

Foamy Virus RNase H  
Aktivität, Struktur und Funktion

Dissertation

Zur Erlangung des Doktorgrades der Graduiertenschule für Mathematik und  
Naturwissenschaften (BayNAT) der Universität Bayreuth

vorgelegt von

Diplom-Biochemikerin

Berit Leo

Bayreuth, 2013



Die vorliegende Arbeit wurde in der Zeit von August 2009 bis März 2013 an der Universität Bayreuth am Lehrstuhl Biopolymere unter der Betreuung von Frau Prof. Dr. Birgitta M. Wöhrl angefertigt.

Vollständiger Abdruck der von der Bayreuther Graduiertenschule für Mathematik und Naturwissenschaften (BayNAT) der Universität Bayreuth genehmigten Dissertation zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften (Dr. rer. nat.).

Dissertation eingereicht am: 05.03.2013

Zulassung durch das Leitungsgremium: 11.03.2013

Wissenschaftliches Kolloquium: 24.07.2013

Amtierender Direktor: Prof. Dr. Franz X. Schmid

Prüfungsausschuss:

Prof. Dr. Birgitta M. Wöhrl (Erstgutachterin)

Prof. Dr. Wulf Blankenfeldt (Zweitgutachter)

Prof. Dr. Klaus Ersfeld (Vorsitz)

Prof. Dr. G. Matthias Ullmann



## Inhaltsverzeichnis

|   |     |
|---|-----|
| Inhaltsverzeichnis.....   | I   |
| Zusammenfassung.....  | III |
| Summary.....  | V   |
| 1 Einleitung.....   | 1   |
| 1.1 Historische Daten zur Entdeckung einzelner Viren.....   | 1   |
| 1.2 Taxonomie der Retroviren.....   | 2   |
| 1.3 Genomorganisation und Aufbau des Viruspartikels in FV.....  | 3   |
| 1.4 Der Replikationszyklus.....   | 6   |
| 1.5 Die reverse Transkription.....  | 7   |
| 1.6 Die Reverse Transkriptase.....  | 9   |
| 1.6.1 Die Polymerase-Domäne der Reversen Transkriptase.....   | 10  |
| 1.6.2 Die RNase H-Domäne der Reversen Transkriptase.....  | 12  |
| 1.7 Die RNase H in den verschiedenen Reichen der Lebewesen.....   | 16  |
| 1.8 Aktivität und Struktur isolierter RNase H-Domänen des Typs 1.....                                     | 18  |
| 2 Ziele der Arbeit.....   | 21  |
| 3 Zusammenfassung und Diskussion der Ergebnisse.....  | 23  |
| 3.1 Reinigung und Untersuchung der Aktivität der isolierten PFV RNase H-Domäne... 23                      | 23  |
| 3.2 Struktur der isolierten PFV RNase H-Domäne.....   | 25  |
| 3.3 Aktivität der isolierten PFV RNase H-Domäne als D599N-H724N Variante.....                             | 28  |
| 3.4 Substratbindungsstudien der isolierten PFV RNase H-(D599N-H724N) mit Hilfe von NMR-Spektroskopie..... | 30  |
| 3.5 Die Domäneneinteilung der FV PR-RT.....   | 34  |
| 4 Abkürzungsverzeichnis.....  | 39  |
| 5 Literaturverzeichnis.....   | 41  |
| 6 Publikationsliste.....  | 51  |
| 7 Einzelarbeit A.....   | 53  |
| 8 Einzelarbeit B.....   | 67  |
| 9 Einzelarbeit C.....   | 81  |
| 10 Danksagung.....  | 109 |
| 11 Erklärung.....   | 111 |



## Zusammenfassung

Das für die Replikation des RNA-Genoms von Foamy Viren (FV) notwendige Enzym, die Protease-Reverse Transkriptase (PR-RT), beinhaltet die Protease-, die Polymerase- und RNase H-Domäne. Letztere ist für den Abbau der RNA im entstehenden RNA/DNA-Hybrid verantwortlich. Während die FV PR-RT als Monomer vorliegt, besteht die HIV-1 RT aus einem p66/p51-Heterodimer. Erstaunlicherweise ist die isolierte HIV-1 RNase H im Vergleich zur z.B. *E. coli* oder separaten MoMLV RNase H nicht aktiv. Aus den Sequenzvergleichen verschiedener RNase H-Domänen ergibt sich, dass die Prototyp FV (PFV) RNase H im Gegensatz zur HIV-1 RNase H einen Sekundärstrukturbereich aufweist, bei dem es sich um die sogenannte C-Helix mit einer sich anschließenden basischen Schleife (*basic protrusion*) handelt. Da zu Beginn der Arbeit keine 3D-Struktur einer retroviralen RNase H mit *basic protrusion* bekannt war, sollte in dieser Arbeit die Struktur der PFV RNase H und die Funktion der *basic protrusion* bei der Substratbindung geklärt werden.

Die Tertiärstruktur der PFV RNase H-Domäne konnte mit NMR-Spektroskopie gelöst werden. Somit war es möglich, die *basic protrusion* einschließlich der C-Helix zu identifizieren. Die isolierte RNase H-Domäne zeigte in fluoreszenzbasierten Tests sowie in qualitativen RNase H-Versuchen mit radioaktiv markiertem Substrat Aktivität. Um die Funktion der C-Helix und der sich anschließenden basischen Schleife bei der Substratbindung zu analysieren, wurden NMR-Titrationsexperimente durchgeführt. Dafür wurde die PFV RNase H-Domäne zunächst durch den Austausch der zwei katalytisch wichtigen Reste Aspartat 599 und Histidin 724 zu Asparagin inaktiviert (RNase H-(D599N-H724N)), um den Abbau des Substrates während der Messungen zu vermeiden. Die Auswertung von [<sup>15</sup>N, <sup>1</sup>H]-HSQC- und [<sup>1</sup>H, <sup>15</sup>N, <sup>1</sup>H]-NOESY-HSQC-Spektren erbrachte eine Übereinstimmung der Tertiärstrukturen der RNase H-(D599N-H724N) mit der wt RNase H.

Die NMR-Titrationsexperimente zeigten, dass die C-Helix in der PFV RNase H wie ein Lineal agiert, das die sich anschließende basische Schleife zum Substrat orientiert. Darüber hinaus besitzt die *basic protrusion* zusätzlich eine Reihe an positiv geladenen Resten, die gut lösungsmittelzugänglich sind und dadurch erste Kontakte mit dem Substrat ermöglichen. Damit bietet die *basic protrusion* eine Art Plattform für die Substratbindung. Der HIV-1 RNase H fehlt nicht nur die C-Helix, zusätzlich ist die sich anschließende Schleife vermutlich zu kurz, um das Substrat zu binden. Da diese Schleife außerdem nur über einen basischen Rest verfügt, ist wahrscheinlich auch die Gesamtaffinität dieses Bereichs für die Substratbindung zu gering. Strukturvergleiche der PFV RNase H mit der HIV-1 RT zeigen,

dass die fehlende *basic protrusion* der HIV-1 RNase H durch eine Schleife aus der Verbindungs-Subdomäne der p66-Untereinheit kompensiert wird. Dieser Bereich könnte somit ein neuer Angriffspunkt für Inhibitoren in der antiretroviralen Therapie bei HIV-1 sein.

Um zukünftig weitere Strukturanalysen mit der PR-RT bzw. mit einzelnen Domänen durchführen zu können, wurden in einem weiteren Projekt verschiedene N- und C-terminale Deletionsvarianten der PR-RT des Simian Foamy Virus hergestellt. Durch Aktivitätstests mit diesen Deletionsvarianten konnten die Abgrenzungen für die PR-, die Polymerase-, die RNase H-Domäne und die Verbindungs-Subdomäne in der PR-RT identifiziert werden. Dabei zeigte sich, dass die Region H107-N143 C-terminal von der PR wichtig für die Funktion der Polymerase ist. Die Deletion der RNase H-Domäne und Verbindungs-Subdomäne führte zu einer drastischen Abnahme der Substrataffinität, Integrität und Polymerisationsfähigkeit des Enzyms. Trotzdem konnte eine minimale Polymerase-Domäne bestimmt werden (RT(107-454)), die ohne PR- und RNase H-Domäne sowie ohne die Verbindungs-Subdomäne in der Lage ist, zu polymerisieren. Für die Dimerisierung und damit Aktivierung der PR, die durch Bindung von zwei PR-RTs an das sog. PARM-Element (engl. *protease activating RNA motif*) auf der genomischen RNA geschieht, sind jedoch die RNase H-Domäne und die Verbindungs-Subdomäne unverzichtbar. Dadurch ist die RNase H nicht nur für die reverse Transkription essenziell; in FV stellt sie auch ein Regulationselement für die PR-Aktivierung dar und ist somit indirekt an der Prozessierung von Gag (Strukturproteine) und Pol (virale Enzyme) beteiligt.



## Summary

The foamy virus (FV) protease-reverse transcriptase (PR-RT), which is necessary for the replication of the viral RNA genome, consists of a protease, polymerase and RNase H domain. The latter is responsible for the degradation of the RNA in the emerging RNA/DNA-hybrid. While FV PR-RT is monomeric, HIV-1 RT consists of a p66/p51 heterodimer. Surprisingly, in comparison e.g. to the *E. coli* or the separate MoMLV RNase H, the isolated HIV-1 RNase H is not active. Sequence alignments of different RNase H domains show that in contrast to the HIV-1 RNase H the Prototype FV (PFV) RNase H possesses a secondary structure element that comprises the so-called C-helix followed by a basic loop (basic protrusion). At the beginning of my PhD thesis no 3D structure of a retroviral RNase H with a basic protrusion was available. Thus, the aim of my work was to resolve the 3D structure of the PFV RNase H and the function of the basic protrusion in substrate binding.

The tertiary structure of the PFV RNase H could be solved by NMR spectroscopy and the basic protrusion including the C-helix could be identified. The isolated RNase H domain proved active by fluorescence based tests and qualitative RNase H assays with radioactively labelled substrate. To elucidate the function of the C-helix and the adjacent basic loop in substrate binding NMR titration experiments were performed. The catalytically important residues aspartate 599 and histidine 724 were exchanged for asparagines (RNase H-(D599N-H724N)), to inactivate the RNase H and thus avoid substrate degradation during NMR measurements. Analysis of [<sup>15</sup>N, <sup>1</sup>H]-HSQC- and [<sup>1</sup>H, <sup>15</sup>N, <sup>1</sup>H]-NOESY-HSQC-spectra showed that the tertiary structure of the wt RNase H and of the RNase H-(D599N-H724N) were comparable.

NMR titration experiments with the mutant PFV RNase H showed that the C-helix acts as a ruler to position the basic loop to the substrate. In addition, the basic protrusion possesses many positively charged residues which are well exposed to the solvent, thereby enabling first non-specific contacts with the substrate. Thus, the basic protrusion provides a platform for substrate binding. In contrast, HIV-1 RNase H lacks the C-helix and, in addition, the adjacent loop is probably too short to bind the substrate. Furthermore, since this loop harbours only one basic residue, the overall affinity to the substrate is expected to be rather low. However, structural comparison of PFV RNase H with HIV-1 RT showed that the lack of the basic protrusion in HIV-1 RNase H is compensated by a loop derived from the connection subdomain of the p66 subunit. Therefore, this region could be a new target for inhibitors against HIV-1.

To further analyze the structure of PR-RT or of PR-RT domains in future projects, N- and C-terminal deletion variants of the PR-RT from Simian Foamy Virus have been constructed. Activity measurements of the deletion variants were used to identify the domain borders of the PR, polymerase and RNase H domains as well as the connection subdomain. The region H107-N143 proved important for the functionality of the polymerase. Deletion of the RNase H domain together with the connection subdomain resulted in a drastic decrease in substrate affinity, integrity and polymerization activity of the enzyme. It was possible to determine a minimal polymerase domain (RT(107-454)) lacking the PR and RNase H domains and the connection subdomain, which is still able to polymerize. Furthermore, we found that PR dimerization and activation *via* binding of at least two PR-RT molecules to the viral *protease activating RNA motif* (PARM) requires the presence of the full length PR-RT including the RNase H domain and the connection subdomain. Consequently, the RNase H is not only essential for reverse transcription. In FV it is also an important tool for regulating PR activation. Thus the RNase H is indirectly involved in the processing of the Gag (structural proteins) and Pol (viral enzymes) precursors.

# 1 Einleitung

## 1.1 Historische Daten zur Entdeckung einzelner Viren

Bis zum Ende des 19. Jahrhunderts wurden für heute bekannte Viruserkrankungen „Gifte“ bzw. ein „filtrierbares Agens“ als Ursache verantwortlich gemacht, da es mit den damaligen Methoden nicht möglich war, Organismen, die kleiner als Bakterien sind, nachzuweisen (Ivanofsky, 1892; Beijerinck, 1898; Mayer, 1882). Der Begriff „Virus“ (lateinisch für „Gift“ oder „Schleim“) setzte sich um 1900 anstelle des Ausdruckes „infektiöses Agens“ durch und wurde erstmals für das Tabakmosaikvirus verwendet. 1898 entdeckten Friedrich Loeffler und Paul Frosch das erste tierpathogene Virus, das der Maul- und Klauenseuche (Loeffler & Frosch, 1898). Die ersten Retroviren wurden 1908 von Vilhelm Ellerman und Oluf Bang sowie 1911 von Peyton Rous in Filtraten von Geflügelsarkomen identifiziert (Ellermann & Bang 1908; Rous, 1911). Im ersten Fall handelte es sich um das Aviäre Leukosevirus (engl. *Avian Leukosis Virus*, ALV), das zweite gefundene Retrovirus wurde nach seinem Entdecker Rous-Sarkom-Virus (engl. *Rous Sarcoma Virus*, RSV) benannt.

Seitdem wurden viele weitere Retroviren in Vögeln und in Säugetieren entdeckt. Prominente humanpathogene Vertreter sind die humanen Immundefizienzviren 1 und 2 (engl. *Human Immunodeficiency Virus type 1 and type 2*, HIV-1 und HIV-2) sowie die humanen T-Zell-lymphotropen Viren 1 und 2 (engl. *Human T-cell lymphotropic Virus 1 and 2*, HTLV-1 und HTLV-2). Die Existenz von Spumaretroviren, auch Foamy Viren (FV) genannt, wurde erstmals 1954 bzw. 1955 in Zellkulturen aus dem Nierengewebe von Makaken (Enders & Peebles, 1954; Rustigian *et al.*, 1955) und 1971 im Nasopharynx-Karzinom-Gewebe eines kenianischen Patienten (*Human Foamy Virus*, HFV) (Achong *et al.*, 1971) nachgewiesen. Da das 1971 isolierte HFV eine sehr hohe Sequenzhomologie zu anderen FV aus Schimpansen aufwies (Herchenröder *et al.*, 1994), wurde vermutet, dass es sich bei HFV um eine Interspezies-Transmission eines Affenvirus handelt. Aus diesem Grund wurde der Virusstamm HFV später in Prototyp FV (engl. *Prototype Foamy Virus*, PFV) umbenannt (Rinke *et al.*, 2002; Rethwilm, 2003).

