200 nM complementary DNA. Hybridization reactions were heated to 72°C for five minutes, and then annealed by slowly cooling to room temperature.

RNase H assays were performed in standard reaction conditions (50 mM Tris-HCl (pH 8.0), 1 mM MgCl_2 , 1.5 μM bovine serum albumin, 5 mM β -mercaptoethanol) with 0.1 μM RNA-DNA/DNA hybrid. Enzymes were diluted from concentrated stocks in 20 mM Tris-HCl (pH 8.0), 100 mM NaCl, 1 mM EDTA, 10 mM β -mercaptoethanol to the final stock concentrations in 50 mM Tris-HCl (pH 8.0), 100 mM NaCl, 1.5 μM bovine serum albumin, 1 mM DTT, 30% glycerol prior to usage. Enzyme concentrations for a typical reaction ranged from 0.1 nM to 10 nM and all reactions were incubated at 30°C for specified time intervals and quenched with an equal volume of stop solution (USB, Cleveland, OH). A base hydrolysis ladder was prepared by incubation of 5'-end-labeled RNA at 90°C for five minutes in 100 mM NaCO_3 (pH 9.0).

Steady-state enzyme kinetics

Steady-state kinetic reactions were carried out in the same conditions as described above by using the RD18 duplex at concentrations ranging from 0 to 5 μM . Initial velocity measurements were determined by measuring differences in substrate and product intensities on a denaturing 15% polyacrylamide gel containing 7 M urea, by PhosphorImager (Molecular Dynamics, San Diego, CA) as a function of time. Velocity measurements were expressed as the rate of the converted substrate concentration (nanomolar) over time (minutes) and one unit of enzymatic activity is defined as the amount of enzyme producing 1 μmol of product per minute. Specific activity was defined as the enzymatic activity per milligram of protein. As aRNaseHIII substrate hydrolysis follows Michaelis-Menten kinetics, the kinetic parameters, K_m and V_{max} were determined from an Eadie-Scatchard plot. Data from multiple assays were analyzed by linear regression methods using software provided in the CricketGraph data analysis package (Cricket Software, Philadelphia, PA).

Protein Data Bank accession codes

The coordinates of aRNase HIII and of the aRNase HIII cobalt(III) hexammine chloride inhibitor complex have

been deposited in the RCSB Protein Data Bank under ID codes 1I39 and 1I3A, respectively.

Acknowledgments

We thank the staff of Beamline 9-1, 9-2 at the SSRL and Beamline 5.0.2 at the ALS for help with data collection, Xinwei Li for technical assistance, and C. D. Mol, S. S. Parikh, C. D. Putnam, D. S. Shin and D. S. Daniels for helpful discussions. Work in the authors' laboratory is supported by the Skaggs Institute for Chemical Biology, the National Institutes of Health (CA81967 to J.A.T.) and (CA85344 to B.H.S.), and a Graduate Research Fellowship from the National Science Foundation (B.R.C.).

