DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD

Crystal structure of xenotropic murine leukaemia virus-related virus (XMRV) ribonuclease H

(Article begins on next page)

Original Paper

Crystal structure of xenotropic murine leukaemia virus-related virus (XMRV) ribonuclease H

Ju Hee KIM*[1], Sunghyun KANG*[1], Suk-Kyeong JUNG*, Keum Ran YU*, Sang J. CHUNG†, Bong Hyun CHUNG†, Raymond L. ERIKSON‡§, Bo Yeon KIM‡[2] and Seung Jun KIM*[2]

*Medical Proteomics Research Center, Korea Research Institute of Bioscience and Biotechnology, 52 Eoeun-Dong, Yuseong-Gu, Daejeon 305-333, Republic of Korea, †BioNanotechnology Research Center, Korea Research Institute of Bioscience and Biotechnology, 52 Eoeun-Dong, Yuseong-Gu, Daejeon 305-333, Republic of Korea, ‡Chemical Biology Research Center, Korea Research Institute of Bioscience and Biotechnology, 52 Eoeun-Dong, Yuseong-Gu, Daejeon 305-333, Republic of Korea, and §Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA 02138, U.S.A.

Synopsis

RNase H (retroviral ribonuclease H) cleaves the phosphate backbone of the RNA template within an RNA/DNA hybrid to complete the synthesis of double-stranded viral DNA. In the present study we have determined the complete structure of the RNase H domain from XMRV (xenotropic murine leukaemia virus-related virus) RT (reverse transcriptase). The basic protrusion motif of the XMRV RNase H domain is folded as a short helix and an adjacent highly bent loop. Structural superposition and subsequent mutagenesis experiments suggest that the basic protrusion motif plays a role in direct binding to the major groove in RNA/DNA hybrid, as well as in establishing the co-ordination among modules in RT necessary for proper function.

Key words: *Escherichia coli*, murine leukaemia virus, reverse transcriptase, ribonuclease H, xenotropic murine leukaemia virus-related virus (XMRV)

Cite this article as: Kim, J.H., Kang, S., Jung, S.-K., Yu, K.R., Chung, S.J., Chung, B.H., Erikson, R.L., Kim, B.Y. and Kim, S.J. (2012) Crystal structure of xenotropic murine leukaemia virus-related virus (XMRV) ribonuclease H. Biosci. Rep. **32**, 455–463

INTRODUCTION

XMRV (xenotropic murine leukaemia virus-related virus) is a gamma-retrovirus, a genus of the retroviridae family [1]. As its name suggests, the XMRV genomic sequence is closely related to that of the MLVs (murine leukaemia viruses). XMRV was initially isolated from the tissues of prostate cancer patients [1] and subsequently found to be present in 6–27% of human prostate cancer cases [2]. Serological tests on prostate cancer patients have shown the presence of neutralizing antibodies to XMRV in plasma [3]. XMRV has also been detected in 9.9% of immuno-suppressed transplant patients from Germany [4].

Mapping of XMRV integration sites has revealed a strong preference for structurally open transcription regulatory regions of chromosomes that are associated with cancer breakpoints, common fragile sites, microRNA and cancer-related genes [5]. However, the idea that XMRV also infects human populations has been met with skepticism. The replication of XMRV is highly inhib-ited by the human antiviral restriction factors APOBEC3 (apoli-poprotein B mRNA editing enzyme catalytic polypeptide 3) and tetherin, suggesting that a replication-competent dose is not effectively achieved in humans [6–10]. Recently, Paprotka et al. [9] discovered two new murine endogenous pro-retroviruses: PreXMRV-1 and PreXMRV-2, in several strains of laboratory mice. PreXMRV-1 and PreXMRV-2 share 99.92% sequence identity with XMRV, and retroviral recombination between them can generate XMRV. On the basis of studies on the XMRV-producing human prostate cancer cell lines CWR22R*v1* and CWR-R1 and their progenitor tumour xenograft CRW22, they concluded that XMRV infections were caused by contamination during *in vivo* passaging in nude mice. However, the possibility of occasional XMRV infection in humans has not been completely ruled out.

Retroviral RTs (reverse transcriptases) are replicative polymerases that are essential for the replication and the survival of retroviruses during infection. RTs have two functional domains: a DNA polymerase domain that replicates DNA strands

---

using either DNA or RNA as a template and an RNase H (retroviral ribonuclease H) domain that hydrolyses the RNA strand in an RNA/DNA hybrid. Despite their functional similarities, the general structures of RTs are diverse. Among the many retroviruses, HIV-1 and MoMLV (Moloney murine leukaemia virus) RTs have been most extensively studied biochemically and structurally [11]. The RT of HIV is a heterodimer of p66 and p51 subunits. The larger p66 contains five structural domains: fingers, palm, thumb, connection and RNase H domains. The p51 subunit is generated by proteolytic cleavage of p66 and lacks the RNase H domain. The p51 subunit is enzymatically inactive and serves a structural role in the heterodimeric RT complexes. The RT of MoMLV is a 75-kDa monomer that contains all five domains on a single polypeptide chain. XMRV is almost identical with MoMLV in terms of genome sequence, protein composition and morphology, suggesting that these two viruses might have a highly similar catalytic mechanism as well. Mutational inactivation of RNase H prevents replication of HIV [12]. Therefore RNase H has been considered an attractive target for anti-HIV drugs, and many such inhibitors have been extensively explored [13].