## 1.2 Taxonomie der Retroviren

Derzeit werden Retroviren in zwei Unterfamilien und sieben Genera eingeteilt (Rethwilm, 2003) (Tabelle 1). Das markanteste Unterscheidungsmerkmal zwischen beiden Unterfamilien ist, dass Orthoretroviren einzelsträngige (engl. *single stranded*, ss) RNA als Genom im Viruspartikel besitzen, während man in Spumaretroviren doppelsträngige (engl. *double stranded*, ds) DNA als infektiöses Material in Viruspartikeln vorfindet (Yu *et al.*, 1999; Moebes *et al.*, 1997).

**Tabelle 1: Klassifizierung der Retroviren** (geändert nach (Rethwilm, 2003))

| Unterfamilie             | Genus                  | Beispiele   | Genom   |
|--------------------------|------------------------|---|---------|
| <i>Orthoretrovirinae</i> | $\alpha$ -Retrovirus   | Aviäres Leukosevirus ( <i>Avian Leukosis Virus</i> , ALV)<br>Rous-Sarkom-Virus ( <i>Rous Sarcoma Virus</i> , RSV)   | einfach |
|                          | $\beta$ -Retrovirus    | Maus-Mammatumovirus ( <i>Mouse Mammary Tumor Virus</i> , MMTV)<br>Mason-Pfizer-Affenvirus ( <i>Mason-Pfizer Monkey Virus</i> , MPMV)                                  | einfach |
|                          | $\gamma$ -Retrovirus   | Mausleukämievirus ( <i>Murine Leukemia Virus</i> , MLV)<br>Felines Leukämievirus ( <i>Feline Leukemia Virus</i> , FeLV)   | einfach |
|                          | $\delta$ -Retrovirus   | Humanes T-Zell-lymphotropes Virus ( <i>Human T-cell lymphotropic Virus</i> , HTLV)<br>Bovines Leukämievirus ( <i>Bovine Leukemia Virus</i> , BLV)                     | komplex |
|                          | $\epsilon$ -Retrovirus | Walleye-Dermal-Sarkom-Virus ( <i>Walleye Dermal Sarcoma Virus</i> , WDSV)<br>Walleye-Epidermal-Hyperplasie-Virus ( <i>Walleye Epidermal Hyperplasia Virus</i> , WEHV) | komplex |
|                          | Lentivirus             | Humanes Immundefizienzvirus ( <i>Human Immunodeficiency Virus</i> , HIV)<br>Simian-Immundefizienzvirus ( <i>Simian Immunodeficiency Virus</i> , SIV)                  | komplex |
| <i>Spumaretrovirinae</i> | Foamy Virus            | Prototypes Foamy Virus ( <i>Prototype Foamy Virus</i> , PFV)<br>Simian Foamy Virus ( <i>Simian Foamy Virus</i> , SFV)   | komplex |

einfach: Genom kodiert keine Leserahmen für akzessorische Proteine

komplex: Genom kodiert Leserahmen für akzessorische Proteine

Alle Orthoretroviren sind mit Krankheiten verbunden, die Tumorbildung und Immunschwäche hervorrufen, wobei einige bekannte Vertreter in Tabelle 1 aufgelistet sind. FV sind bei vielen Primatenarten, Wild- und Hauskatzen sowie Rindern und Pferden weit verbreitet (Rethwilm, 2003; Meiering & Linial, 2001). Ist der Wirtsorganismus infiziert, persistieren sie als ins Wirtsgenom integriertes Retrovirus (Provirus) bis zum Tode ihres Wirtes. Bisher konnte den FV kein Krankheitsbild zugeordnet werden (Meiering & Linial, 2001). Da FV höchstwahrscheinlich durch Zoonose vom Affen an den Menschen übertragen wurden und eine Mensch-zu-Mensch Infektion noch nicht beobachtet werden konnte, gilt der

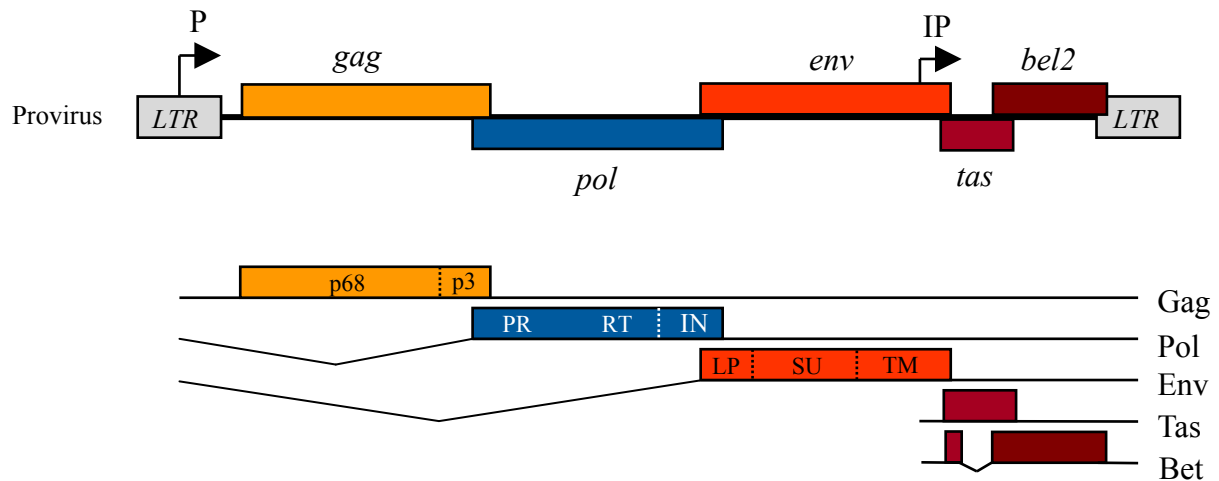
Mensch nicht als natürliches Reservoir des FV (Switzer *et al.*, 2004; Heneine *et al.*, 2003; Switzer *et al.*, 2005; Meiering & Linial, 2001). Die Übertragung von FV erfolgt über den Speichel und tritt daher bei Bissverletzungen oder durch Lecken auf (Meiering & Linial, 2001). Afrikanische Jäger und Tierpfleger, die berufsbedingt in häufigem Kontakt mit Affen stehen, stellen die größte Risikogruppe dar.

In starkem Kontrast zum apathogenen Erscheinungsbild der FV, das sie in ihren natürlichen Wirten zeigen, steht ein starker zytopathischer Effekt (engl. *cytopathic effect*, CPE), der in Zellkulturen auftritt (Linial, 2007). Typische Merkmale des CPE sind die Synzytienbildung durch Zell-Zell-Fusion und Bildung von mehrkernigen Riesenzellen sowie eine zytoplasmatische Vakuolisierung, die schaumig erscheint (Meiering & Linial, 2001). Dieser Effekt ist der Grund für den Namen der Spumaretroviren bzw. FV (engl. *foamy* für schaumig; lat. *spuma* Schaum) (Flügel, 2012).

Durch ihre Apathogenität aufgrund der Jahrtausenden langen Koexistenz in ihren Wirtsorganismen (Switzer *et al.*, 2005) und durch ihre Eigenschaft als Retrovirus ins Wirtsgenom zu integrieren, bieten FV ein vielversprechendes Instrument für gentherapeutische Ansätze. Zusätzlich zeigen FV neben einem breiten Wirtsspektrum auch einen ausgeprägten Zelltyp-Tropismus. Damit FV medizinische Anwendung finden können, muss der Replikationszyklus im Einzelnen aufgeklärt sein. Sowohl PFV als humanes FV als auch der Simian FV aus Makaken (SFV<sub>mac</sub>) bieten beide sehr gute Voraussetzungen für weitere Untersuchungen hinsichtlich ihrer Replikationsmechanismen. Im besonderen Interesse steht der strukturelle Aufbau der Reversen Transkriptase des jeweiligen Virus. Dieses Enzym ist lebenswichtig für jedes Retrovirus.

### 1.3 Genomorganisation und Aufbau des Viruspartikels in FV

Das provirale Genom aller Retroviren hat eine Größe von 9-12 kb und besteht aus den Genen *gag*, *pol* und *env*, welche von langen terminalen Wiederholungssequenzen (engl. *long terminal repeats*, LTRs) flankiert werden. Dabei kodiert *gag* die viralen Strukturproteine (in FV: p68, p3), *pol* die virale Protease (PR), die Reverse Transkriptase (RT) sowie die Integrase (IN) und *env* die viralen Hüllproteine (in FV: *leaderpeptide* (LP), Oberflächenprotein (SU) und Transmembranprotein (TM)) (Abb. 1.1). Während die kanonische Reihenfolge der Gene bei allen Retroviren 5'-*gag-pol-env*-3' lautet und ihre Expression im 5'-LTR-Promotorbereich beginnt, unterscheidet sich die Art der Expression dieser Gene innerhalb der Genera der Retroviren. Env wird in allen Retroviren von einer separaten mRNA übersetzt.



**Abb. 1.1: Genomorganisation des PFV und dazugehörige transkribierte mRNAs.** LTR: *long terminal repeat*; P: LTR-Promotor; IP: interner Promotor; *gag*: Gen der viralen Strukturproteine p68 und p3; *pol*: Gen der viralen Enzyme PR (Protease), RT (Reverse Transkriptase) und IN (Integrase); *env*: Gen für die viralen Hüllproteine einschließlich LP (*leaderpeptide*), SU (Oberflächenprotein) und TM (Transmembranprotein); *tas* und *bet2*: Gene kodieren die akzessorischen Proteine Tas und Bet. Die Schnittstellen der Vorläuferproteine werden durch schwarz bzw. weiß gestrichelte Linien dargestellt.

In Orthoretroviren werden Gag- bzw. Gag-Pol-Fusionsproteine von einer ungespleißten mRNA translatiert. Dabei kommt es am 3'-Ende von *gag* zum Überlesen eines Stopcodons ( $\gamma$ - und  $\epsilon$ -Retroviren) oder zu einer bis zwei ribosomalen -1 Leserasterverschiebungen (engl. *frameshift*) ( $\alpha$ -,  $\beta$ -,  $\delta$ -Retroviren, Lentiviren) (Goff, 2007). Im Gegensatz dazu wird in FV Pol unabhängig von Gag von einer gespleißten mRNA translatiert. Zusätzlich zu Gag, Pol und Env kodieren komplexe Retroviren wie FV oder Lentiviren akzessorische Proteine. Bei den FV sind das Tas und Bet. Dabei wird Tas von der *tas*-mRNA und Bet von einer gespleißten mRNA aus *tas* und *bel2* translatiert. Eine Besonderheit der FV zeigt sich dadurch, dass die Gene dieser beiden Proteine von einem internen Promotor (IP), der sich am 3'-Ende von *env* befindet, exprimiert werden (Rethwilm *et al.*, 1991). Tas bindet als Transaktivator der Transkription an den Promotor am 5'-LTR (5'-LTRP) sowie an den IP. Dem IP wurde nicht nur eine höhere Affinität zu Tas nachgewiesen (Kang *et al.*, 1998), er zeigt darüber hinaus auch eine höhere, Tas-unabhängige Basalaktivität als der 5'-LTRP (Löchelt *et al.*, 1993). Dadurch verstärkt Tas seine eigene Transkription und aktiviert dosisabhängig die Transkription der anderen Gene beginnend beim 5'-LTRP (Rethwilm *et al.*, 1991). Es konnte gezeigt werden, dass Tas für die Infektiosität der FV unabdingbar ist (Baunach *et al.*, 1993). Für die Funktion von Bet gibt es Hinweise, dass es die antiretrovirale Aktivität der Cytidindesaminasen der *Apolipoprotein B messenger RNA-editing, catalytic polypeptide* (APOBEC) 3-Genfamilie inhibiert (Löchelt *et al.*, 2005; Russell *et al.*, 2005; Perkovic *et al.*, 2009).

Der Reifeprozess zu infektiösen Viren umfasst neben dem Zusammenbau und der Knospung auch die Prozessierung von Gag und Pol durch virale Proteasen sowie von Env durch zelluläre Proteasen. In Lentiviren wird Gag proteolytisch in Matrix- (engl. *matrix*, MA), Kapsid- (engl. *capsid*, CA) und Nukleokapsidproteine (engl. *nucleocapsid*, NC) gespalten. In FV wird Gag nur einmal geschnitten, wobei das 71 kDa-Vorläuferprotein am C-Terminus um 3 kDa verkürzt wird und dadurch das Peptid p3 und das Protein p68 entstehen (Flügel & Pfrepper, 2003). Frühere Studien haben gezeigt, dass diese Gag-Prozessierung für die Infektiosität wichtig ist (Enssle *et al.*, 1997; Zemba *et al.*, 1998), da mit ungeschnittenen Gag-Vorläuferproteinen nicht-infektiöse Viruspartikel gebildet werden (Enssle *et al.*, 1997; Konvalinka *et al.*, 1995). Die genaue Funktionsweise von p3 ist bisher ungeklärt. Während in Orthoretroviren das Gag-Vorläuferprotein für die Bildung eines nicht-infektiösen Viruspartikels ausreichend ist, verläuft die Partikelbildung von Gag in FV Env-abhängig (Wilk *et al.*, 2001; Lindemann *et al.*, 2001).

Env wird in allen Retroviren in ein Transmembran- (engl. *transmembrane*, TM) und ein Oberflächenprotein (engl. *surface*, SU) gespalten, die durch nicht-kovalente Bindungen miteinander assoziiert sind. Darüber hinaus wird von Env in FV zusätzlich am N-Terminus ein *leaderpeptide* (LP) abgespalten, das während der Partikelmorphogenese mit dem N-Terminus von Gag interagiert (Wilk *et al.*, 2001).

Das Pol-Vorläuferprotein in FV besitzt eine PR-, eine RT- und eine IN-Domäne und wird nur einmal zwischen PR-RT und IN proteolytisch geschnitten (Hartl *et al.*, 2010a; Pfrepper *et al.*, 1998). Auch darin unterscheiden sich FV von den Orthoretroviren, in denen die PR als separate, aktive Domäne vorliegt. Die PR-RT ist daher für die Prozessierung von Gag und Pol sowie auch für die reverse Transkription der viralen ssRNA in dsDNA verantwortlich. Die IN integriert anschließend die revers transkribierte dsDNA in das Wirtsgenom.

Der Aufbau von infektiösen Partikeln verschiedener Retroviren ist ähnlich. Dabei liegt die Größe der Partikel in FV bei 100-140 nm (Lecellier & Saib, 2000). Das Viruskapsid besteht aus einer Lipiddoppelschicht, die sich von der Zellmembran der Wirtszelle ableitet. Neben den Membranproteinen des Wirtes befinden sich auch die Hüllproteine (Env) des Virus in dieser Lipiddoppelmembran. In FV besteht das Viruskapsid aus radialsymmetrisch angeordneten Gag-Proteinen (Wilk *et al.*, 2001) und beinhaltet neben der PR-RT, der IN und viraler RNA in 20 % der extrazellulär untersuchten Viruspartikel auch längere DNA-Fragmente. Diese extrahierten längeren DNA-Fragmente sind nach Transfektion infektiös (Yu *et al.*, 1999). Für HIV wurde bisher nur bei 0,001 % der Virionen längere DNA Fragmente gefunden, die für die Infektiosität zudem nicht notwendig waren (Yu *et al.*, 1999).