References

- Ariyoshi, M., Vassilyev, D. G., Iwasaki, H., Nakamura, H., Shinagawa, H. & Morikawa, K. (1994). Atomic structure of the RuvC resolvase: a holliday junction-specific endonuclease from *E. coli*. *Cell*, **78**, 1063-1072.
- Arudchandran, A., Cerritelli, S., Narimatsu, S., Itaya, M., Shin, D. Y., Shimada, Y. & Crouch, R. (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.
- Beese, L. S. & Steitz, T. A. (1991). Structural basis for the 3'-5' exonuclease activity of *Escherichia coli* DNA polymerase I: a two metal ion mechanism. *EMBO J.* **10**, 25-33.
- Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). Crystallography and NMR system: a new software suite for macromolecular structure determination. *Acta Crystallog. sect. D*, **54**, 905-921.
- Bruns, C. M. (1999). *SEQUOIA ver 0.9.7*, The Scripps Research Institute.
- Busen, W., Peters, J. H. & Hausen, P. (1977). Ribonuclease H levels during the response of bovine lymphocytes to concanavalin A. *Eur. J. Biochem.* **74**, 203-208.
- Cathala, G., Rech, J., Huet, J. & Jeanteur, P. (1979). Isolation and characterization of two types of ribonucleases H in Krebs II ascites cells. *J. Biol. Chem.* **254**, 7353-7361.
- Collaborative Computational Project Number 4 (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallog. sect. D*, **50**, 760-763.
- Conn, G. L., Brown, T. & Leonard, G. A. (1999). The crystal structure of the RNA/DNA hybrid r(GAA-GAGAAGC). d(GCTCTCTTC) shows significant differences to that found in solution. *Nucl. Acids Res.* **27**, 555-561.
- Crooke, E. (1995). DNA synthesis initiated at oriC: *in vitro* replication reactions. *Methods Enzymol.* **262**, 500-506.
- Crouch, R. J. & Dirksen, M. L. (1982). Ribonucleases H. In *Nucleases*, pp. 211-241, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Davies, D. R., Gorshin, I. Y., Reznikoff, W. S. & Rayment, I. (2000). Three-dimensional structure of the Tn5 synaptic complex transposition intermediate. *Science*, **289**, 77-85.
- Davis, M. E., Madura, J. D., Luty, B. A. & McCammon, J. A. (1991). Electrostatics and diffusion of molecules in solution: simulations with the University of Houston Brownian dynamics program. *Comput. Phys. Commun.* **62**, 187-197.
- de La Fortelle, E. & Bricogne, G. (1997). Maximum-likelihood heavy-atom parameter refinement for the multiple isomorphous replacement and multiwavelength anomalous diffraction methods. *Methods Enzymol.* **276**, 472-494.
- Dyda, F., Hickman, A. B., Jenkins, T. M., Engelman, A., Craigie, R. & Davies, D. R. (1994). Crystal structure of the catalytic domain of HIV-1 integrase: similarity to other polynucleotidyl transferases. *Science*, **266**, 1981-1986.
- Eder, P. S., Walder, R. Y. & Walder, J. A. (1993). Substrate specificity of human RNase H1 and its role in excision repair of ribose residues misincorporated in DNA. *Biochimie*, **75**, 123-126.
- Fedoroff, O. Y. U., Salazar, M. & Reid, B. R. (1993). Structure of a DNA:RNA hybrid duplex. Why RNase H does not cleave pure RNA. *J. Mol. Biol.* **233**, 509-523.
- Frank, P., Braunshofer-Reiter, C., Poltl, A. & Holzmann, K. (1998a). Cloning, subcellular localization and functional expression of human RNase HIII. *Biol. Chem.* **379**, 1407-1412.
- Frank, P., Braunshofer-Reiter, C. & Wintersberger, U. (1998b). Yeast RNase H(35) is the counterpart of the mammalian RNase H1, and is evolutionarily related to prokaryotic RNase HIII. *FEBS Letters*, **421**, 23-26.
- Frank, P., Braunshofer-Reiter, C., Wintersberger, U., Grimm, R. & Busen, W. (1998c). Cloning of the cDNA encoding the large subunit of human RNase H1, a homologue of the prokaryotic RNase HIII. *Proc. Natl Acad. Sci. USA*, **95**, 12872-12877.
- Frank, P., Braunshofer-Reiter, C., Karwan, A., Grimm, R. & Wintersberger, U. (1999). Purification of *Saccharomyces cerevisiae* RNase H(70) and identification of the corresponding gene. *FEBS Letters*, **450**, 251-256.
- Gyi, J. I., Conn, G. L., Lane, A. N. & Brown, T. (1996). Comparison of the thermodynamic stabilities and solution conformations of DNA. RNA hybrids containing purine-rich and pyrimidine rich strands with DNA and RNA duplexes. *Biochemistry*, **35**, 12538-12548.
- Gyi, J. I., Lane, A. N., Conn, G. L. & Brown, T. (1998). Solution structures of DNA. RNA hybrids with purine-rich and pyrimidine-rich strands: comparison with the homologous DNA and RNA duplexes. *Biochemistry*, **37**, 73-80.
- Haruki, M., Noguchi, E., Nakai, C., Liu, Y. Y., Oobatake, M., Itaya, M. & Kanaya, S. (1994). Investigating the role of conserved residue Asp134 in *Escherichia coli* ribonuclease HI by site-directed random mutagenesis. *Eur. J. Biochem.* **220**, 623-631.
- Haruki, M., Noguchi, E., Kanaya, S. & Crouch, R. J. (1997). Kinetic and stoichiometric analysis for the binding of *Escherichia coli* ribonuclease HI to RNA-DNA hybrids using surface plasmon resonance. *J. Biol. Chem.* **272**, 22015-22022.
- Haruki, M., Hayashi, K., Kochi, T., Muroya, A., Koga, Y., Morikawa, M., Imanaka, T. & Kanaya, S. (1998). Gene cloning and characterization of recombinant