The three-dimensional structure of RNase H has been elucidated in several organisms, including retroviruses (HIV-1 and MoMLV), bacteria [*Escherichia coli*, *Thermus thermophilus* HB8 and *Bacillus halodurans* (*Bh*)] and eukaryotes (humans) [14–19]. These RNase Hs have been shown to share tertiary folding and a common catalytic mechanism. RNase H has a compact structure comprising a central β-sheet of mixed antiparallel and parallel β-strands surrounded by four or five helices on or beneath the β-sheet. The helices and strands are arranged as an $\alpha\beta\alpha$ Rossmann-like fold. The catalytic core of RNase H consists of four conserved acidic residues that are engaged in binding bivalent cations crucial for catalysis. RNase Hs can be further grouped into two categories based on the presence or absence of a basic protrusion motif [11,20]. RNase Hs from *E. coli*, human, MoMLV and XMRV contain a basic protrusion motif, whereas HIV-1, *Bh* and avian sarcoma leukaemia virus do not. Deletion of the basic protrusion markedly reduces the binding affinity of *E. coli* RNase H for RNA/DNA hybrids in association with a decrease in Mg^{2+}-dependent catalytic activity. Deletion of this region in MoMLV RNase H results in minimal catalytic activity and replication defects, indicating that the basic protrusion contributes to important *in vivo* and *in vitro* functions [21,22]. Crystal structures of MoMLV RNase H [17] and XMRV RNase H [23] have been determined. Notably, however, the former structure lacked 11 residues that constitute the basic protrusion motif and the His (histidine)-containing loop responsible for binding RNA/DNA hybrid and catalysis, and the latter also lacked the basic protrusion motif. However, the crystal structure of XMRV RNase H that includes all residues has been determined very recently [24].

In the present paper, we report the crystal structure of XMRV RNase H at 2.6 Å (1 Å = 0.1 nm) resolution. Our structure includes all residues of the RNase H domain. Several key differences between the RNase H domains of HIV and XMRV provide an explanation for why the RNase Hs of XMRV and MoMLV

possess intrinsic hydrolysing activities, whereas RNase H from HIV does not. Our structure also provides a framework for developing specific RNase H inhibitors for XMRV and other related retroviruses.

MATERIALS AND METHODS

Plasmid construction and protein purification

A full-length clone of XMRV-VP62 RNase H (GenBank® accession no. BC008803) was sub-cloned into pET28a (residues 498–672). The XMRV-VP62 RNase H domain was expressed in the BL21 (DE3) strain of *E. coli*. Cells were grown at 295 K. after induction with 0.1 mM IPTG (isopropyl-β-D-thiogalactopyranoside) for 20 h. The cells were harvested and suspended in lysis buffer containing 50 mM Tris/HCl (pH 7.5), 500 mM NaCl, 1 mM PMSF, 0.05 % (v/v) 2-mercaptoethanol and 5 % (v/v) glycerol. After lysing cells by sonication, the His-tagged XMRV-RNase H domain was purified using nickel-affinity chromatography, and the His-tag was subsequently removed by thrombin protease digestion. The XMRV-RNase H domain was further purified by SP-Sepharose FF ion exchange chromatography and gel filtration chromatography, and then equilibrated with buffer containing 20 mM Tris/HCl (pH 7.0), 0.2 M NaCl, 5 mM $MgCl_2$, 2 mM DTT (dithiothreitol) and 10 % (v/v) glycerol. To prevent protein oxidation, 10 mM DTT was added and the protein was concentrated to 43 mg/ml for crystallization. Wild-type and mutant forms of full-length RT were prepared using similar methods.

Crystallization and data collection

Crystallization was performed at 291 K using the sitting-drop vapour diffusion method. Initial trials were carried out using commercial screening kits (Hampton Research). Optimal crystals were formed by mixing 1.8 μl of protein solution with an equal volume of reservoir solution containing 0.1 M Tris/HCl (pH 7.0), 0.05 M $CdSO_4$ and 0.3 M sodium acetate and growing for 3 days at 277 K. X-ray diffraction data were collected on a Photon Factory Beamline 1A (Tsukuba) equipped with an ADSC quantum detector. For cryoprotection, the crystals were soaked in a reservoir solution containing 20 % (v/v) glycerol. The crystals diffracted to 2.6 Å resolution and belonged to the space group $P4_32_12$, with unit cell parameters of $a = b = 63.70$ Å; $c = 163.92$ Å; and $\alpha = \beta = \gamma = 90°$. The collected diffraction data were processed and scaled with Mosflm [25] and Scala [26]. Data collection statistics and refinements are shown in Table 1.

Structure solution and model refinement

The crystal structure of XMRV-RNase H was determined by the molecular replacement method using the MoMLV structure ([17], PDB code 2HB5; 85 % sequence identity to XMRV RNase H). The program Phaser [27] placed one molecule in the asymmetric

Table 1 Data collection and refinement statistics

Parameter	Value
Data collection	
Wavelength (Å)	1.0
Space group	$P4_32_12$
Cell parameter (Å)	$a = b = 63.70$ Å, $c = 163.92$ Å, $\alpha = \beta = \gamma = 90°$
Resolution (Å)	40–2.6 (2.60–2.74)*
Total reflections	73 798 (10 768)
Unique reflections	11 092 (1580)
Completeness (%)/redundancy	100.0 (100.0)/6.7 (6.8)
R_{merge} (%)†	6.3 (37.3)
$<I>/<\sigma I>$	9.0 (2.0)
Refinement	
Number of reflections	11022
Number of atoms	1219
R_{cryst}/R_{free}	0.239/0.260
Rms deviations from ideal geometry	
Bond distances (Å)	0.008
Bond angles (°)	1.41
Impropers (°)	0.84
Dihedrals (°)	23.5
Average B factors (Å^2)	64.4

*The values in parentheses are for the highest resolution shell.
†$R_{merge} = \Sigma_{hkl} \Sigma_j |I_{hkl,j} - <I_{hkl}>|/ \Sigma_{hkl} \Sigma_j I_{hkl,j}|$, where I is the intensity for the jth measurement of an equivalent reflection with the indices h,k,l.

unit. A randomly selected 5% of data were reserved for the R_{free} calculation. The structure was refined using CNS [28], and iterative manual model building was done using Coot [29]. Several strong densities in the 2Fo–Fc map ($>10\sigma$) were found elsewhere. We modelled these as cadmium ions, because cadmium sulfate was present in the precipitant solution and is indispensable for crystallization. Furthermore, all cadmium sites were engaged in crystal contacts through formation of strong salt bridges with negatively charged residues. A difference Fourier map showed many positive peaks, especially near the active site, that could be modelled as water molecules; however, these were not included in the model because of the medium resolution of the data. The final model includes residues 507–663, one magnesium and four cadmium ions. Figures were drawn using the program Pymol (http://www.pymol.org).