Studien mit dem RT-Inhibitor 3'-Azido-3'-desoxythymidin (AZT) zeigten, dass Zellen, die mit AZT vorbehandelt worden waren, mit FV infiziert werden konnten. Viren jedoch, die aus mit AZT behandelten Zellen gewonnen worden waren, wiesen keine Infektiosität mehr auf (Yu *et al.*, 1999; Moebes *et al.*, 1997). Alle diese Experimente deuten darauf hin, dass FV ihre genomische ssRNA während des Verlassens der Zelle in dsDNA umschreiben und die reverse Transkription in der späten Phase des viralen Lebenszyklus in der virusproduzierenden Zelle erfolgt.

### 1.4 Der Replikationszyklus

Um in die Zielzelle zu gelangen, muss das Virion über Env an einen bisher nicht identifizierten Rezeptor der Zielzelle binden. Kürzlich veröffentlichte Ergebnisse zeigten, dass Heparansulfat einen wichtigen Faktor für die Zellanbindung darstellt (Nasimuzzaman & Persons, 2012; Plochmann *et al.*, 2012). Nach der Bindung an den Rezeptor wird das Viruspartikel endozytotisch aufgenommen. Daraufhin muss das Viruskapsid und damit Gag durch die virale PR prozessiert werden, damit sich der zur Integration der dsDNA notwendige Präintegrationskomplex bilden kann (Lehmann-Che *et al.*, 2005; Lindemann & Rethwilm, 2011). Anschließend wird die dsDNA des Virus ins Wirtschromosom integriert. Ausgehend vom integrierten Provirus wird durch die zelluläre RNA-Polymerase II die genomische RNA transkribiert und im Zytoplasma translatiert.

Das Env-Vorläuferprotein wird durch sein aminoterminal LP zum rauen endoplasmatischen Retikulum transportiert und gelangt über den sekretorischen Transportweg (engl. *secretory pathway*) an seinen Bestimmungsort, die Zelloberfläche (Lindemann & Rethwilm, 2011). Die intrazelluläre Kapsidbildung von FV im Zytoplasma entsteht durch Gag-Gag-Interaktionen am N-Terminus des Gag-Vorläuferproteins (Tobaly-Tapiero *et al.*, 2001). Eine weitere spezifische Interaktion des N-Terminus von Gag mit dem LP des Env-Proteins ist essentiell für die Knospung der Viruspartikel (Wilk *et al.*, 2001; Geiselhart *et al.*, 2003). *In vitro*-Zellkulturversuche zeigten, dass der C-Terminus des Gag-Vorläuferproteins wichtig für die Inkorporation von Pol sowie der RNA ist (Lee & Linial, 2008; Müllers *et al.*, 2011). Weiterhin ist für die Aufnahme von Pol ins Viruspartikel und die Aktivierung der PR die Bindung an bestimmte Sequenzen der genomischen RNA entscheidend (Peters *et al.*, 2005; Peters *et al.*, 2008; Hartl *et al.*, 2011; Spannaus *et al.*, 2012). Obwohl die genauen Zusammenhänge unklar sind, herrscht zurzeit die Vorstellung, dass die RNA als verbrückendes Molekül zwischen Gag und Pol fungiert und es durch die Interaktion aller drei



Komponenten auch zur Verpackung der RNA kommt (Lee & Linial, 2008). Die zur Reifung notwendigen Prozesse der reversen Transkription der genomischen RNA und der Prozessierung von Gag und Pol durch die virale PR sind in ihrem zeitlichen Ablauf noch nicht geklärt. Sicher ist nur, dass ein infektiöses Viruspartikel dsDNA sowie prozessiertes Gag und Pol benötigt (Yu *et al.*, 1999; Moebes *et al.*, 1997; Enssle *et al.*, 1997).

## 1.5 Die reverse Transkription

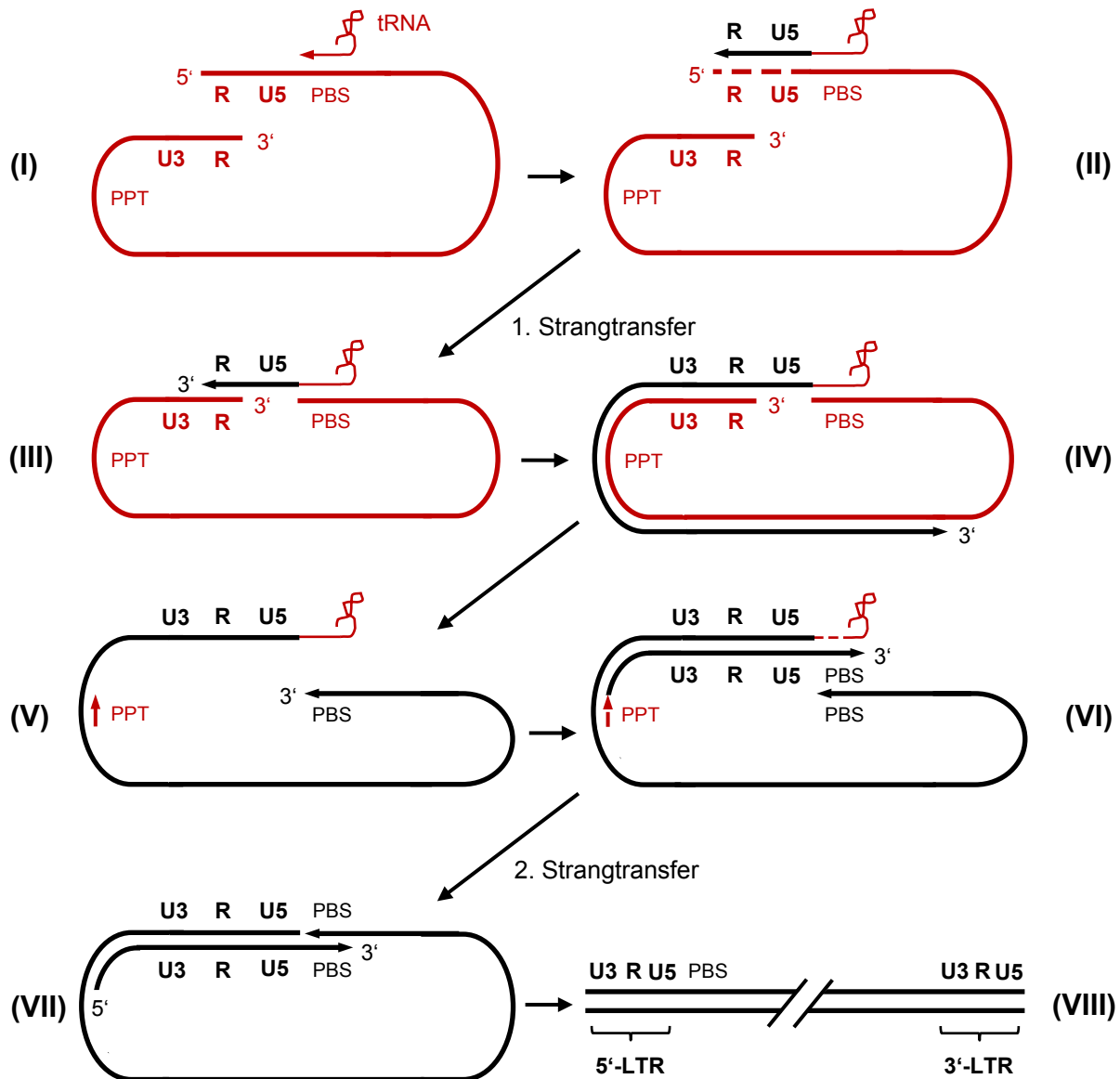
Bis zur Entdeckung der RT in den 1960er Jahren (Temin, 1964; Temin & Mizutani, 1970; Baltimore, 1970) lautete das zentrale Dogma der Molekularbiologie, dass der Fluss der genetischen Information von DNA über RNA zum Protein unumstößlich sei. Die Entdeckung der Retroviren hat dieses Dogma durchbrochen, da die Viren in ihrem einzigartigen Lebenszyklus ihr RNA-Genom revers in dsDNA transkribieren und diese anschließend in das Wirtsgenom integrieren.

Die RT ist ein multifunktionales Enzym und vereinigt drei enzymatische Aktivitäten:

- eine RNA-abhängige DNA-Polymerase
- eine DNA-abhängige DNA-Polymerase
- eine Ribonuklease H (RNase H)

Dabei ist die N-terminal gelegene Polymerase-Domäne für die DNA-Synthese verantwortlich, während die RNase H-Domäne am C-Terminus der RT den RNA-Strang im RNA/DNA-Hybrid hydrolytisch spaltet. Die reverse Transkription beginnt mit der Synthese der (-)-*strong-stop*-DNA. Dazu bindet bei FV eine zelluläre tRNA<sup>Lys 1,2</sup> an eine 18 Nukleotide lange Primer-Bindestelle (engl. *primer binding site*, PBS) (Maurer *et al.*, 1988) im 5'-Bereich des RNA-Genoms ((+)-RNA-Strang) (Abb. 1.2, Schritt (I)). Das 3'-Ende der tRNA dient der RT als Primer zur DNA-Synthese in 5' → 3' Richtung bis zum 5'-Ende der genomischen RNA (Schritt (II)). Der synthetisierte DNA-Bereich erstreckt sich über die sogenannte R-U5-Sequenz des späteren 3'-LTR. Der RNA-Anteil im neu gebildeten RNA/DNA-Doppelstrang wird durch die RNase H abgebaut (Moelling *et al.*, 1971). Der verbleibende, einzelsträngige DNA-Strang kann durch die vorliegende R-Sequenz an die komplementäre R-Sequenz am 3'-Ende des RNA-Genoms hybridisieren (Schritt (III)) und die Synthese des (-)-DNA-Stranges vollendet werden. Dieser Strangtransfer wird als der erste Sprung (engl. *first jump*) bezeichnet und kann inter- bzw. intramolekular stattfinden (Goff, 2007). Während der Polymerisation des (-)-DNA-Stranges wird der RNA-Strang durch die RNase H abgebaut, wobei ein ca. 17-19 Basen langer, purinreicher Sequenzabschnitt (engl. *polypurine tract*, PPT) durch einen

spezifischen RNase H-Schnitt erhalten bleibt (Schritt (IV) und (V)) (Huber & Richardson, 1990). Dieser PPT dient wiederum als Primer für die Synthese der (+)-*strong-stop*-DNA, die als Matrize die U3-R-U5-Sequenz des 3'-LTR und die PBS-Sequenz der noch angeknüpften tRNA zur Verfügung hat (Schritt (VI)). Die tRNA des (-)-DNA-Stranges sowie der PPT des (+)-DNA-Stranges werden durch die RNase H abgeschnitten. Da die PBS des entstehenden (+)-DNA-Stranges nach dem RNase H-Verdau nicht mehr von der tRNA gebunden und komplementär zur PBS des (-)-DNA-Stranges ist, hybridisieren diese beiden Regionen und dienen einander als Matrize zur vollständigen Synthese der proviralen DNA (Schritt (VII)). Da an dieser Stelle wieder ein Strangtransfer stattfindet, wird diese zweite, intramolekulare Umlagerung als zweiter Sprung (engl. *second jump*) bezeichnet. Am Ende dieses Prozesses liegt eine provirale dsDNA vor, die an beiden Seiten von vollständigen LTRs flankiert wird (Schritt (VIII)) (Goff, 2007). Diese LTRs kodieren nicht nur den Promotor zur Transkription der viralen Gene, sondern sie sind auch wichtig für den Integrationsprozess und für verschiedene Spleißvorgänge.



**Abb. 1.2: Schematische Darstellung der reversen Transkription.** Die RNA ist in rot und die DNA in schwarz dargestellt. PBS: Primer-Bindestelle; PPT: *polypurine tract*; LTR: *long terminal repeat*. U3: *LTR unique 3' region*; U5: *LTR unique 5' region*; R: *LTR repeat region*. Weitere Erklärungen siehe Text. Verändert nach (Champoux & Schultz, 2009).

## 1.6 Die Reverse Transkriptase

Der Aufbau der Untereinheiten und der Oligomerisierungszustand der RT ist in den Retroviren sehr verschieden. FV, Moloney Mausleukämievirus (engl. *Moloney Murine Leukemia Virus*, MoMLV) sowie das dem Xenotropischen Mausleukämievirus verwandte Virus (engl. *Xenotropic Murine Leukemia Virus related Virus*, XMRV) besitzen monomere RTs (Hartl *et al.*, 2010a; Moelling, 1974; Verma, 1975; Roth *et al.*, 1985; Das & Georgiadis, 2004; Ndongwe *et al.*, 2012). Dabei beinhaltet die FV RT N-terminal noch eine PR-Domäne (PR-RT). Diese Eigenschaft der PR-RT ist unter den Retroviren bisher einmalig. Für RSV

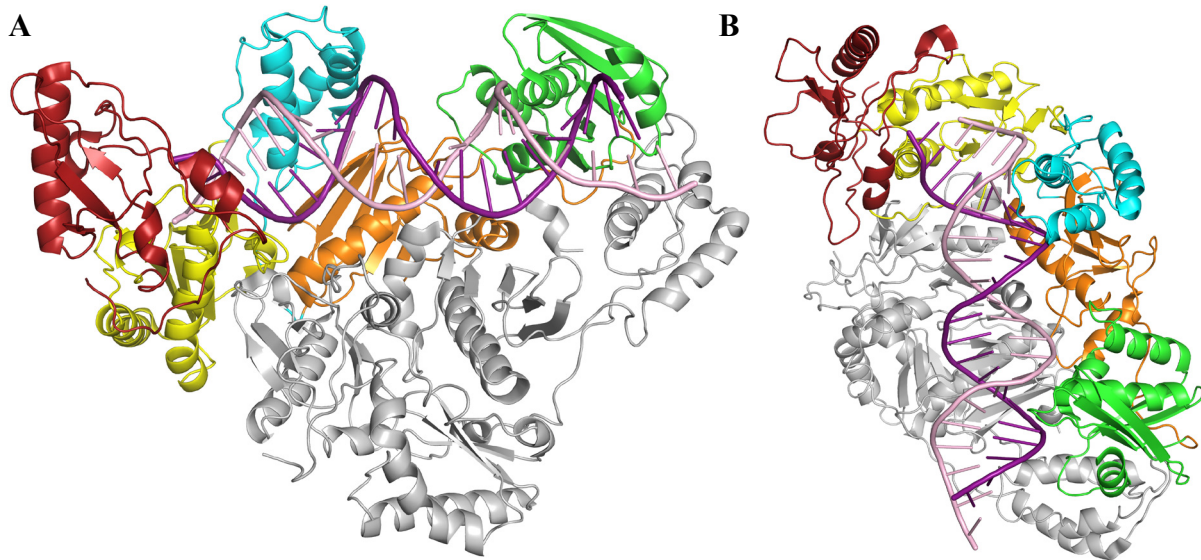
und HIV wurden heterodimere RTs als dominierende aktive Enzymform nachgewiesen. Bei RSV geht man von einem  $\alpha/\beta$  Heterodimer aus, dessen  $\alpha$ -Untereinheit aus der Polymerase- und RNase H-Domäne besteht und die  $\beta$ -Untereinheit zusätzlich zur  $\alpha$ -Untereinheit C-terminal die Integrase-Domäne besitzt (Rho *et al.*, 1975; Golomb & Grandgenett, 1979). Die RT von HIV-1 besteht aus den beiden Untereinheiten p66 und p51, wobei p66 die Polymerase- sowie RNase H-Domäne enthält. p51 ist sequenzidentisch zu p66, besitzt jedoch keine RNase H-Domäne (di Marzo Veronese *et al.*, 1986; Hansen *et al.*, 1988; Kohlstaedt *et al.*, 1992). Für HTLV-1 wird sogar eine oligomere RT-Zusammensetzung angenommen ( $\alpha_3/\beta$  Variante) (Trentin *et al.*, 1998). Trotz der Unterschiede im Aufbau und Oligomerisierungszustand zwischen den retroviralen RTs ist die Funktion der Polymerase- sowie RNase H-Domäne der verschiedenen RTs gleich.