- RNase HII from a hyperthermophilic archaeon. *J. Bacteriol.* **180**, 6207-6214.
- Hosfield, D. J., Guan, Y., Haas, B. J., Cunningham, R. P. & Tainer, J. A. (1999). Structure of the DNA repair enzyme endonuclease IV and its DNA complex: double-nucleotide flipping at abasic sites and three-metal-ion catalysis. *Cell*, **98**, 397-408.
- Huang, L., Kim, Y., Turchi, J. J. & Bambara, R. A. (1994). Structure-specific cleavage of the RNA primer from Okazaki fragments by calf thymus RNase HI. *J. Biol. Chem.* **269**, 25922-25927.
- Itaya, M. (1990). Isolation and characterization of a second RNase H (RNase HII) of *Escherichia coli* K-12 encoded by the *rnhB* gene. *Proc. Natl Acad. Sci. USA*, **87**, 8587-8591.
- Itaya, M., Omori, A., Kanaya, S., Crouch, R. J., Tanaka, T. & Kondo, K. (1999). Isolation of RNase H genes that are essential for growth of *Bacillus subtilis* 168. *J. Bacteriol.* **181**, 2118-2123.
- Kanaya, S., Kohara, A., Miura, Y., Sekiguchi, A., Iwai, S., Inoue, H., Ohtsuka, E. & Ikehara, M. (1990). Identification of the amino acid residues involved in an active site of *Escherichia coli* ribonuclease H by site-directed mutagenesis. *J. Biol. Chem.* **265**, 4615-4621.
- Kanaya, S., Katsuda, C., Kimura, S., Nakai, T., Kitakuni, E., Nakamura, H., Katayanagi, K., Morikawa, K. & Ikehara, M. (1991a). Stabilization of *Escherichia coli* ribonuclease H by introduction of an artificial disulfide bond. *J. Biol. Chem.* **266**, 6038-6044.
- Kanaya, S., Katsuda-Nakai, C. & Ikehara, M. (1991b). Importance of the positive charge cluster in *Escherichia coli* ribonuclease HI for the effective binding of the substrate. *J. Biol. Chem.* **266**, 11621-11627.
- Kanaya, E., Uchiyama, Y., Ohtsuka, E., Ueno, Y., Ikehara, M. & Kanaya, S. (1994). Kinetic analyses of DNA-linked ribonucleases H with different sizes of DNA. *FEBS Letters*, **354**, 227-231.
- Kashiwagi, T., Jeanteur, D., Haruki, M., Katayanagi, K., Kanaya, S. & Morikawa, K. (1996). Proposal for new catalytic roles for two invariant residues in *Escherichia coli* ribonuclease HI. *Protein Eng.* **9**, 857-867.
- Katayanagi, K., Miyagawa, M., Matsushima, M., Ishikawa, M., Kanaya, S., Ikehara, M., Matsuzaki, T. & Morikawa, K. (1990). Three-dimensional structure of ribonuclease H from *E. coli*. *Nature*, **347**, 306-309.
- Katayanagi, K., Miyagawa, M., Matsushima, M., Ishikawa, M., Kanaya, S., Nakamura, H., Ikehara, M., Matsuzaki, T. & Morikawa, K. (1992). Structural details of ribonuclease H from *Escherichia coli* as refined to an atomic resolution. *J. Mol. Biol.* **223**, 1029-1052.
- Katayanagi, K., Okumura, M. & Morikawa, K. (1993). Crystal structure of *Escherichia coli* RNase HI in complex with Mg^{2+} at 2.8 Å resolution: proof for a single Mg^{2+} -binding site. *Proteins: Struct. Funct. Genet.* **17**, 337-346.
- Kayana, S. & Ikehara, M. (1995). Functions and structures of ribonuclease H enzymes. *Subcell. Biochem.* **24**, 377-422.
- Keck, J. L., Goedken, E. R. & Marqusee, S. (1998). Activation/attenuation model for RNase H. A one-metal mechanism with second-metal inhibition. *J. Biol. Chem.* **273**, 34128-34133.