RNase H FRET assay

For the preparation of RNA/DNA hybrid substrates, an 18-nucleotide 3′-fluourescein-labelled RNA was annealed to a complementary 18-nucleotide 5′-dabsyl-labelled DNA. The assay was performed using the FRET protocol described by Parniak et al. [30]. Briefly, the reaction was initiated by adding DNA/RNA substrate and the desired amount of either wild-type or mutant RT of XMRV to reaction buffer containing 50 mM Tris/HCl (pH 8.5), 5 mM MgCl$_2$, 50 mM NaCl and 10 mM DTT. Fluorescence signals were monitored at excitation and emission wavelengths of 485 and 535 nm respectively. All experiments were performed in triplicate, and kinetic constants k_{cat}, K_m and k_{cat}/K_m were determined from a direct fit of the data to the Michaelis–Menten equation.

RESULTS

Overall structure of the XMRV RNase H domain

The crystal structure of the RNase H domain from XMRV RT was solved by the molecular replacement method at a resolution of 2.6 Å. The final R_{cryst} and R_{free} values were 0.239 and 0.260, respectively. The model was validated using MolProbity [31]. A Ramachandran plot showed that 96.1% of all residues were within the favoured regions and only one residue (Glu[35]) was in an outlier region. The all-atom contacts clash score was 25.84 (59th percentile) and the MolProbity score was 2.58 (58th percentile). Overall, the geometry and electron-density map were of good quality.

XMRV RNase H has a compact structure comprising a central β-sheet of five mixed antiparallel and parallel β-strands surrounded by four helices on one face and one helix on the other (Figure 1A). The helices and strands are arranged as an $\alpha\beta\alpha$ Rossmann-like fold. A search for homologous structures using the Dali server [32] identified several members of the RNase H family, including MoMLV RNase H ([17]; PDB code 2HB5; z-score = 23.3), XMRV RNase H ([23]; PDB code 3P1G; z-score = 23.1), and human RNase H1 ([19]; PDB code 2QKB; z-score = 17.1). When we aligned the structure of XMRV RNase H with that of MoMLV RNase H, 134 out of 153 Cα atoms were superimposed with an rms (root-mean-square) deviation of 0.8 Å. The regions that could not be aligned were mainly found among residues 596–610 (basic protrusion; discussed below), residues 635–640 (His-containing loop) and residues near the N-terminus. Except for these regions, the XMRV-RNase H structure was nearly identical with that of MoMLV RNase H. Structural superposition showed that the location of His[638] in XMRV RNase H was distinct from that of HIV-1 or *E. coli* [14,16]. This conserved His residue within the His-containing loop is necessary for proper binding to the nucleic acid template [33]. Since the side chain of His[638] directly interacts with cadmium ion in our model, it is possible that this difference in conformation reflects a crystallization artefact.

Active site

Sequence alignment showed that the residues in the active site of RNase H are highly conserved among species. The active site of RNase H from XMRV contains a highly conserved DEDD motif (Asp[524], Glu[562], Asp[583] and Asp[653]), which co-ordinates one magnesium ion required in RNA hydrolysis (Figures 1B and 1C). Since the binding of two metal ions occurs in the presence of RNA/DNA hybrid substrates [18,19], only one magnesium ion is shown in our structure. As shown in Figure 1(B), the
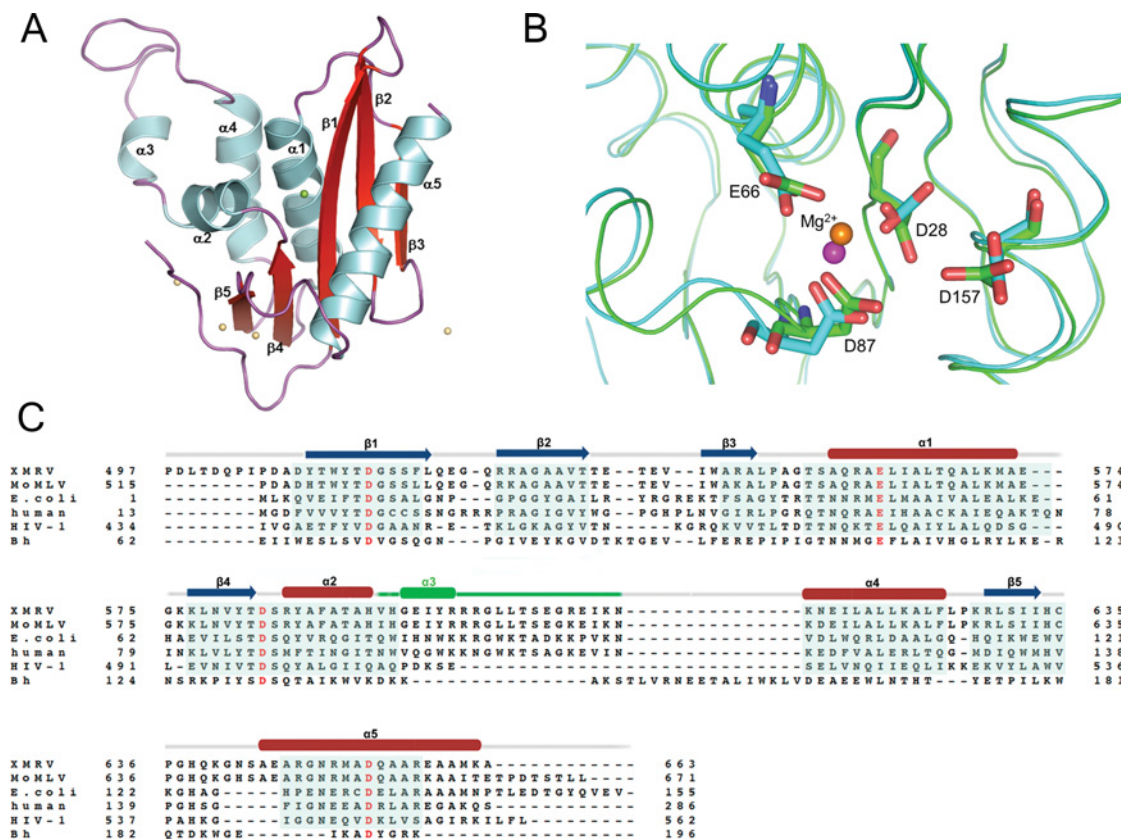
Figure 1   Structure of the RNase H domain from XMRV