Bisher wurden die Strukturen der RTs von HIV-1, HIV-2 (pdb: 1MU2) (Ren *et al.*, 2002), MoMLV (pdb: 1RW3) (Das & Georgiadis, 2004) und XMRV (pdb: 4HKQ) (Nowak *et al.*, 2013) gelöst. Die Elektronendichten der RNase H-Domänen der MoMLV und XMRV RTs waren für eine Strukturlösung jedoch nicht ausreichend. Während noch keine vollständige Struktur einer monomeren RT vorliegt, wurden schon eine Vielzahl an RT Strukturen mit und ohne Ligand für HIV-1 gelöst.

### 1.6.1 Die Polymerase-Domäne der Reversen Transkriptase

Die Polymerase-Domäne der RT nutzt RNA oder DNA als Matrize zur Synthese eines DNA-Stranges. Obwohl die Strukturen der monomeren RTs von MoMLV und XMRV eine bessere Grundlage für die monomere PR-RT der FV bieten, sind sie wegen ihrer unzureichenden Elektronendichten im RNase H-Bereich zur Darstellung der einzelnen Domänen und Unterdomänen ungeeignet. Deshalb wird am Beispiel der am besten untersuchten RT Struktur, der HIV-1 RT, näher auf den Aufbau eingegangen. Obwohl p66 und p51 bis auf den Bereich der RNase H sequenzidentisch sind, ist die Anordnung der Subdomänen der beiden Untereinheiten im Heterodimer unterschiedlich (Abb. 1.3). Die Polymerase-Domäne der p66-Untereinheit besteht aus vier Subdomänen. Da die Anordnung und Interaktion der Subdomänen einer rechten Hand ähneln, werden drei dieser Subdomänen auch Finger, Daumen und Handfläche (engl. *finger, thumb, palm*) genannt (Kohlstaedt *et al.*, 1992). An diese drei Subdomänen schließt sich die Verbindungs-Subdomäne (engl. *connection*) an. Sie sorgt für den Anschluss der RNase H an die Polymerase-Domäne (Rodgers *et al.*, 1995). Die

p66-Untereinheit besitzt die aktiven Zentren der Polymerase- sowie der RNase H-Domäne, während p51 eine vorwiegend stabilisierende Funktion zukommt.



**Abb. 1.3: Kristallstruktur der HIV-1 RT.** (A) Bänderdarstellung der HIV-1 RT im Komplex mit dem RNA/DNA-Hybrid des *polypurine tract* (pdb: 1HYS). Die Finger- (rot), Handfläche- (gelb), Daumen- (cyan), Verbindungs- (orange) Subdomänen sowie die RNase H-Domäne (grün) von p66 sind farbig und p51 ist grau dargestellt. Die RNA und DNA wurden violett und rosa eingefärbt. (B) Bänderdarstellung der HIV-1 RT in Komplex mit dem RNA/DNA-Hybrid des *polypurine tract* (pdb: 1HYS) von oben und um 90° im Uhrzeigersinn zu (A) gedreht. Farbdarstellung siehe (A).

Das aktive Zentrum der Polymerase von HIV-1 befindet sich in der Handflächen-Subdomäne und besteht aus drei Aspartatresten, die zwei divalente Metallionen ( $Mg^{2+}$  bzw.  $Mn^{2+}$ ) koordinieren. Zwei dieser Aspartatreste sind Teil des hochkonservierten YXDD-Sequenzmotivs (Übersichtsartikel (Sarafianos *et al.*, 2009)). Bei HIV-1 steht das X für Methionin, während man in FV und MLV Valin findet (Linial, 2007).

Bei der Substratbindung benötigt die HIV-1 RT einen Primer- und den dazugehörigen Matrizenstrang. Dabei umschließen die Daumen- und die Finger-Subdomäne ähnlich einer rechten Hand das Substrat, wobei sich die Daumen-Subdomäne bis zu 30° auf die Finger-Subdomäne zubewegen kann (Rodgers *et al.*, 1995). Durch diesen „Haltegriff“ wird das 3'-Ende des Primers in die Primerbindestelle (P-Stelle) in der Nähe des aktiven Zentrums der Polymerase positioniert. Bei der anschließenden Polymerisationsreaktion bindet ein Desoxyribonukleosidtriphosphat (dNTP) an die Nukleotidbindungsstelle (N-Stelle) und es kommt zu einer Konformationsänderung der RT, bei der sich die Finger-Subdomäne um das eintretende dNTP schließt. Dadurch werden das  $\alpha$ -Phosphat des Nukleotides und das 3'-OH des Primers im aktiven Zentrum der Polymerase optimal positioniert, sodass es zur Bildung einer Phosphodiesterbindung zwischen dNTP und Primer kommt. Das dabei entstehende Pyrophosphat kann durch das Öffnen der Finger das aktive Zentrum wieder verlassen. Um die

Polymerisationsreaktion fortzusetzen, kommt es zur Translokation des verlängerten Primers von der N-Stelle zur P-Stelle (Sarafianos *et al.*, 2009).

Da die RT ein Enzym ist, das ausschließlich vom Virus und nicht von der Zelle benötigt wird, stellt es ein Zielobjekt antiretroviraler Therapien dar. Angriffsstellen der bisher eingesetzten Inhibitoren betreffen nicht nur die retroviralen Enzyme wie die Polymerase-, IN- und PR-Domäne, sondern auch die Aufnahme des Virus in die Zielzelle (Arts & Hazuda, 2012). Für die RNase H konnte bisher noch kein medizinisch einsetzbares Medikament entwickelt werden.

### 1.6.2 Die RNase H-Domäne der Reversen Transkriptase

Die RNase H ist eine Endonuklease, die während der reversen Transkription den RNA-Anteil in RNA/DNA-Hybriden abbaut. Sie befindet sich am C-terminalen Ende der RT der meisten Retroviren.

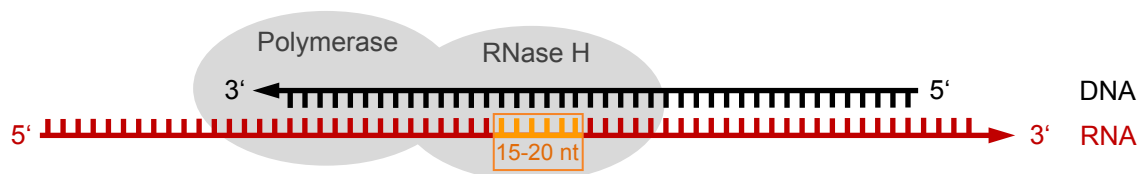
Da die retrovirale RNase H mit der Polymerase-Domäne verbunden vorliegt, unterscheidet man zwischen polymerase-abhängiger und polymerase-unabhängiger RNase H-Aktivität. Bei der polymerase-abhängigen Aktivität schneidet die RNase H das Substrat präferentiell in einer Entfernung von ca. 18 Nukleotiden vom aktiven Zentrum der Polymerase-Domäne (Wöhrl & Moelling, 1990). Kristallstrukturen der HIV-1 RT in Komplex mit dsDNA (Jacobo-Molina *et al.*, 1993) bzw. einem RNA/DNA-Hybrid des PPT (Sarafianos *et al.*, 2001) zeigen, dass der physikalische Abstand zwischen dem 3'-Ende der DNA, das im aktiven Zentrum der Polymerase positioniert ist, und der Substratregion nahe dem aktiven Zentrum der RNase H 18 bzw. 17 Nukleotide beträgt.

Darüber hinaus gibt es bei der RNase H drei Schnittmodi (Abb 1.4). Man unterscheidet zwischen (i) einem durch das 3'-Ende des DNA-Primers bestimmten, (ii) einem durch das 5'-Ende der RNA bestimmten und (iii) einem internen Schnittmodus (Übersichtsartikel (Schultz & Champoux, 2008) und (Basu *et al.*, 2008)). Bei (i) positioniert sich die Polymerase-Domäne an das zurückgesetzte 3'-Ende des DNA-Primers und der RNase H-Schnitt erfolgt in der RNA ca. 15-20 Nukleotide entfernt vom 3'-Ende der DNA. Dieser Schnitt kann während einer Pause bei der Polymerisation bzw. auch polymerisationsunabhängig stattfinden, d. h. die Polymerase bindet zwar das Hybrids substrat, synthetisiert aber keine DNA. Beim zweiten Schnittmodus (ii) bindet die Polymerase auch den DNA-Strang, orientiert sich aber am zurückgesetzten 5'-Ende der RNA. Der RNase H-Schnitt ist in der RNA ca. 13-19 Nukleotide vom 5'-Ende der RNA zu beobachten und ist polymerisationsunabhängig. Wie beim

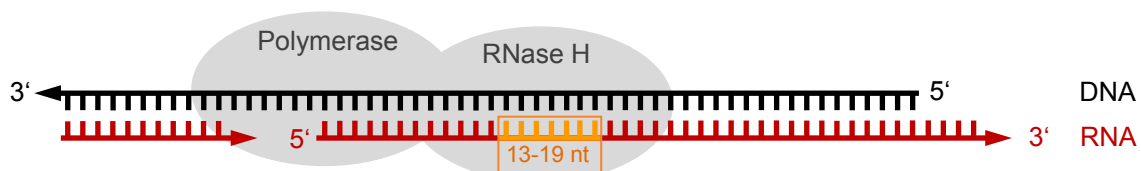
Schnittmodus (i) ist die Spezifität des RNase H-Verdau hier abhängig vom Abstand zum zurückgesetzten Oligonukleotidende und der Nukleotidsequenz in unmittelbarer Nähe der Schnittstelle. Für den Schnittmodus (ii) ist außerdem noch die Zugänglichkeit des 5'-Endes der RNA wichtig. Es muss eine Lücke von mindestens zwei Nukleotiden am 5'-Ende der RNA vorhanden sein, damit die RNase H der RT in diesem Modus schneiden kann. Für (i) und (ii) konnten auch schon Schnitte beobachtet werden, die ca. fünf bis acht Nukleotide vom Oligonukleotidende entfernt waren. Das zeigt, dass die RT nach einer anfänglich gerichteten Bindung nicht starr verharrt, sondern einige Nukleotide am Substrat entlang gleiten und einen sogenannten sekundären Schnitt setzen kann. Der interne Schnittmodus (iii) ist syntheseunabhängig und bedarf keiner Orientierung der Polymerase-Domäne am 3'-Ende der DNA bzw. 5'-Ende der RNA. Auch hier bestimmt die Nukleotidsequenz in der Nähe der Schnittstelle die Spezifität des RNase H-Schnittes.

### RNase H Schnittmodi

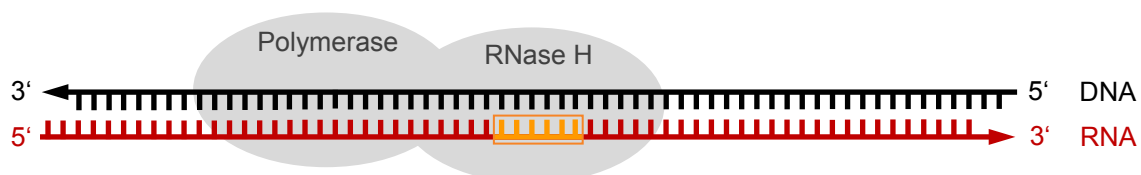
(i) durch das 3'-Ende des DNA-Primers bestimmter Schnittmodus



(ii) durch das 5'-Ende der RNA bestimmter Schnittmodus

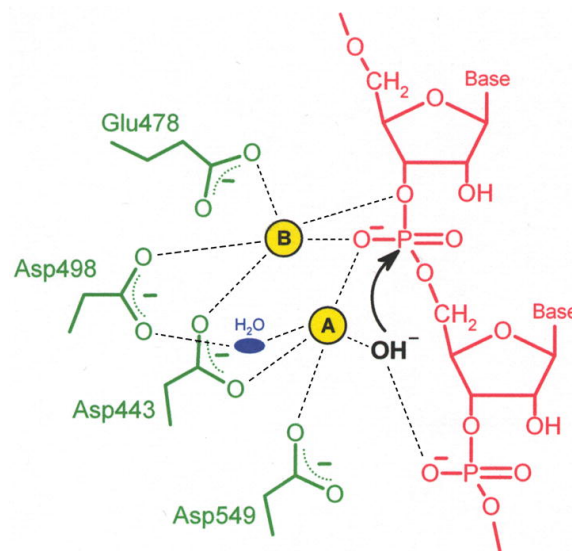


(iii) interner Schnittmodus



**Abb 1.4: Die drei Schnittmodi der retroviralen RNase H.** Das Schnittfenster der RNase H ist orange hervorgehoben. **(i)** Der RNase H-Schnitt wird durch das 3'-Ende des DNA-Primers bestimmt. Dabei positioniert sich die Polymerase-Domäne an das zurückgesetzte 3'-Ende der DNA, von dem aus im Abstand von 15-20 nt entfernt der RNA-Strang geschnitten wird. **(ii)** Die RNA wird in einem Abstand von 13-19 nt vom zurückgesetzten 5'-Ende der RNA entfernt geschnitten, wobei sich die Polymerase am 5'-Ende der RNA orientiert. **(iii)** Bei längeren RNA/DNA-Hybrid-Sequenzen erfolgt der RNase H-Schnitt unabhängig vom 3'-Ende der DNA oder 5'-Ende der RNA und wird durch die RNA-Sequenz in der Nähe der Schnittstelle bestimmt. Verändert nach (Schultz & Champoux, 2008).

2005 und 2007 wurden von Nowotny *et al.* die Kristallstrukturen der RNase H von *Bacillus halodurans* (*B. halodurans*) (Nowotny *et al.*, 2005) und der humanen RNase H1 (Nowotny *et al.*, 2007) im Komplex mit einem RNA/DNA Hybrid gelöst. Aufgrund der strukturellen Ähnlichkeiten zur HIV-1 RNase H (Davies *et al.*, 1991) und der MoMLV RNase H-Variante ohne C-Helix (Lim *et al.*, 2006) sowie Mutationsexperimenten (Mizrahi *et al.*, 1994; Mizrahi *et al.*, 1990) geht man derzeit davon aus, dass im aktiven Zentrum der RNase H vier hochkonservierte saure Reste (DEDD-Motif) zwei  $Mg^{2+}$ -Ionen koordinieren. Es konnte zwar gezeigt werden, dass die RNase H der HIV-1 und MoMLV RT auch mit  $Mn^{2+}$  aktiv ist; diese Aktivität geht aber mit geringerer Substratspezifität einher (Schultz & Champoux, 1996; Ben-Artzi *et al.*, 1992) und stellt die biologische Relevanz der Manganionen damit in Frage. Bei diesem katalytischen 2-Metallionen-Mechanismus aktiviert Metallion A ein Wassermolekül im aktiven Zentrum, das dann als Nukleophil das Phosphatatom der Phosphodiesterbindung zwischen zwei Nucleotiden angreift (Abb. 1.5). Der pentakovalente Übergangszustand des Phosphatatoms wird von Metallion B stabilisiert. Diese RNA-Phosphoryltransferreaktion unter Beteiligung von zwei Metallionen führt zu 5'- $PO_4^-$ - und 3'-OH-Reaktionsprodukten (Steitz & Steitz, 1993).