- Krynetskaia, N. F., Krynetski, E. Y. & Evans, W. E. (1999). Human RNase H-mediated RNA cleavage from DNA-RNA duplexes is inhibited by 6-deoxythioguanosine incorporation into DNA. *Mol. Pharmacol.* **56**, 841-848.
- Lai, L., Yokota, H., Hung, L., Kim, R. & Kim, S. (2000). Crystal Structure of archaeal RNase HII: a homologue of human major RNase H. *Structure*, **8**, 897-904.
- Lima, W. F. & Crooke, S. T. (1997). Cleavage of single strand RNA adjacent to RNA-DNA duplex regions by *Escherichia coli* RNase H1. *J. Biol. Chem.* **272**, 27513-27516.
- Lima, W. F., Mohan, V. & Crooke, S. T. (1997). The influence of antisense oligonucleotide-induced RNA structure on *Escherichia coli* RNase H1 activity. *J. Biol. Chem.* **272**, 18191-18199.
- McRee, D. E. (1999). Xtalview/Xfit-a versatile program for manipulating atomic coordinates and electron density. *J. Struct. Biol.* **125**, 156-165.
- Mian, I. S. (1997). Comparative sequence analysis of ribonucleases HII, III, II PH and D. *Nucl. Acids Res.* **25**, 3187-3195.
- Mol, C. D., Izumi, T., Mitra, S. & Tainer, J. A. (2000). DNA-bound structures and mutants reveal abasic DNA binding by APE1 coordinates DNA repair. *Nature*, **403**, 451-456.
- Murante, R. S., Henricksen, L. A. & Bambara, R. A. (1998). Junction ribonuclease: an activity in Okazaki fragment processing. *Proc. Natl Acad. Sci. USA*, **95**, 2244-2249.
- Nakamura, H., Oda, Y., Iwai, S., Inoue, H., Ohtsuka, E., Kanaya, S., Kimura, S., Katsuda, C., Katayanagi, K. & Morikawa, K. (1991). How does RNase H recognize a DNA:RNA hybrid?. *Proc. Natl Acad. Sci. USA*, **88**, 11535-11539.
- Navaza, L. (1994). AMoRe: an automated package for molecular replacement. *Acta Crystallog. sect. A*, **50**, 157-163.
- Ohtani, N., Haruki, M., Morikawa, M., Crouch, R. J., Itaya, M. & Kanaya, S. (1999a). Identification of the genes encoding Mn^{2+} -dependent RNase HII and Mg^{2+} -dependent RNase HIII from *Bacillus subtilis*: classification of RNases H into three families. *Biochemistry*, **38**, 605-618.
- Ohtani, N., Haruki, M., Morikawa, M. & Kanaya, S. (1999b). Molecular diversities of RNases H. *J. Biosci. Bioeng.* **88**, 12-19.
- Ohtani, N., Haruki, M., Muroya, A., Morikawa, M. & Kanaya, S. (2000). Characterization of ribonuclease HIII from *Escherichia coli* overproduced in a soluble form. *J. Biochem. (Tokyo)*, **127**, 895-899.
- Otwinowski, Z. (1993). *Data Collection and Processing* (Sawyer, L., Isaacs, N. & Bailey, S., eds), pp. 56-62, Science and Engineering Research Council, Warrington, UK.
- Parikh, S. S., Mol, C. D., Slupphaug, G., Bharati, S., Krokan, H. E. & Tainer, J. A. (1998). Base-excision repair initiation revealed by crystal structures and DNA-binding kinetics of human uracil-DNA glycosylase bound to DNA. *EMBO J.* **17**, 5214-5226.
- Parikh, S. S., Walcher, G., Jones, G. D., Slupphaug, G., Krokan, H. E., Blackburn, G. M. & Tainer, J. A. (2000). Uracil-DNA glycosylase-DNA substrate and product structures: conformational strain promotes catalytic efficiency by coupled stereoelectronic effects. *Proc. Natl Acad. Sci. USA*, **97**, 5083-5088.
- Qiu, J., Qian, Y., Chen, V., Guan, M. X. & Shen, B. (1999a). Human exonuclease 1 functionally complements its yeast homologues in DNA recombination, RNA primer removal, and mutation avoidance. *J. Biol. Chem.* **274**, 17893-17900.