(**A**) Ribbon diagram. The boundaries of secondary structural elements are $\beta1$ (519–529), $\beta2$ (534–541), $\beta3$ (547–551), $\alpha1$ (558–573), $\beta4$ (577–582), $\alpha2$ (585–592), $\alpha3$ (595–599), $\alpha4$ (614–625), $\beta5$ (629–633) and $\alpha5$ (644–662). One magnesium ion (green) and four cadmium ions (pale yellow) are represented as spheres. (**B**) Active site. The active site of RNase H of XMRV (green) was superimposed with that of RNase H of MoMLV (cyan). The magnesium ion is represented as a sphere (orange, XMRV; magenta, MoMLV). (**C**) Structure-based sequence alignment. XMRV RNase H sequences are aligned with those of MoMLV, *E. coli*, human, HIV-1 and *Bh*. Sequence alignment is based on structural superposition with XMRV RNase H except for *Bh*; aligned sequences are shaded cyan. The secondary structural elements of XMRV RNase H are indicated above the sequences and those of the basic protrusion are shown in green. The DEDD motif, which is crucial for catalysis, is shown in red.

positions of crucial acidic residues and a magnesium ion are essentially the same as those in MoMLV. In RNase H of XMRV, the first three acidic residues (Asp^{524}, Glu^{562} and Asp^{583}) interact with a magnesium ion. The structures of *Bh* and human RNase H in complex with substrates revealed that the magnesium ion stabilizes transition state intermediates during catalysis [18,19].

Basic protrusion motif

Previous models of RNase H from MoMLV (Lim et al. [17]; PDB code 2HB5) and XMRV ([23]; PDB code 3P1G) were obtained only after deleting the loop containing helix $\alpha3$ (residues 593–603 for MoMLV; residues 595–605 for XMRV). Helix $\alpha3$ (residues 595–599) occurs immediately after helix $\alpha2$, and the following residues adopt a highly bent loop structure formed by several stabilizing hydrogen-bonding contacts (O in Gly^{602}–NZ in Lys^{614}, 3.1 Å; N in Thr^{605}–O in Arg^{609}, 2.9 Å; OG1 in Thr^{605}–N in R609,

3.0 Å) (Figures 2A and 2B). Except for the contact between Gly^{602} and Lys^{604}, helix $\alpha3$ and the following loop have no contact with the remaining part of the RNase H domain. Bound cadmium ions are not found in this region, implying that this region is in a native conformation.

Overall, the conformation of this protrusion in XMRV RNase H is highly similar to that of *E. coli* (Figure 2B). It includes four arginine residues (Arg^{599}, Arg^{600}, Arg^{601} and Arg^{609}) (Figures 1C and 2A), hence the name, basic protrusion. All these residues are exposed to solvent with no contact with other residues. A comparison of the structure of XMRV RNase H with that of the HIV-1 RT RNase H domain and human RNase H, shown with RNA/DNA hybrid substrates, showed a good alignment in both cases (Figure 3). We took advantage of this to identify the crucial residues responsible for binding to the RNA/DNA hybrid. The superposition showed that residues 606–612 of XMRV RNase H have the potential to make contact with the RNA/DNA hybrid
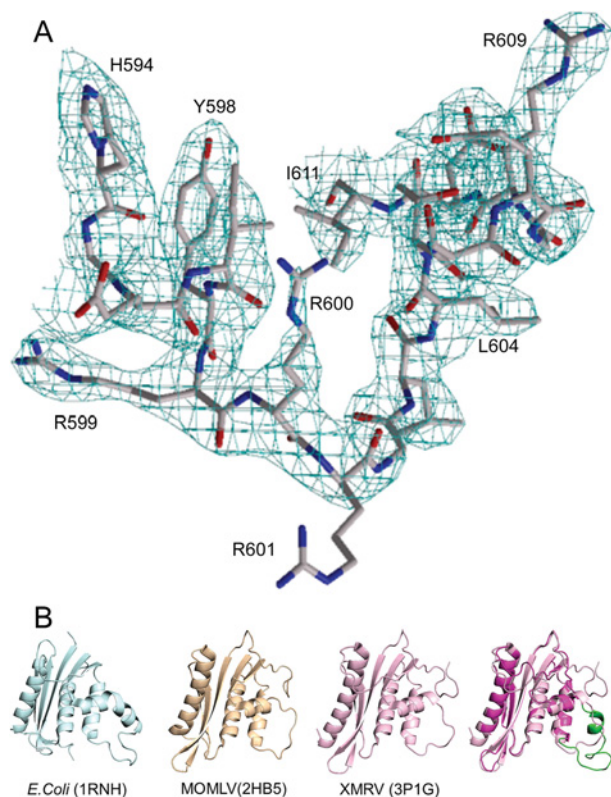
Figure 2 Basic protrusion motif