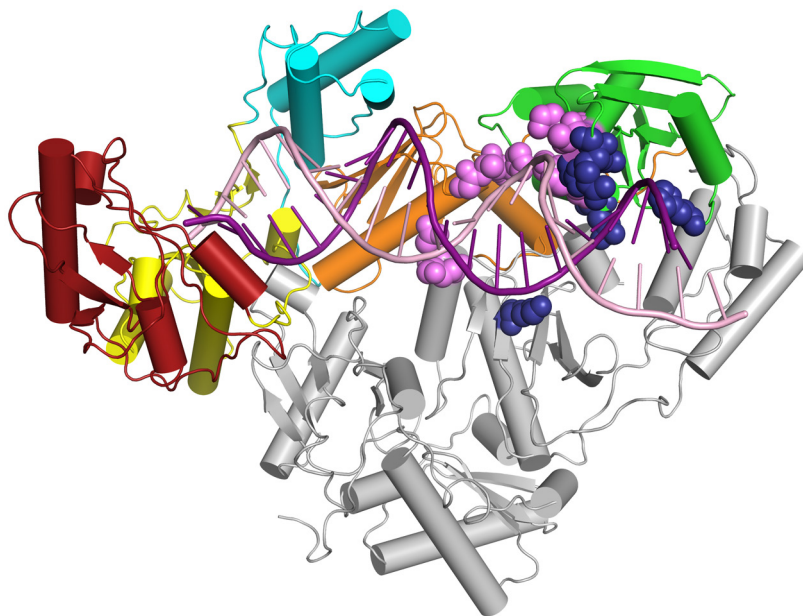
**Abb. 1.5: Der 2-Metallionen-Mechanismus der retroviralen RNase H.** Schematische Darstellung des 2-Metallionen-Mechanismus am Beispiel der HIV-1 RNase H (Schultz & Champoux, 2008). Die Reste des aktiven Zentrums (Asp549, Asp443, Asp498, Glu478) sind grün dargestellt; die beiden Metallionen sind mit A und B gekennzeichnet und gelb eingefärbt. Die Koordination der Metallionen sowie Wasserstoffbrückenbindungen sind als gestrichelte Linien dargestellt. Das Nucleophil, welches das RNA-Substrat (rot) angreift, ist schwarz und das verbrückende Wassermolekül zwischen Metallion A und Asp498 ist als blaues Oval gezeigt. Für weitere Erklärungen siehe Text.

Da die RNase H-Domäne retroviraler RTs immer kovalent an den C-Terminus der Polymerase-Domäne gebunden ist, wird auch ihre spezifische Aktivität von der Polymerase-Domäne mitbestimmt (Schultz & Champoux, 1996; Zhan & Crouch, 1997).



Bisher wurden drei Strukturen der HIV-1 RT mit dsDNA (pdb: 2HMI) (Ding *et al.*, 1998), mit einem RNA/DNA-Hybrid des PPT (pdb: 1HYS) (Sarafianos *et al.*, 2001) und mit einem unspezifischen RNA/DNA-Hybrid-Substrat (pdb: 4B3O) (Lapkouski *et al.*, 2013) gelöst, deren Substrate ohne Vernetzung (engl. *cross-linking*) mit dem Enzym im Komplex kristallisierten. Die Lage der Substrate in den beiden erstgenannten Strukturen ist sehr ähnlich; das 3'-OH-Ende des Primers befindet sich im aktiven Zentrum der Polymerase-Domäne und der Matrizenstrang verfehlt das aktive Zentrum der RNase H um ca. 4 Å (Nowotny *et al.*, 2007). Die letztgenannte Struktur wurde mit einem nicht-nukleosidischen Reverse Transkriptase Inhibitor gelöst, wobei das 3'-OH-Ende des Primers nicht im aktiven Zentrum der Polymerase-Domäne liegt. Der Matrizenstrang befindet sich ca. 2 Å vom aktiven Zentrum der RNase H entfernt (Lapkouski *et al.*, 2013).

Die HIV-1 RT hat eine Vielzahl an Substrat-Enzym-Kontakten in der Polymerase- sowie der RNase H-Region. Bei der Polymerase werden zwei Interaktionsbereiche unterschieden; zum einen der Polymerase *primer grip*, eine Art ‚Haltegriff‘ für den Primer-Strang, und den Polymerase *template grip*, der bei der Bindung des Matrizenstranges wichtig ist (Jacobomolina *et al.*, 1993b). Die RNase H besitzt den sogenannten RNase H *primer grip* (Abb. 1.6). Dabei handelt es sich um Reste der RT in der Nähe des aktiven Zentrums der RNase H, die mit der DNA sowie der RNA Interaktionsflächen aufweisen.



**Abb. 1.6: RNase H *primer grip* in der HIV-1 RT.** Bänder- und Zylinderdarstellung des Komplexes der HIV-1 RT mit dem RNA/DNA-Hybrid des *polypurine tract* (pdb: 1HYS). Farbcode siehe Abb. 1.3; Reste des RNase H *primer grip*, welche die DNA bzw. die RNA kontaktieren, sind in rosafarbener bzw. dunkelblauer Kugeldarstellung hervorgehoben.

Die DNA wird in einer Entfernung von  $-4$  bis  $-9$  Nukleotiden vom aktiven Zentrum der RNase H von Resten, die sich in der p51-Untereinheit (K395, E396) und der p66-Untereinheit ((Verbindungsdomäne (G359, A360, H361) und RNase H-Domäne (T473, N474, Q475, K476, Y501, I505)) befinden, kontaktiert (Sarafianos *et al.*, 2001).

Der Bereich, in dem die RNA mit dem Enzym interagiert, befindet sich bei  $-2$  bis  $+2$  Nukleotide um das zu spaltende Phosphatatom. Auch hier leisten beide Untereinheiten, p51 (K390) und p66 (R448, N474, Q475, Q500, H539), einen Beitrag zur Substratbindung (Abb. 1.6). Mutationen der Aminosäuren Q475 (Volkman *et al.*, 1993), H539 (Tisdale *et al.*, 1991; Wöhrl *et al.*, 1991), H361 (p66), K395 (p51) und E396 (p51) (Julias *et al.*, 2003) beeinträchtigen die RNase H Aktivität und untermauern die vorliegenden strukturellen Daten zum RNase H *primer grip*.

### 1.7 Die RNase H in den verschiedenen Reichen der Lebewesen

Die RNase H gibt es nicht nur bei Viren, sondern auch bei Archaeen, Bakterien und Eukaryoten. Man unterscheidet Typ 1 und Typ 2 RNasen H.

In Bakterien und Archaeen gehört zum Typ 1 die RNase HI, während zu Typ 2 die RNasen HII und RNase HIII zählen. Die Unterteilung der bakteriellen sowie archaealen RNasen H in Typ 1 und 2 basiert auf Unterschieden in deren Aminosäuresequenz (Tadokoro & Kanaya, 2009; Kochiwa *et al.*, 2007). Während die RNase HII in Archaeen und Bakterien weit verbreitet ist, kommt die RNase HIII nur in einer limitierten Anzahl von Archaeen und Bakterien vor (Kochiwa *et al.*, 2007; Ohtani *et al.*, 1999). Für die bakterielle und archaeale RNase HI findet man in einzelnen Fällen eine N-terminale Verlängerung zur stärkeren Substratbindung, die sogenannte Hybridbindungsdomäne (HBD) (Tadokoro & Kanaya, 2009; Kochiwa *et al.*, 2007).

Auch in Eukaryoten gibt es Typ 1 RNasen H, zu denen die RNasen H1 (homolog zur bakteriellen RNase HI) zählen, und Typ 2 RNasen H, zu denen die RNasen H2 (homolog zur bakteriellen RNase HII) gehören. Es wurde in Eukaryoten bisher kein Homolog zur RNase HIII gefunden. Im Gegensatz zu Bakterien und Archaeen werden die eukaryotischen RNasen H aufgrund ihrer biochemischen Eigenschaften unterschieden (Ohtani *et al.*, 1999). Die eukaryotische RNase H1 ist monomer, besitzt eine HBD und in höheren Eukaryoten eine mitochondriale Zielsequenz (engl. *mitochondrial targeting sequence*), während die eukaryotische RNase H2 ein Heterotrimer ist, das aus den Untereinheiten 2A, 2B und 2C besteht. Das katalytische Zentrum befindet sich in der Untereinheit 2A. Die genaue Funktion

der Untereinheiten 2B und 2C ist unbekannt; sie sind aber für die katalytische Aktivität der RNase H2 notwendig. Bisher wurden für 2B und 2C keine Homologe in Bakterien gefunden (Cerritelli & Crouch, 2009).

Typ 1 und Typ 2 RNasen H unterscheiden sich auch hinsichtlich ihrer Substratspezifität. Die Typ 1 RNase H schneidet ein RNA/DNA-Substrat nur, wenn mindestens vier Ribonukleotide hintereinander in der DNA eingebettet vorkommen, während die Typ 2 RNase H in der Lage ist, ein einzelnes Ribonukleotid, das in der DNA auftritt, zu erkennen und 5' davon zu schneiden (Rigby *et al.*, 2008; Cerritelli & Crouch, 2009). Aufgrund dieser Eigenschaften nimmt man derzeit an, dass die RNase H2 beim Ausbau von falsch eingebauten Ribonukleotiden während der DNA-Replikation beteiligt ist (Rigby *et al.*, 2008). Mutationen in einer der drei Untereinheiten in der humanen RNase H2 führen zur autosomal rezessiven Autoimmunkrankheit Aicardi-Goutières Syndrom (Crow *et al.*, 2006). Ein Modell für die Ursache dieser Krankheit ist, dass sich durch die fehlende RNase H2-Aktivität eine Reihe unprozessierter Nukleinsäuren ansammeln, die im Falle von RNA/DNA-Hybriden auf noch ungeklärtem Weg die angeborene Immunantwort auslösen. Dadurch wird eine hohe Menge an Interferon  $\alpha$  gebildet, was zur Zerstörung der weißen Substanz und zur Gehirnverkalkung führt (Rigby *et al.*, 2008).

*In vivo*-Experimente mit Mäuseembryonen, deren RNase H1-Gen homozygot deletiert wurde, führten zu Totgeburten. Während bei der Entwicklung der Embryonen die Replikation der chromosomalen DNA nicht beeinträchtigt war, konnte eine starke Abnahme der mitochondrialen DNA (mtDNA) festgestellt werden. Aufgrund dieser Beobachtung geht man davon aus, dass die RNase H1 wichtig für Replikation der mtDNA ist (Cerritelli *et al.*, 2003) und Fehlfunktionen in den Mitochondrien aufgrund fehlender mtDNA zum Zelltod und damit auch zum Sterben des Embryos führen. In *Drosophila* führte die Deletion des RNase H1-Gens im Verpuppungsstadium zum Tod des Insekts. Dies zeigt, dass auch hier die RNase H1 für die Entwicklung des Organismus wichtig ist. Die genauen molekularen Gründe für das Sterben konnten noch nicht geklärt werden (Filippov *et al.*, 2001).

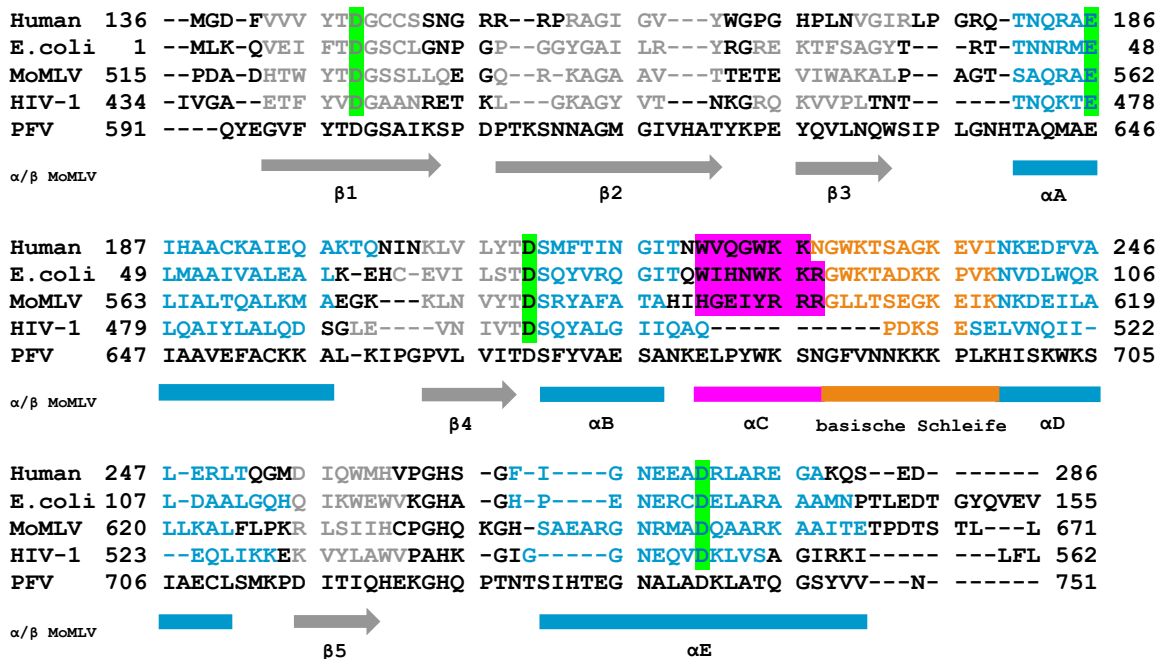
Außerdem scheinen beide RNasen H bei der Entfernung von RNA-Primern während der DNA-Replikation involviert zu sein (Aguilera & Garcia-Muse, 2012; Cerritelli & Crouch, 2009). Während Mutationen, die zur Inaktivierung der RNase H1 oder RNase H2 führen, verheerende Auswirkungen in höheren Eukaryoten zur Folge haben, weisen RNase H-Deletionsmutanten beider Typen in Hefe und Bakterien keine letalen Folgen auf (Cerritelli & Crouch, 2009; Tadokoro & Kanaya, 2009, Nowotny *et al.*, 2008, Rigby *et al.*, 2008).