- Qiu, J., Qian, Y., Frank, P., Wintersberger, U. & Shen, B. (1999b). *Saccharomyces cerevisiae* RNase H(35) functions in RNA primer removal during lagging-strand DNA synthesis, most efficiently in cooperation with Rad27 nuclease. *Mol. Cell. Biol.* **19**, 8361-8371.
- Rice, P., Craigie, R. & Davies, D. R. (1996). Retroviral integrases and their cousins. *Curr. Opin. Struct. Biol.* **6**, 76-83.
- Saenger, W. (1984). *Principles of Nucleic Acid Structure*, pp. 221-282, Springer-Verlag, New York.
- Slupphaug, G., Mol, C. D., Kavli, B., Arvai, A. S., Krokan, H. E. & Tainer, J. A. (1996). A nucleotide-flipping mechanism from the structure of human uracil-DNA glycosylase bound to DNA. *Nature*, **384**, 87-92.
- Stein, H. & Hausen, P. (1969). Enzyme from calf thymus degrading the RNA moiety of DNA-RNA Hybrids: effect on DNA-dependent RNA polymerase. *Science*, **166**, 393-395.
- Steitz, T. A. & Steitz, J. A. (1993). A general two-metal-ion mechanism for catalytic RNA. *Proc. Natl Acad. Sci. USA*, **90**, 6498-6502.
- Tainer, J. A. & Friedberg, E. C. (2000). Dancing with the elephants: envisioning the structural biology of DNA repair pathways. *Mutat. Res.* **460**, 139-141.
- Terwilliger, T. C. & Berendzen, J. (1999). Automated MAD and MIR structure solution. *Acta Crystallog. sect. D*, **55**, 849-861.
- Turchi, J. J., Huang, L., Murante, R. S., Kim, Y. & Bambara, R. A. (1994). Enzymatic completion of mammalian lagging-strand DNA replication. *Proc. Natl Acad. Sci. USA*, **91**, 9803-9807.
- Upson, C., Faulhaber, T., Kamins, D., Laidlaw, D., Schlegel, D., Vroom, J., Gurwitz, R. & Vandam, A. (1989). The application visualization system - a computational environment for scientific visualization. *IEEE Comp. G.* **9**, 30-42.
- Waga, S. & Stillman, B. (1994). Anatomy of a DNA replication fork revealed by reconstitution of SV40 DNA replication *in vitro*. *Nature*, **369**, 207-212.
- Waga, S., Bauer, G. & Stillman, B. (1994). Reconstitution of complete SV40 DNA replication with purified replication factors. *J. Biol. Chem.* **269**, 10923-10934.
- Wu, H., Lima, W. F. & Crooke, S. T. (1999). Properties of cloned and expressed human RNase H1. *J. Biol. Chem.* **274**, 28270-28278.
- Yang, W. & Steitz, T. A. (1995). Recombining the structures of HIV integrase, RuvC and RNase H. *Structure*, **3**, 131-134.
- Yang, W., Hendrickson, W. A., Crouch, R. J. & Satow, Y. (1990a). Structure of ribonuclease H phased at 2 Å resolution by MAD analysis of the selenomethionyl protein. *Science*, **249**, 1398-1405.
- Yang, W., Hendrickson, W. A., Kalman, E. T. & Crouch, R. J. (1990b). Expression, purification, and crystallization of natural and selenomethionyl recombinant ribonuclease H from *Escherichia coli*. *J. Biol. Chem.* **265**, 13553-13559.

Edited by K. Morikawa

(Received 6 November 2000; received in revised form 26 December 2000; accepted 22 January 2001)