(**A**) Electron density map. The 2Fo–Fc map of XMRV RNase H was contoured at the 1σ level around the basic protrusion (residues 594–611) and was superposed with the refined model. (**B**) Comparison of our model of the basic protrusion of RNase H (right panel) with that in *E. coli*, MoMLV and XMRV. All ribbon diagrams are shown from the same point of view. Our model (basic protrusion in green and the remainder in magenta) was superposed with that of XMRV with a truncated basic protrusion (pink; PDB code 3P1G).
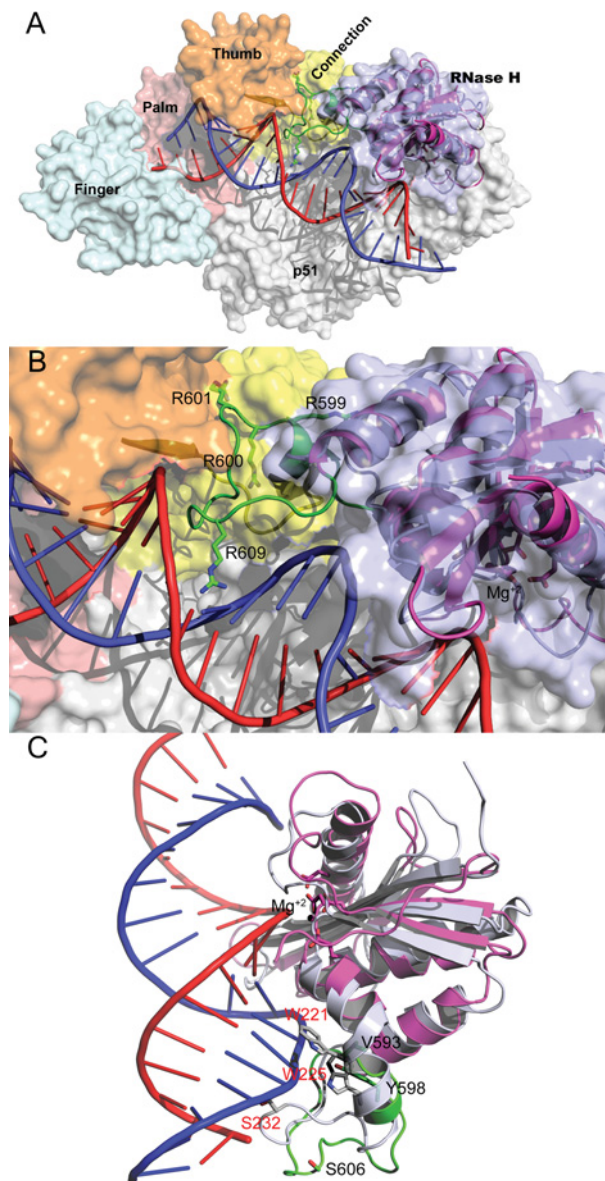

Figure 3 Superposition of XMRV RNase H with complexed forms of HIV-1 RT (A, B) and human RNase H (C)

(**A, B**) The p66/p51 heterodimer of HIV-1 is represented as a surface diagram, and DNA (blue)/RNA (red) template bound to HIV-1 is represented in backbone form. The p51 subunit is shown in light grey, whereas different colours are used to denote distinct domains in the p66 subunit. The primer grip (residues 349–367) in thumb and connection domains corresponding to the basic protrusion in XMRV RNase H is shown as a ribbon. The basic protrusion of RNase H of XMRV (residues 593–612) is shown in green and the remainder is coloured magenta. The DEDD motif and magnesium ion of XMRV RNase H, shown as sticks and a sphere, respectively, indicate the location of the catalytic core. (**C**) Ribbons of human RNase H are coloured grey whereas those of XMRV RNase H are coloured magenta. DNA and RNA strands bound to human RNase H are shown in red and blue, respectively. The residues that participate in binding to the DNA/RNA hybrid are coloured black (XMRV) or red (human).

molecule present in the HIV-1 RT model. It is approx. -6 to -9 base pairs away from the scissile phosphate position (between -1 and $+1$ nucleotides in the RNA strand). Notably, Arg[609] and Lys[612], in particular, make good contact with the substrate (Figure 3B). The side chains of three consecutive arginine residues point towards the connection domain. Therefore it is reasonable to infer that the basic protrusion motif may stabilize the RT structure and binding with an RNA/DNA hybrid, supporting proper enzymatic function. In HIV-1 RT lacking the basic protrusion motif, the residues Gly[359], Ala[360] and His[361] in the connection domain make alternative interactions with the corresponding region and compensate for the absence of a basic protrusion motif.

Conformations of the basic protrusion in XMRV and human RNase Hs are somewhat different (Figure 3C). In the human RNase H structure [19], residues 231–236 dive deep into the groove of the RNA/DNA template, and Ser[232] (that corresponds to Ser[606] of XMRV RNase H) and Thr[233] make direct contact. Compared with the basic protrusion motif of XMRV RNase H, these residues must move ∼5–8 Å into the template in

a twisting manner, suggesting a dramatic conformational change upon substrate binding. Further, hydrophobic interactions also
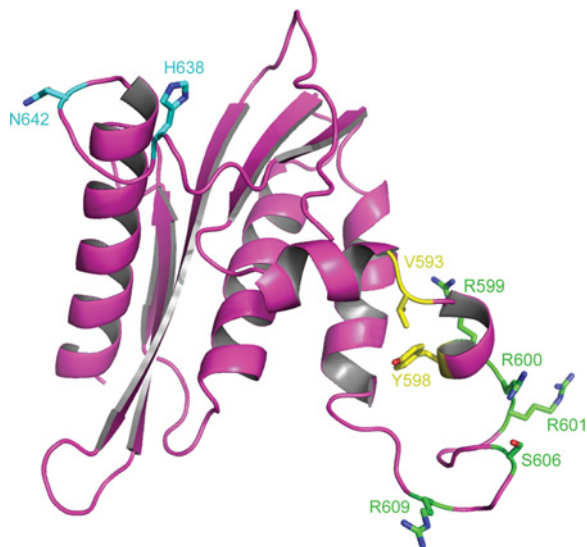
Figure 4 Categorization and location of mutations
Residues at which mutations were introduced are shown as sticks on the ribbon diagram. The mutated residues are coloured differently according to the roles of mutations (His-containing loop, cyan; hydrophilic contact, green; hydrophobic contact, yellow).

appear to be important. The residues Trp^{221} and Trp^{225} of helix $\alpha 3$, which are in good alignment with the corresponding residues (Val^{593} and Tyr^{598}) in XMRV RNase H, are in proximity with the backbone of the RNA/DNA template.