### 1.8 Aktivität und Struktur isolierter RNase H-Domänen des Typs 1

Die Typ 1 RNasen H können in folgende Subfamilien eingeteilt werden: die viralen RNasen H, die bakteriellen und archealen RNasen HI und die eukaryotischen RNasen H1. Subfamilien der Typ 2 RNasen H sind: die bakteriellen RNasen HII und HIII, die archealen RNasen HIII und HIII sowie die eukaryotischen RNasen H2 (Kochiwa *et al.*, 2007; Ohtani *et al.*, 1999). Aufgrund dieser phylogenetischen Verwandtschaft der retroviralen RNasen H zu den Typ 1 RNasen H, wird im Folgenden strukturell näher auf diese RNase H-Klasse eingegangen. Die Homologie der Aminosäuresequenzen der verschiedenen Typ 1 RNase H-Domänen von Retroviren, Bakterien und Eukaryoten liegt bei ca. 30 % und ihre Tertiärstrukturen sind sehr ähnlich. Sie zeigen eine typische RNase H-Faltung (engl. *RNase H fold*), die aus einem fünfsträngigen  $\beta$ -Faltblatt mit drei längeren antiparallelen und zwei kurzen parallelen  $\beta$ -Strängen besteht. Dieses  $\beta$ -Faltblatt wird auf der einen Seite von einer langen  $\alpha$ -Helix (Helix E) flankiert, während sich auf der anderen Seite 3-4  $\alpha$ -Helices (Helix A, B, D und C) befinden, wobei die C-Helix nicht überall vorkommt.

Betrachtet man die Aktivität der isolierten retroviralen RNase H-Domänen, konnte für die separate HIV-1 RNase H keine Aktivität nachgewiesen werden (Davies *et al.*, 1991), während die isolierte MoMLV RNase H aktiv war (Tanese & Goff, 1988; Telesnitsky *et al.*, 1992). Ein Sequenzvergleich dieser beiden Domänen mit *E. coli* und humaner RNase H1 zeigt (Abb. 1.7), dass die HIV-1 RNase H im Gegensatz zur MoMLV, *E. coli* und humaner RNase H1 keine C-Helix besitzt. Darüber hinaus zeigen *E. coli* und die humane RNase H1 als isolierte Domänen Aktivität (Kanaya *et al.*, 1991; Nowotny *et al.*, 2007). Studien an der isolierten MoMLV RNase H, deren C-Helix-Bereich deletiert (I593-L603 deletiert; RH $\Delta$ C) wurde, zeigten, dass die RH $\Delta$ C keine Aktivität mehr hatte (Telesnitsky *et al.*, 1992). Viren, deren RT dieselbe Deletion der C-Helix in der RNase H-Domäne aufwies (MoMLV RT-RH $\Delta$ C), waren nicht mehr infektiös (Telesnitsky *et al.*, 1992). Weitere Studien an der C-Helix in *E. coli* zeigten, dass nicht nur die C-Helix, sondern auch eine C-terminal anschließende Schleife, die auffällig viele positiv geladene Reste besitzt, wichtig für die Substratbindung ist (Kanaya *et al.*, 1991). Die C-Helix und die anschließende basische Schleife werden als *basic protrusion* (früher *handle region* (Yang *et al.*, 1990)) zusammengefasst (Katayanagi *et al.*, 1992). Kanaya *et al.* zeigte, dass Mutationen dieser positiv geladenen Reste in der *E. coli* RNase HI den  $K_M$ -Wert 60-fach verminderten (Kanaya *et al.*, 1991). Darüber hinaus führten Deletionen im Bereich der *basic protrusion* der *E. coli* RNase HI zum Verlust der  $Mg^{2+}$ -abhängigen Aktivität, während mit  $Mn^{2+}$  eine geringe Aktivität beobachtet werden konnte

(Keck & Marqusee, 1996). Das Entfernen der fast kompletten *basic protrusion* in der *E. coli* RNase HI (H83-N100) führte zum Aktivitätsverlust des Enzyms (Haruki *et al.*, 1997). Interessanterweise konnte eine Mn<sup>2+</sup>-abhängige Aktivität der isolierten HIV-1 RNase H durch Insertion der *basic protrusion* aus *E. coli* wiederhergestellt werden (Keck & Marqusee, 1995; Stahl *et al.*, 1994).

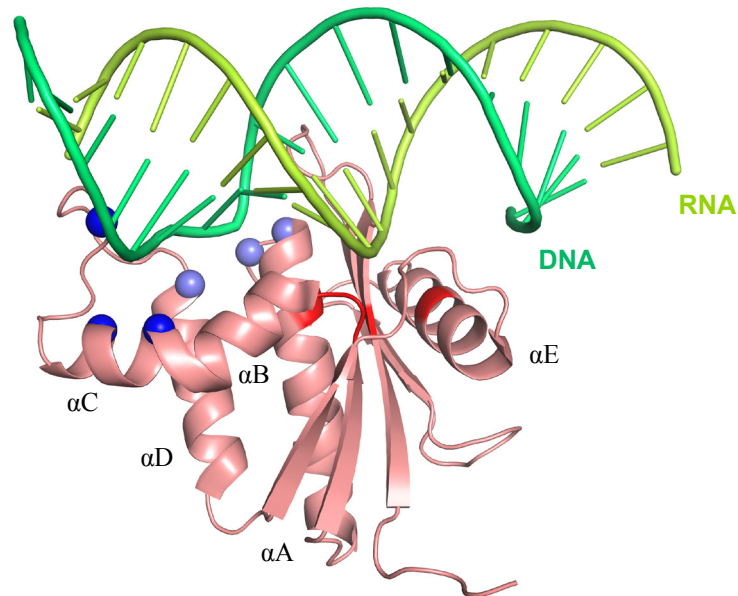


**Abb. 1.7: Vergleich der Aminosäuresequenzen verschiedener RNase H Domänen.** Gezeigt sind die Aminosäuresequenzen der humanen, *E. coli*, MoMLV, HIV-1 und PFV RNase H. Die Nummerierung der Aminosäuren der jeweiligen Domäne entspricht der des Gesamtzyms. Die Reste des aktiven Zentrums sind grün eingefärbt;  $\beta$ -Faltblattstränge und  $\alpha$ -Helices sind grau bzw. blau dargestellt. Die C-Helix sowie die anschließende basische Schleife sind pink und orange hervorgehoben. Das Sequenzalignment wurde mit dem Programm CLC Protein Workbench 5 (Version 5.3) erstellt.

Erstaunlicherweise führte der Tausch des ursprünglichen C-Helix-Bereichs (I593-L603) in der RNase H-Domäne der MoMLV RT mit dem C-Helix-Bereich (Q80-W90) der *E. coli* RNase HI zu replikationsdefizienten Viren (Lim *et al.*, 2002), obwohl die Polymerase-Domäne Aktivität zeigte. Dies spricht für eine korrekte Faltung der Polymerase-Domäne. Unklar bleibt, ob die RNase H-Domäne der MoMLV RT, die die C-Helix der *E. coli* RNase HI besitzt, *in vitro* Aktivität zeigt und ob sie richtig gefaltet ist.

Bisher wurde nur eine Struktur der Typ 1 RNase H mit einer *basic protrusion* und Substrat gelöst. Bei dieser Struktur handelt es sich um die humane RNase H1 ohne HBD im Komplex mit einem RNA/DNA-Hybrid-Substrat. Dabei bindet das RNA-Rückgrat in die RNA-Bindungsfurche, welche auch das aktive Zentrum umfasst, während die DNA in die sogenannte DNA-Bindungsfurche bindet. Die DNA-Bindungsfurche besitzt zwei DNA-

Bindestellen: die Phosphatbindungstasche (R179, T181, N240) und den DNA-Bindungskanal (W221, W225, S233) in der *basic protrusion* (Abb. 1.8).



**Abb. 1.8: Kristallstruktur der humanen RNase H1 im Komplex mit einem RNA/DNA-Hybrid.** Die Kristallstruktur der humanen RNase H1 (pdb: 2QK9) ist im lachsfarbenen Bändermodell dargestellt. Die Ca-Kohlenstoffatome der Aminosäuren der Phosphatbindungstasche und des DNA-Bindungskanals sind als hellblaue bzw. dunkelblaue Kugeln dargestellt. Die RNA und DNA sind hellgrün bzw. dunkelgrün eingefärbt und die  $\alpha$ -Helices ( $\alpha A$ - $\alpha E$ ) sind beschriftet. Das Rückgrat der Reste des aktiven Zentrums ist rot hervorgehoben.

Man geht davon aus, dass aufgrund von sterischen Konflikten zwischen der 2'OH-Gruppe der RNA und einem Indolring im DNA-Bindungskanal zwischen RNA und DNA sequenzunspezifisch unterschieden wird. Obwohl für die *E. coli* RNase HI durch Mutationsexperimente gezeigt wurde, dass die positiv geladenen Reste in der *basic protrusion* wichtig für die Substratbindung sind (Kanaya *et al.*, 1991), sieht man in der Struktur der humanen RNase H1 kaum direkte Kontakte dieser basischen Reste mit dem Substrat (Nowotny *et al.*, 2007). Aufgrund dieser Tatsache postulierte man, dass dieser positiv geladene Bereich erste nicht-spezifische Interaktionen zwischen Substrat und Enzym ermöglicht (Nowotny *et al.*, 2007).

Die erste Kristallstruktur einer isolierten retroviralen RNase H-Domäne mit C-Helix wurde erst 2012 veröffentlicht (Zhou *et al.*, 2012). Da diese Struktur ohne Substrat gelöst wurde, fehlen auch hier die molekularen Details der Substratbindung.

## 2 Ziele der Arbeit

Die RNase H-Domäne ist ein integraler Bestandteil der RT von HIV-1 und der PR-RT von FV. Sie ist bei der reversen Transkription wichtig für den Abbau der RNA in RNA/DNA-Hybriden. Interessanterweise ist die isolierte HIV-1 RNase H-Domäne nicht aktiv, während z.B. die separate MoMLV oder *E. coli* RNase H-Domäne Aktivität zeigen.

Aus Sequenzvergleichen verschiedener RNase H-Domänen geht hervor, dass die RNase H von FV einen bestimmten Sekundärstrukturbereich, die sogenannte C-Helix mit einer sich anschließenden basischen Schleife (*basic protrusion*), besitzt, der in einigen anderen retroviralen RNase H-Domänen, wie z.B. auch der von HIV-1, fehlt. Da es zu Beginn meiner Doktorarbeit keine Struktur einer retroviralen RNase H mit C-Helix gab, lag der zentrale Inhalt meiner Arbeit in der Strukturlösung der freien PFV RNase H-Domäne mittels Kernresonanzspektroskopie (engl. *nuclear magnetic resonance spectroscopy*, NMR-Spektroskopie) und der Identifizierung der *basic protrusion* einschließlich der C-Helix. Anschließend sollte die Funktion der *basic protrusion* bei der Substratbindung durch strukturelle Analyse des Protein/Substratkomplexes mittels NMR-Titrationsexperimenten aufgeklärt werden.

Ein weiteres Ziel war die Identifikation der Domänen- und Subdomänengrenzen der FV PR-RT, um damit zukünftige Strukturanalysen mit unterschiedlichen Domänen zu erleichtern.





### 3 Zusammenfassung und Diskussion der Ergebnisse

#### 3.1 Reinigung und Untersuchung der Aktivität der isolierten PFV RNase H-Domäne

Um strukturelle Untersuchungen mit Hilfe der NMR-Spektroskopie an der isolierten RNase H-Domäne der PFV PR-RT durchführen zu können, wurde die Domäne isotopenmarkiert ( $^{15}\text{N}$ ,  $^{13}\text{C}$ ) gereinigt und im hohen mikromolaren Bereich angereichert. Die Reinheit des Enzyms nach dem Reinigungsprozess konnte durch Natriumdodecylsulfat-Polyacrylamidgelelektrophorese (SDS-PAGE) gezeigt werden (Einzelarbeit A, Abb. 2A). Mit Hilfe der analytischen Gelfiltration wurde nachgewiesen, dass es sich bei der PFV RNase H-Domäne um ein Monomer handelt (Einzelarbeit A, Abb. 2B). Dieses Resultat harmoniert sehr gut mit dem Ergebnis des Gesamtzyms PFV PR-RT, welches ebenfalls als Monomer vorliegt (Hartl *et al.*, 2010a). Außerdem stimmt dieses Ergebnis auch mit dem Oligomerisierungszustand anderer RNase H-Domänen, wie der separat produzierten MoMLV sowie HIV-1 RNase H, *E. coli* RNase HI und humaner RNase H1, überein.

Wie in der Einleitung beschrieben, befindet sich die RNase H-Domäne auf einer Polypeptidkette mit der Polymerase-Domäne. Im Falle der HIV-1 RT führt die Abspaltung der Polymerase-Domäne zum Verlust der Aktivität der isolierten HIV-1 RNase H. Aus diesem Grund war es wichtig, die Aktivität der freien PFV RNase H-Domäne zu überprüfen. Zwei verschiedene Aktivitätstests wurden hierfür verwendet; eine fluoreszenzbasierte, quantitative Aktivitätsmessung und ein qualitativer RNase H-Test zur Analyse des Schnittmusters der freien PFV RNase H-Domäne im Vergleich zum Gesamtzym PR-RT.

Der im fluoreszenzbasierten Aktivitätstest ermittelte  $K_M$ -Wert für die PFV PR-RT ( $5,3 \text{ nM} \pm 0,8 \text{ nM}$ ) lag sehr nahe an dem ermittelten Literaturwert von  $17,1 \text{ nM} \pm 1,2 \text{ nM}$  (Hartl *et al.*, 2010a). Obwohl die isolierte PFV RNase H-Domäne in einer 10-fach höheren Konzentration ( $10 \text{ nM}$ ) verwendet wurde, konnte damit in diesem Versuch keine Sättigung erreicht werden. Dadurch war es trotz des Einsatzes verschiedener graphischer Auswertungsverfahren wie Lineweaver-Burk, Eadie-Hofstee, Hanes oder dem Fit-Programm GraFit 5.0 nicht möglich, für die isolierte RNase H-Domäne eine präzise Michaelis-Menten-Konstante zu bestimmen (Einzelarbeit A, Abb. 3A). Der Test zeigte aber trotzdem, dass die PFV RNase H-Domäne im Gegensatz zur isolierten HIV-1 RNase H-Domäne aktiv ist (Einzelarbeit A, Abb. 3A).

Ähnliche Ergebnisse wurden auch für die isolierte XMRV RNase H-Domäne und für die XMRV RT gefunden (Kim *et al.*, 2012). Auch hier war die unabhängige RNase H-Domäne

nicht ausreichend aktiv, um kinetische Parameter aufzunehmen. Der  $K_M$ -Wert der RT lag mit  $10,4 \text{ nM} \pm 0,5 \text{ nM}$  im Bereich der PFV PR-RT. Es existieren aber auch andere Untersuchungen, die für die separate RNase H-Domäne von XMRV eine hohe Aktivität zeigten (Zhou *et al.*, 2012). Diese Versuche wurden jedoch manganabhängig durchgeführt, während die Ergebnisse von Kim *et al.* (Kim *et al.*, 2012) in Gegenwart von  $\text{Mg}^{2+}$  erhalten wurden. Die MoMLV RT sowie die isolierte RNase H-Domäne zeigten sowohl mit  $\text{Mg}^{2+}$  als auch mit  $\text{Mn}^{2+}$  RNase H-Aktivität. Dabei war die Aktivität mit  $\text{Mn}^{2+}$  höher als mit  $\text{Mg}^{2+}$  (Schultz & Champoux, 1996; Zhan & Crouch, 1997), während Messungen mit *E. coli* RNase HI eine höhere Aktivität mit  $\text{Mg}^{2+}$  als mit  $\text{Mn}^{2+}$  aufwiesen (Zhan & Crouch, 1997; Keck & Marqusee, 1996). Die  $K_M$ -Werte für *E. coli* RNase HI, die magnesiumabhängig ermittelt wurden, befinden sich im Bereich von  $0,05 \text{ }\mu\text{M}$  über  $0,11 \text{ }\mu\text{M}$  bis  $0,53 \text{ }\mu\text{M}$  (Kanaya *et al.*, 1991; Kanaya *et al.*, 1990; Kanaya & Kanaya, 1995). Trotz verschiedener Testbedingungen ist dieser deutliche Unterschied in der Aktivität zwischen PFV und *E. coli* RNase H wahrscheinlich eher darauf zurückzuführen, dass die PFV RNase H normalerweise N-terminal kovalent mit der Polymerase-Domäne verbunden ist, die die Affinität zum Substrat enorm steigert.