Effects of mutations in RNase H on catalytic activity

In order to confirm the functional importance of residues suggested to interact with RNA/DNA hybrid based on the crystal structure, we individually mutated the identified residues and tested catalytic activity. We created three groups of XMRV RNase H mutants based on the structural information on XMRV RNase H (Figure 4). These groups are designed to show how three motifs, the His-containing loop, hydrophilic contacts in the basic protrusion and hydrophobic contacts in the basic protrusion, affect RT catalytic activity (Figure 4, Table 2). All mutations except for Val^{593} involved alanine substitution to exclude the effects of side chain. Val^{593} was replaced with the charged residue aspartate to test whether hydrophobic contact is important. Initial studies showed that the catalytic activity of the RNase H domain alone is too low to allow the determination of kinetic parameters. Therefore we inserted these mutations into full-length XMRV RT enzymes. All mutant RT enzymes were successfully overexpressed in *E. coli* and their purities were determined to be greater than 90 % by SDS/PAGE (results not shown). Catalytic activity was determined using fluorescein-labelled RNA/DNA duplex as a substrate [30]. Kinetic constants of wild-type and mutant enzymes are summarized in Table 2.

The H638A mutation resulted in a significant decrease in k_{cat}/K_m, whereas the N642A mutation resulted in an increase in k_{cat}/K_m, explaining how the conformation observed in this structure is influenced by crystallization and demonstrating that His^{638} in the His-binding loop is responsible for template binding.

The basic protrusion of XMRV RNase H contains five positively charged residues, three of which are consecutive arginine residues (Arg^{599}–Arg^{600}–Arg^{601}). To test whether these positively charged residues affect catalytic activity, we individually mutated each residue to alanine and tested the effects on kinetic parameters. Replacement of either Arg^{600} or Arg^{601} with alanine resulted in a marked increase in K_m value, suggesting that mutations of these residues weakened the binding affinity of the enzyme for substrate. In contrast, substitution of alanine for Arg^{599} had little or no effect. Collectively, these results suggest that Arg^{600} and Arg^{601} are important for substrate binding and enzymatic activity, whereas Arg^{599} is not. These results are in good agreement with the superposition of RNase H of XMRV with the RT of HIV-1 (Figure 3B), in which the side chain of Arg^{599} is directed towards helix $\alpha 4$ of XMRV RNase H, whereas those of Arg^{600} and Arg^{601} are directed towards thumb and connection domains of the RT. Although the superposition of RNase H of XMRV with HIV-1 RT showed good potential contact between the side chain of Arg^{609} and the phosphate backbone of the RNA template, the R609A mutation had no influence on catalysis. In contrast, mutation of S606 to alanine completely abolished catalytic activity. The side chain of Ser^{606} has close contact with the DNA backbone and appears to be important for binding substrate in the crystal structure of human RNase H [19]. These results demonstrate that XMRV RNase H and human RNase H share a common recognition mechanism and further indicate that the basic protrusion of both undergoes a dramatic conformational change upon substrate binding.

In human RNase H complexed with an RNA/DNA hybrid, Ser^{233}, Trp^{221} and Trp^{225} form a channel and are involved in the interaction with the DNA substrate [19]. Sequence and structural alignments show that these residues correspond to Ser^{606}, Val^{593} and Tyr^{598}, respectively, in XMRV RNase H. A mutational analysis of hydrophobic residues on helix $\alpha 3$ in the basic protrusion motif showed that substitution of alanine for Y598A or replacement of Val^{593} with aspartate (altered charge) resulted in complete abolition of catalytic activity. This suggests that the hydrophobic pocket formed by Tyr^{598} and Val^{593} is also crucial for substrate binding. These findings are also in agreement with the information obtained from the structure of human RNase H and RNA/DNA hybrid in complex [19].

DISCUSSION

The structural basis of gamma-retroviral RT function has not been fully elucidated. Although the structure of MoMLV RT has been determined [34], this structural information is largely confined to fingers, palm and thumb domains; thus, it is difficult to envision the overall architecture of XMRV RT at the moment. Considering the approx. 37 % sequence homology between

Table 2 Activity assays of full-length wild-type and mutant RNase Hs from XMRV

Location of mutation	Mutation	K_m (nM)	K_{cat} (min^{-1})	K_{cat}/K_m (min$^{-1}\cdot\mu$M^{-1})
—	Wild-type	10.4 ± 0.5	0.12 ± 0.01	11.41 ± 0.03
Histidine-containing loop	H638A	72.2 ± 28.23	0.20 ± 0.02	3.92 ± 1.63
	N642A	20.7 ± 4.4	0.50 ± 0.12	24.06 ± 0.68
Basic protrusion	R599A	22.0 ± 5.2	0.36 ± 0.04	18.11 ± 4.89
	R600A	50.1 ± 7.4	0.04 ± 0.01	0.80 ± 0.19
	R601A	177.6 ± 26.7	0.06 ± 0.01	0.38 ± 0.08
	R599A,R600A	36.8 ± 10.5	0.28 ± 0.12	7.07 ± 1.11
	R600A,R601A	111.7 ± 44.9	0.68 ± 0.59	0.86 ± 0.36
	R599A,R600A,R601A	64.5 ± 26.9	0.04 ± 0.03	0.57 ± 0.30
	R609A	19.9 ± 4.2	0.31 ± 0.04	16.45 ± 3.28
	S606A		N/A	
Hydrophobic pocket	V593D		N/A	
	Y598A	84.5 ± 15.7	0.06 ± 0.01	0.65 ± 0.31

HIV-1 and XMRV RTs, the only valid option is to make speculations based on the superposition of HIV-1 and XMRV RTs. Employing this superposition approach using both human RNase H and HIV-1 RT, we obtained important insights into the roles of the basic protrusion motif. First, the basic protrusion is very mobile and moves $\sim$6–8 Å towards the major groove upon template binding (Figure 3C). Mutational studies confirmed that the distant Ser[606] participates in binding, whereas the presumably adjacent Arg[609] does not. The way in which the basic protrusion motif moves might be a combination of translation and rotation during substrate recognition. The basic protrusion also stabilizes connection and palm domains within the RT protein. Although Arg[600] and Arg[601] did not participate in contacting the RNA/DNA template, mutations of these residues resulted in a marked decrease in catalytic activity. The superposition with HIV-1 RT suggests that these residues point towards and bind with the connection and thumb domains and thereby help establish the proper architecture of the entire RT enzyme. The thumb domain in HIV-1 RT is known to be highly mobile, depending on its presence as an apo form, a holo form in complex with a DNA duplex or a holo form in complex with a polymerase-directed inhibitor [14,35,36]. Thus, the basic protrusion may play an additional role in polymerase activity through interaction with the thumb domain. This observation is also supported by the fact that the presence of the basic protrusion motif not only improves catalytic activity and specificity of RNase H *in vitro*, but also enhances retroviral replication *in vivo* [21,22] Unlike heterodimeric HIV-1 RT, both XMRV and MoMLV RTs are monomeric enzymes approx. 110 residues longer than HIV-1 RT [11]. Except for an approx. 40-residue extension at the N-terminus that appears to be dispensable for proper function [37], the distinct differences lie in the approx. 30-residue insertion between connection and RNase H domains, and the presence of a basic protrusion motif within RNase H. Although these regions are completely disordered in the crystal structure of MoMLV RT, a linker-scanning analysis of the region between connection and RNase H domains (residues 475–502)

demonstrated that this region is essential for retroviral replication [38]. In HIV-1, the p51 domain rarely contacts the RNA/DNA template and plays a role in the stabilization and proper positioning of the p66 domain relative to the RNA/DNA hybrid. These two additional motifs in monomeric RT (approx. 30-residue insertion and basic protrusion) might replace the role of the p51 domain in heterodimeric RT. However, additional structural and biological studies are required to support this speculation.

In summary, we describe the structure of the entire gamma-retroviral RNase H domain that includes the basic protrusion motif. The structural investigations and mutational analyses allowed us to gain insights into the roles of individual residues. The basic protrusion motif is folded into a short helix and an immediately adjacent highly bent loop that exhibits little interaction with the body of the catalytic domain. Substrate recognition by the basic protrusion motif involves both shape complementarity and specific interactions between residues within the motif. The structure described here should aid in understanding the detailed roles of the RNase H domain and the entire RT of gamma-retroviruses.

REFERENCES

1 Urisman, A., Molinaro, R. J., Fischer, N., Plummer, S. J., Casey, G., Klein, E. A., Malathi, K., Magi-Galluzzi, C., Tubbs, R. R., Ganem, D. et al. (2006) Identification of a novel Gammaretrovirus in prostate tumors of patients homozygous for R462Q RNASEL variant. PLoS Pathog. **2**, e25

2 Schlaberg, R., Choe, D. J., Brown, K. R., Thaker, H. M. and Singh, I. R. (2009) XMRV is present in malignant prostatic epithelium and is associated with prostate cancer, especially high-grade tumors. Proc. Natl. Acad. Sci. U.S.A. **106**, 16351–16356

3 Arnold, R. S., Makarova, N. V., Osunkoya, A. O., Suppiah, S., Scott, T. A., Johnson, N. A., Bhosle, S. M., Liotta, D., Hunter, E., Marshall, F. F. et al. (2010) XMRV infection in patients with prostate cancer: novel serologic assay and correlation with PCR and FISH. Urology **75**, 755–761

4 Fischer, N., Schulz, C., Stieler, K., Hohn, O., Lange, C., Drosten, C. and Aepfelbacher, M. (2010) Xenotropic murine leukemia virus-related gammaretrovirus in respiratory tract. Emerg. Infect. Dis. **16**, 1000–1002

5 Kim, S., Kim, N., Dong, B., Boren, D., Lee, S. A., Das Gupta, J., Gaughan, C., Klein, E. A., Lee, C., Silverman, R. H. and Chow, S. A. (2008) Integration site preference of xenotropic murine leukemia virus-related virus: a new human retrovirus associated with prostate cancer. J. Virol. **82**, 9964–9977

6 Bogerd, H. P., Zhang, F., Bieniasz, P. D. and Cullen, B. R. (2011) Human APOBEC3 proteins can inhibit xenotropic murine leukemia virus-related virus infectivity. Virology **410**, 234–239

7 Chaipan, C., Dilley, K. A., Paprotka, T., Delviks-Frankenberry, K. A., Venkatachari, N. J., Hu, W. S. and Pathak, V. K. (2011) Severe restriction of xenotropic murine leukemia virus-related virus replication and spread in cultured human peripheral blood mononuclear cells. J. Virol. **85**, 4888–4897

8 Groom, H. C., Yap, M. W., Galao, R. P., Neil, S. J. and Bishop, K. N. (2010) Susceptibility of xenotropic murine leukemia virus-related virus (XMRV) to retroviral restriction factors. Proc. Natl. Acad. Sci. U.S.A. **107**, 5166–5171

9 Paprotka, T., Delviks-Frankenberry, K. A., Cingoz, O., Martinez, A., Kung, H. J., Tepper, C. G., Hu, W. S., Fivash, M.J., Jr., Coffin, J. M. and Pathak, V. K. (2011) Recombinant origin of the retrovirus XMRV. Science **333**, 97–101

10 Stieler, K. and Fischer, N. (2010) Apobec 3G efficiently reduces infectivity of the human exogenous gammaretrovirus XMRV. PLoS One **5**, e11738

11 Cote, M. L. and Roth, M. J. (2008) Murine leukemia virus reverse transcriptase: structural comparison with HIV-1 reverse transcriptase. Virus Res. **134**, 186–202

12 Julias, J. G., McWilliams, M. J., Sarafianos, S. G., Arnold, E. and Hughes, S. H. (2002) Mutations in the RNase H domain of HIV-1 reverse transcriptase affect the initiation of DNA synthesis and the specificity of RNase H cleavage *in vivo*. Proc. Natl. Acad. Sci. U.S.A. **99**, 9515–9520