Um mehr Informationen bezüglich der enzymatischen Eigenschaften der PFV RNase H-Domäne im Vergleich zur PFV PR-RT zu bekommen, wurde ein qualitativer RNase H-Test durchgeführt, bei dem die RNA am 5'-Ende radioaktiv ( $^{32}\text{P}$ ) markiert war (Einzelarbeit A, Abb. 3B). Für die PFV PR-RT kann man ein für RTs typisches Spaltungsmuster erkennen. Der im Allgemeinen so bezeichnete primäre Schnitt der PFV PR-RT erfolgte in der RNA ca.  $-16$  bis  $-19$  Nukleotide (nt) vom 3'-OH-Ende der DNA entfernt. (Kap. 1.6.2). Anschließend baute die PR-RT zeitabhängig das Substrat weiter ab, bis sie nach 60 min bei  $-8$  nt schnitt. Dieses Spaltungsmuster findet man auch bei der HIV-1 und MoMLV RT (Boyer *et al.*, 2001; Boyer *et al.*, 2004). Es bestätigt bereits veröffentlichte Ergebnisse zur SFVmac und PFV PR-RT (Hartl *et al.*, 2010a; Boyer *et al.*, 2004).

Im Gegensatz zur PFV PR-RT war der Bereich der Spaltungsprodukte der PFV RNase H erheblich größer und erstreckte sich von  $-2$  bis  $-19$  nt. Auch die Anzahl der verschiedenen Spaltungsprodukte war größer. Obwohl die RNase H wenig aktiv ist, führte wahrscheinlich der Überschuss an Enzym schon nach 10 min zu einer Gleichverteilung von verschiedenen Spaltungsprodukten, die dann zeitabhängig weiter abgebaut wurden, sodass zusehends nur noch kleinere Fragmente detektiert werden konnten. Außerdem kann wegen des Einsatzes eines 20 nt langen RNA/DNA-Hybrid-Substrates nicht ausgeschlossen werden, dass mehr als

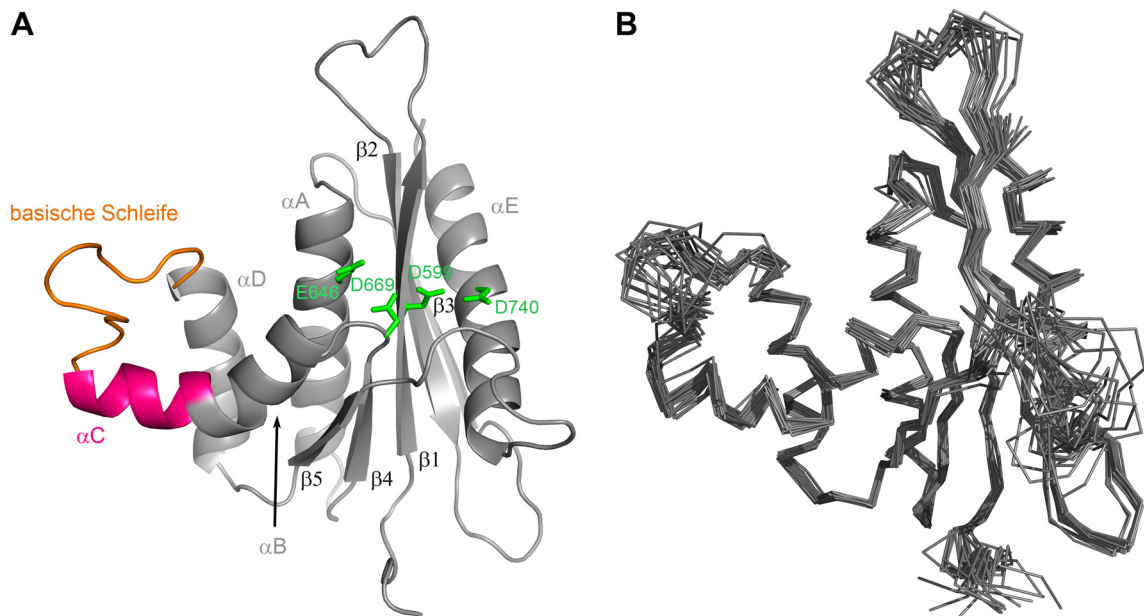
ein Molekül RNase H an das Substrat binden konnte. Dieser Unterschied im Substratabbau zwischen dem Gesamtzym und der isolierten RNase H-Domäne wurde auch schon für die MoMLV RT beobachtet (Schultz & Champoux, 1996; Zhan & Crouch, 1997).

### **3.2 Struktur der isolierten PFV RNase H-Domäne**

Da die isolierte PFV RNase H-Domäne als Monomer vorlag und Aktivität zeigte, obgleich stark vermindert im Vergleich zum Gesamtzym, war sie geeignet, um die Struktur mit NMR-Spektroskopie zu lösen. Da die RNase H  $Mg^{2+}$ -Ionen zur Umsetzung ihres Substrates benötigt, wurden 2D -  $^{15}N$ ,  $^1H$ -*Heteronuclear Single Quantum Coherence* (HSQC) – Experimente ( $[^{15}N, ^1H]$ -HSQC) mit und ohne Zusatz von  $Mg^{2+}$  im Puffer aufgenommen (Einzelarbeit A, Abb. 5). Im  $[^{15}N, ^1H]$ -HSQC-Spektrum zeigt jedes Signal eine Korrelation der Resonanzfrequenz eines Stickstoffatoms mit der Resonanzfrequenz des direkt daran gebundenen Wasserstoffatoms. Die meisten Signale repräsentieren somit einen Aminosäurerest in Form der Proteinrückgrat-N-H-Gruppen. Einen kleinen Teil der Signale erhält man von den N-H-Bindungen der Aminosäureseitenketten.

Obwohl das  $[^{15}N, ^1H]$ -HSQC-Spektrum ohne  $Mg^{2+}$  eine Dispersion von Signalen zeigte, die typisch für ein gefaltetes Protein ist (6,9 - 9,7 ppm), war die Anzahl an Signalen nicht ausreichend für ein Protein, welches aus 165 Aminosäuren besteht. Der Zusatz von  $Mg^{2+}$  erhöhte nicht nur die Anzahl der Signale, sondern auch die Homogenität der Signalintensitäten. Somit ist  $Mg^{2+}$  nicht nur wichtig für die Katalyse, darüber hinaus scheint es die Stabilität der gesamten Struktur zu festigen. Daher erlaubte die Qualität der Spektren der Doppel- und Tripelresonanzexperimente, die in Gegenwart von  $Mg^{2+}$  aufgenommen worden waren, die Strukturaufklärung der PFV RNase H-Domäne.

Die energieärmste Struktur der PFV RNase H ist in Abb. 3.1-A im Bändermodell dargestellt und zeigt eine typische RNase H-Faltung, deren Hauptmerkmal ein fünfsträngiges  $\beta$ -Faltblatt ist, das aus zwei parallelen und drei antiparallelen  $\beta$ -Strängen besteht. Dieses  $\beta$ -Faltblatt wird von der einen Seite von einer langen  $\alpha$ -Helix ( $\alpha$ -Helix E) und von der anderen Seite von vier  $\alpha$ -Helices ( $\alpha$ -Helix A-D) flankiert (Abb. 3.1-A).



**Abb. 3.1: Strukturlösung der PFV RNase H mittels NMR-Spektroskopie.** (A) Bänderdarstellung der energieärmsten Struktur der PFV RNase H-Domäne. Die C-Helix ist pink und die basische Schleife ist orange dargestellt. Die Reste des aktiven Zentrums (D599, E646, D669, D740) sind grün hervorgehoben. Die einzelnen  $\alpha$ -Helices ( $\alpha$ A- $\alpha$ E) und  $\beta$ -Faltblattstränge ( $\beta$ 1- $\beta$ 5) sind beschriftet. (B) Überlagerung der 19 energieärmsten Strukturen der PFV RNase H.

Damit bestätigte sich die Vermutung, dass die PFV RNase H-Domäne eine C-Helix mit anschließender basischer Schleife besitzt. Der Überlagerung der 19 energieärmsten Strukturen (Abb. 3.1-B) kann man entnehmen, dass das Rückgrat der Sekundärstrukturbereiche in allen Strukturen nur wenig voneinander abweicht ( $1,05 \pm 0,46 \text{ \AA}$ , Einzelarbeit B, Tabelle 1). Die gute räumliche Definition dieser Bereiche ist möglich, da es für sie genügend räumliche Abstandsinformationen aus den *Nuclear Overhauser Effect Spectroscopy* (NOESY)-Experimenten gab. Für einige Aminosäuren (D607, N639, N692-K695, Q725-N728, S730, E734) konnten keine Resonanzfrequenzen des Rückgratstickstoffes mit zugehörigem Wasserstoff bestimmt werden. Außerdem fehlen bei einigen der Aminosäuren nicht nur die Resonanzfrequenzen des Rückgrates, sondern auch die der Seitenketten. Durch das Fehlen dieser Resonanzfrequenzen gehen Abstandsinformationen zu räumlich nahe gelegenen Atomen verloren, sodass diese Bereiche in der Struktur nicht fixiert werden konnten. Aus diesem Grund überlagern die 19 energieärmsten Strukturen in diesen Bereichen schlechter (Abb. 3.1-B).

Die Struktur der PFV RNase H passt sehr gut zu den Strukturen anderer RNase H-Domänen, die mit und ohne *basic protrusion* gelöst werden konnten, wie z.B. die XMRV (pdb: 3V1O), die humane (pdb: 2QK9), die *B. halodurans* (pdb: 1ZBF) und die HIV-1 RNase H (pdb: 1HRH), deren Proteinrückgrat mit rmsd-Werten von  $2,10 \text{ \AA}$ ,  $2,35 \text{ \AA}$ ,  $2,37 \text{ \AA}$  und  $1,90 \text{ \AA}$  zur

PFV RNase H überlagert (Einzelarbeit B, Abb. 6). Dabei zeigt der Strukturvergleich der PFV RNase H mit der XMRV und der humanen RNase H eine gute Überlagerung der Sekundärstrukturen, einschließlich der C-Helix. Unterschiede ergeben sich in den flexiblen Schleifenregionen zwischen den Sekundärelementen. Auffällig ist, dass sich in der PFV RNase H, im Gegensatz zur XMRV, MoMLV und auch zur *E. coli* RNase H-Domäne, drei hintereinander positionierte positiv geladene Reste in der basischen Schleife, anstatt in der C-Helix, befinden (Einzelarbeit B, Abb. 1 und Abb. 6). Dies könnte Auswirkungen auf die Bindungseigenschaften dieser Region zum Substrat haben. Ein anderer interessanter Fall ist die RNase H-Struktur von *B. halodurans*. Sowohl die RNase HI-Domäne von *B. halodurans* als auch eukaryotische RNase H1-Domänen besitzen eine N-terminale HBD zur besseren Substratbindung. Allerdings ist die RNase H von *B. halodurans* auch ohne diese N-terminale Domäne aktiv, obwohl sie weder eine vollständige E-Helix noch eine C-Helix aufweist. Es ergibt sich jedoch für den Bereich der C-Helix bei *B. halodurans* ein Schleifenbereich, der 4 basische Reste besitzt und vermutlich durch seine Länge in der Lage ist, die C-Helix räumlich zu ersetzen (Einzelarbeit B, Abb. 1 und Abb. 6).

Eine Besonderheit der PFV RNase H, die bei allen anderen RNase H-Domänen nicht auftritt, ist das Vorkommen zweier Cysteinreste, die in der Struktur so positioniert sind, dass die Ausbildung einer Disulfidbrücke möglich wäre (Einzelarbeit B, Abb. 3A). Da die PFV RNase H unter reduzierenden Bedingungen produziert und gereinigt wurde und auch die Resonanzfrequenz der C $\beta$ -Atome hochfeldverschoben sind, liegt eine Struktur mit reduzierten Cysteinen vor. Darüber hinaus wurden alle Aktivitätstests mit dem Reduktionsmittel Dithiothreitol durchgeführt. Um auszuschließen, dass die gemessene Aktivität von einem kleinen Anteil an RNase H herrührte, dessen Cysteinreste oxidiert vorlagen, wurde ein Cysteinrest in Serin umgewandelt (C654S) und die Aktivität sowie das Spaltungsmuster mit einem qualitativen RNase H-Test überprüft (Einzelarbeit B, Abb. 3B). Da jedoch zwischen RNase H wt und der RNase H-(C654S) Variante in dem Test kein Unterschied zu entdecken war, scheint die reduzierte Form auch die aktive Form zu sein. Ein weiterer Punkt, der diese Annahme unterstützt, ist die Tatsache, dass in der Aminosäuresequenz anderer RNase H-Domänen von FV (außer SFVmac) diese Cysteine nicht konserviert und daher vermutlich nicht für die Funktion des Enzyms ausschlaggebend sind. Was zu diesem Zeitpunkt nicht ausgeschlossen werden kann, ist, ob die PFV RNase H mit oxidierten Cysteinresten aktiver wäre als mit reduzierten Cysteinresten. Da man jedoch in der eukaryotischen Zielzelle von

reduzierenden Bedingungen ausgeht, sollte eine derartige Cysteinverbrückung keine biologische Relevanz besitzen.