13 Klumpp, K. and Mirzadegan, T. (2006) Recent progress in the design of small molecule inhibitors of HIV RNase H. Cur. Pharm. Des. **12**, 1909–1922

14 Huang, H., Chopra, R., Verdine, G. L. and Harrison, S. C. (1998) Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance. Science **282**, 1669–1675

15 Ishikawa, K., Okumura, M., Katayanagi, K., Kimura, S., Kanaya, S., Nakamura, H. and Morikawa, K. (1993) Crystal structure of ribonuclease H from *Thermus thermophilus* HB8 refined at 2.8 Å resolution. J. Mol. Biol. **230**, 529–542

16 Katayanagi, K., Miyagawa, M., Matsushima, M., Ishikawa, M., Kanaya, S., Ikehara, M., Matsuzaki, T. and Morikawa, K. (1990) Three-dimensional structure of ribonuclease H from *E. coli*. Nature **347**, 306–309

17 Lim, D., Gregorio, G. G., Bingman, C., Martinez-Hackert, E., Hendrickson, W. A. and Goff, S. P. (2006) Crystal structure of the moloney murine leukemia virus RNase H domain. J. Virol. **80**, 8379–8389

18 Nowotny, M., Gaidamakov, S. A., Crouch, R. J. and Yang, W. (2005) Crystal structures of RNase H bound to an RNA/DNA hybrid: substrate specificity and metal-dependent catalysis. Cell **121**, 1005–1016

19 Nowotny, M., Gaidamakov, S. A., Ghirlando, R., Cerritelli, S. M., Crouch, R. J. and Yang, W. (2007) Structure of human RNase H1 complexed with an RNA/DNA hybrid: insight into HIV reverse transcription. Mol. Cell **28**, 264–276

20 Schultz, S. J. and Champoux, J. J. (2008) RNase H activity: structure, specificity, and function in reverse transcription. Virus Res. **134**, 86–103

21 Boyer, P. L., Gao, H. Q., Frank, P., Clark, P. K. and Hughes, S. H. (2001) The basic loop of the RNase H domain of MLV RT is important both for RNase H and for polymerase activity. Virology **282**, 206–213

22 Lim, D., Orlova, M. and Goff, S. P. (2002) Mutations of the RNase H C helix of the Moloney murine leukemia virus reverse transcriptase reveal defects in polypurine tract recognition. J. Virol. **76**, 8360–8373

23 Kirby, K. A., Marchand, B., Ong, Y. T., Ndongwe, T. P., Hachiya, A., Michailidis, E., Leslie, M. D., Sietsema, D. V., Fetterly, T. L., Dorst, C. A. et al. (2012) Structural and inhibition studies of the RNase H function of xenotropic murine leukemia virus-related virus reverse transcriptase. Antimicrob. Agents Chemother. **56**, 2048–2061

24 Zhou, D., Chung, S., Miller, M., Grice, S. F. and Wlodawer, A. (2012) Crystal structures of the reverse transcriptase-associated ribonuclease H domain of xenotropic murine leukemia-virus related virus. J. Struct. Biol. **177**, 638–645

25 Leslie, A. G. (1999) Integration of macromolecular diffraction data. Acta Crystallogr. D Biol. Crystallogr. **55**, 1696–1702

26 Collaborative Computational Project, Number 4 (1994) The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D Biol. Crystallogr. **50**, 760–763

27 McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. and Read, R. J. (2005) Likelihood-enhanced fast translation functions. Acta Crystallogr. D Biol. Crystallogr. **61**, 458–464

28 Brunger, 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. et al. (1998) Crystallography and NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr. D Biol. Crystallogr. **54**, 905–921

29 Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. **60**, 2126–2132

30 Parniak, M. A., Min, K. L., Budihas, S. R., Le Grice, S. F. and Beutler, J. A. (2003) A fluorescence-based high-throughput screening assay for inhibitors of human immunodeficiency virus-1 reverse transcriptase-associated ribonuclease H activity. Anal. Biochem. **322**, 33–39

31 Chen, V. B., Arendall, W.B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. and Richardson, D. C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D Biol. Crystallogr. **66**, 12–21

32 Holm, L. and Sander, C. (1993) Protein structure comparison by alignment of distance matrices. J. Mol. Biol. **233**, 123–138

33 Kanaya, S., Katsuda-Nakai, C. and Ikehara, M. (1991) Importance of the positive charge cluster in *Escherichia coli* ribonuclease HI for the effective binding of the substrate. J. Biol. Chem. **266**, 11621–11627

34 Das, D. and Georgiadis, M. M. (2004) The crystal structure of the monomeric reverse transcriptase from Moloney murine leukemia virus. Structure **12**, 819–829

35 Esnouf, R., Ren, J., Ross, C., Jones, Y., Stammers, D. and Stuart, D. (1995) Mechanism of inhibition of HIV-1 reverse transcriptase by non-nucleoside inhibitors. Nature Struct. Biol. **2**, 303–308

36 Rodgers, D. W., Gamblin, S. J., Harris, B. A., Ray, S., Culp, J. S., Hellmig, B., Woolf, D. J., Debouck, C. and Harrison, S. C. (1995) The structure of unliganded reverse transcriptase from the human immunodeficiency virus type 1. Proc. Natl. Acad. Sci. U.S.A. **92**, 1222–1226

37 Das, D. and Georgiadis, M. M. (2001) A directed approach to improving the solubility of Moloney murine leukemia virus reverse transcriptase. Protein Sci. **10**, 1936–1941

38 Puglia, J., Wang, T., Smith-Snyder, C., Cote, M., Scher, M., Pelletier, J. N., John, S., Jonsson, C. B. and Roth, M. J. (2006) Revealing domain structure through linker-scanning analysis of the murine leukemia virus (MuLV) RNase H and MuLV and human immunodeficiency virus type 1 integrase proteins. J. Virol. **80**, 9497–9510