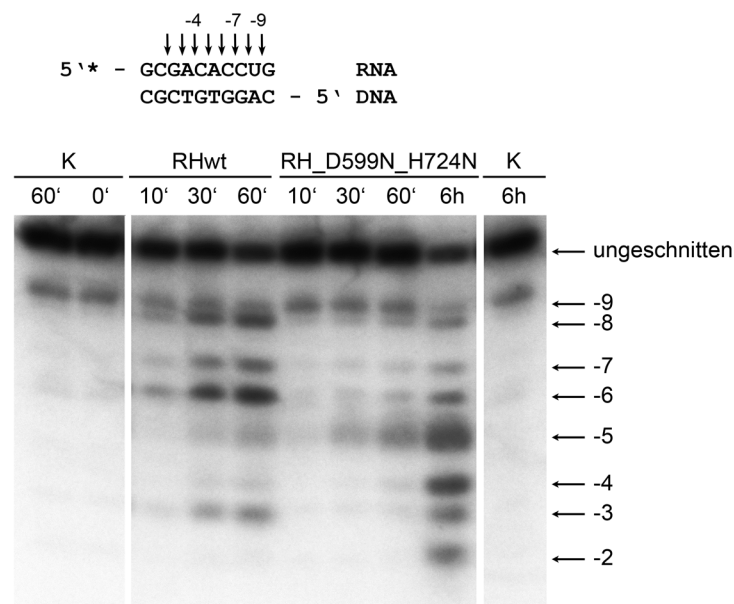
### **3.3 Aktivität der isolierten PFV RNase H-Domäne als D599N-H724N Variante**

Da die Struktur der PFV RNase H geklärt werden konnte und diese gut mit schon bekannten RNase H-Strukturen übereinstimmt, sollte nun der Interaktionsbereich mit dem Substrat mittels NMR-Spektroskopie untersucht werden. Bei dem dazugehörigen Versuchsaufbau wurde isotoptenmarkiertes Enzym im mikromolaren Bereich vorgelegt und in bestimmten molaren Verhältnissen Substrat zugegeben. Um sicher zu gehen, dass in späteren Bindungsstudien nur ein Molekül Enzym an ein Molekül Substrat bindet, wurde anhand von Aktivitätsmessungen mit der RNase HI von *E. coli* und der humanen RNase H1 (Nowotny *et al.*, 2007; Kanaya & Kanaya, 1995) das kleinstmögliche RNA/DNA-Hybrid, welches aus 10 bp besteht, gewählt.

Obwohl die PFV RNase H als isolierte Domäne eine niedrige Aktivität zeigte, war bei Einsatz des Enzyms im mikromolaren Konzentrationsbereich bei den zeitintensiven NMR-Messungen ein merklicher Abbau des Substrates zu erwarten. Dadurch können die theoretischen Änderungen der molaren Verhältnisse bei Zugabe von Substrat im praktischen Versuch nicht eingehalten werden. Aus diesem Grund wurden in der separaten PFV RNase H-Domäne (RNase H wt) zwei Aminosäurereste, Aspartat 599 und Histidin 724, gegen Asparagin ausgetauscht (RNase H-(D599N-H724N)). Aus dem Sequenzvergleich (Einzelarbeit A, Abb. 1) der Aminosäuresequenzen der PFV RNase H mit den RNase H-Domänen von HIV-1, MoMLV, *E. coli* sowie humaner RNase H1, geht hervor, dass es sich bei Aminosäure D599 um einen Rest des aktiven Zentrums handelt und bei H724 um einen Rest, der in vielen Typ 1 RNase H-Domänen konserviert ist (H264 in der humanen RNase H1, H124 in der *E. coli* RNase HI, H539 in der HIV-1 RT). Die bereits bekannten Strukturen zeigen eine Lokalisation dieses Histidin-Rests in einer flexiblen Schleife. Ein Austausch von H539 in der HIV-1 RT gegen Phenylalanin (Schatz *et al.*, 1989), Asparagin (Tisdale *et al.*, 1991; Wöhrle *et al.*, 1991) oder Aspartat (Tisdale *et al.*, 1991; Wöhrle *et al.*, 1991) beeinflusste die RNase H Aktivität stark.

Durch einen Austausch von D599 und H724 gegen Asparagin sollte die Aktivität der PFV RNase H inhibiert werden, die strukturelle Integrität des Enzyms aber erhalten bleiben. Die RNase H-(D599N-H724N) sowie die RNase H wt wurden in einem qualitativen RNase H-Test in ihrer Spaltungsaktivität verglichen. Abb. 3.2 zeigt bei der RNase H-(D599N-H724N)

nach 60 min merklich weniger Substratabbau als bei der RNase H wt. Darüber hinaus weist die Doppelmutante ein anderes Schnittmuster als das Wildtypprotein auf, was besonders bei der RNase H-(D599N-H724N) nach 6 h Inkubation deutlich wird. Außerdem war selbst nach 6 h das RNA/DNA-Hybrid nicht vollständig abgebaut, obwohl das Enzym in 30-fachem Überschuss im Vergleich zum Substrat vorlag. Wie schon oben beschrieben, liegt bei der NMR-Titration eine andere Situation vor. In diesem Experiment wurde zu einer konstanten Enzymkonzentration Substrat in steigenden molaren Verhältnissen zugeben, d.h. das Substrat lag vorwiegend im Überschuss vor. Da das gesamte NMR-Titrationsexperiment nicht länger als 6 h dauerte, kann man davon ausgehen, dass das Substrat während der Messung nur unerheblich von der RNase H-(D599N-H724N) abgebaut wurde.



**Abb. 3.2: Qualitativer RNase H-Test der RNase H wt und der RNase H-(D599N-H724N).** Die Sequenz des Substrathybrides befindet sich oberhalb des Gels, die jeweiligen Schnittstellen sind mit Pfeilen dargestellt. Das erste RNA-Nukleotid, das am 3'-OH-Ende der DNA hybridisiert vorliegt, wird mit -1 gekennzeichnet. Die RNA wurde am 5'-Ende mit  $^{32}\text{P}$  markiert. Die Spaltungsprodukte wurden auf einem 15 %igen Polyacrylamid-/8M Urea-Gel getrennt und mit einem Phosphoimager visualisiert. Die Pfeile auf der rechten Seite des Gels zeigen auf die Spaltungsprodukte. 670 nM 10/10mer RNA/DNA-Hybrid-Substrat wurden mit 20  $\mu\text{M}$  RNase H wt (RHwt) und RNase H-(D599N-H724N) (RH\_D599N\_H724N) bei 25 °C inkubiert und für die angegebenen Zeitpunkte Proben entnommen. K: Substrat ohne Enzym.

Durch den Austausch der katalytisch wichtigen Reste D599 und H724 gegen Asparagin sollte die Spaltungsaktivität stark reduziert, die Substrataffinität aber nicht beeinflusst werden. Um das zu überprüfen, wurden Fluoreszenzanisotropie-Gleichgewichtstitrationen durchgeführt. Für die RNase H wt und RNase H-(D599N-H724N) wurden  $K_D$ -Werte von  $23 \mu\text{M} \pm 3 \mu\text{M}$  und  $31 \mu\text{M} \pm 5 \mu\text{M}$  bestimmt (Einzelarbeit B, Abb. 4B). Aufgrund der vergleichbaren Bindungsaffinitäten war die Doppelmutante für die Bindungsstudien mittels NMR-Spektroskopie mit dem 10/10-mer RNA/DNA-Hybrid geeignet.

### 3.4 Substratbindungsstudien der isolierten PFV RNase H-(D599N-H724N) mit Hilfe von NMR-Spektroskopie

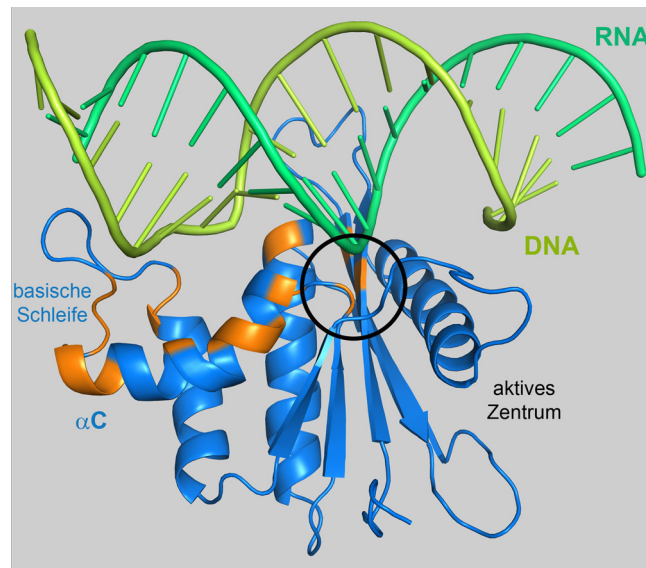
Die Auswertung von [ $^{15}\text{N}$ ,  $^1\text{H}$ ]-HSQC- und [ $^1\text{H}$ ,  $^{15}\text{N}$ ,  $^1\text{H}$ ]-NOESY-HSQC-Spektren der PFV RNase H-(D599N-H724N) zeigte, dass die Tertiärstruktur der RNase H-(D599N-H724N) der des Wildtypproteins entspricht. Zur Untersuchung der Substratbindungsbereiche des Enzyms wurden 50  $\mu\text{M}$  PFV RNase H-(D599N-H724N)  $^{15}\text{N}$ -isotopenmarkiert vorgelegt und das 10/10-mer RNA/DNA Substrat im Verhältnis 1:1, 1:2, 1:3, 1:4 und 1:5 zugegeben. Nach jeder Zugabe wurde ein [ $^{15}\text{N}$ ,  $^1\text{H}$ ]-HSQC-Spektrum aufgenommen. Da nur bei einigen Aminosäureresten eine Intensitätsabnahme der Signale festgestellt wurde, ist davon auszugehen, dass die Gesamtstruktur des Enzyms bei der Substratzugabe erhalten blieb. Kommt es bei Substratbindungsstudien in NMR-Titrationsmessungen zu einer Intensitätsabnahme einzelner Signale, so befinden sich diese Reste in einem sogenannten intermediären Austausch, der zwischen dem schnellen und langsamen Austausch liegt (Roberts, 2003). Beim intermediären Austausch findet der Wechsel zwischen dem freien und substrat-gebundenen Zustand des Enzyms in einem Zeitbereich statt, bei dem im [ $^{15}\text{N}$ ,  $^1\text{H}$ ]-HSQC-Spektrum für die betroffenen Signale eine Linienverbreiterung und damit eine Abnahme der Signalintensität beobachtet werden kann. Da im Fall der PFV RNase H-(D599N-H724N) aufgrund der Substrat-Enzym-Komplexbildung für die Aminosäuren G600, A614, T641, Q643, D669, F671, A674, E675, S676, E680, K685, S686, N687, F689, V690, N691, H699, I700 eine signifikante Abnahme der Signalintensität im [ $^{15}\text{N}$ ,  $^1\text{H}$ ]-HSQC-Spektrum gemessen wurde, kann man von einer Beteiligung dieser Reste bei der Substratbindung ausgehen. Als signifikant wurden alle Reste angesehen, die weniger als 50 % der ursprünglichen Signalintensität des freien Enzyms ohne Substrat erreichten. Markiert man diese Reste in der PFV RNase H-(D599N-H724N) Struktur, zeigt sich, dass sie um ganz bestimmte Bereiche lokalisiert sind, die das aktive Zentrum, die B-, C- und D-Helix und die basische Schleife (Abb. 3.3) betreffen.

Da aus den [ $^{15}\text{N}$ ,  $^1\text{H}$ ]-HSQC-Spektren nur Informationen bezüglich des Proteins, jedoch nicht bezüglich des Substrates gewonnen werden können, wurde ein einfaches Modell zum Enzym-Substrat-Komplex erstellt, um zu überprüfen, ob sich die betreffenden Reste in der Nähe des Substrates befinden. Aufgrund der strukturellen Ähnlichkeit der PFV RNase H mit der humanen RNase H1, die in Komplex mit einem RNA/DNA-Hybrid gelöst werden konnte (pdb: 2QK9), wurden die Strukturen übereinander gelegt und zur besseren Übersicht die humane RNase H1 aus dieser Überlagerung wieder ausgeblendet (Abb. 3.3). Die betroffenen



Reste der PFV RNase H befinden sich auf der substratzugewandten Seite, was man auch gut im Modell nachvollziehen kann.

Da die Signale jener Reste, die das aktive Zentrum betreffen bzw. die sich in unmittelbarer Nähe davon befinden, verschwinden, und die RNase H zwischen dsRNA, dsDNA und RNA/DNA-Hybriden diskriminiert, ist eine Substratbindung wie in Abb. 3.3 gezeigt wahrscheinlich. Interessanterweise zeigten auch Signale von Resten in der B-Helix eine signifikante Intensitätsabnahme bei den NMR-Titrations. Die äquivalenten Reste in der Kristallstruktur der humanen RNase H1 sind räumlich zu weit entfernt, um mit dem Substrat zu interagieren. Da die NMR-Titrations in Lösung stattfinden und damit einen dynamischen Charakter besitzen, könnte die B-Helix in der PFV RNase H eine Funktion bei der Substratbindung haben, die in der Kristallstruktur der humanen RNase H1 mit Substrat nicht zu erkennen ist.

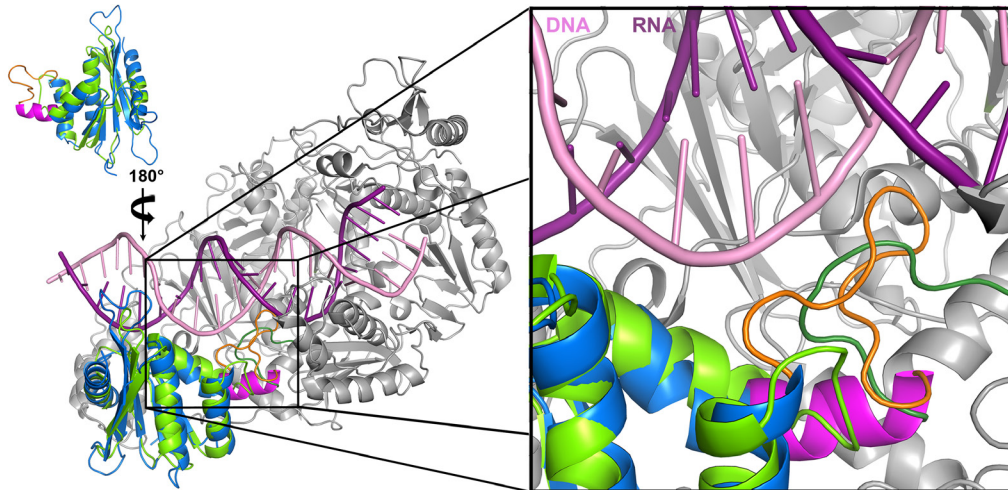


**Abb. 3.3: Modell eines Enzym-Substrat-Komplexes der PFV RNase H mit einem RNA/DNA-Hybrid Substrat.** Bänderdarstellung der PFV RNase H. Die Reste, die während der NMR-Titration bei Substratzugabe signifikant an Signalintensität verloren haben, sind orange markiert. Mittels einer Strukturüberlagerung der humanen RNase H1 (2QK9) mit der PFV RNase H (rmsd: 2,35 Å) konnte ein unspezifisches RNA/DNA-Hybrid-Substrat zur PFV RNase H modelliert werden. Die Struktur der humanen RNase H1 wurde zur besseren Übersicht nicht dargestellt. Der DNA- bzw. RNA-Strang sind hell- und dunkelgrün dargestellt. Der Bereich des aktiven Zentrums wird von einem schwarzen Kreis eingerahmt.

Die Reste K685, S686, N687, die sich am C-Terminus der C-Helix befinden, sowie F689, V690, N691, H699, die in der basischen Schleife lokalisiert sind und I700 (N-terminaler Beginn D-Helix) zeigen nicht nur in den NMR-Titrations Effekte (Abb. 3.3), sondern sind auch sehr gut für das umgebende Lösungsmittel zugänglich, wodurch unspezifische Kontakte zum Substrat zu Stande kommen könnten. Obwohl die Zuordnung einiger Resten der basischen Schleife der PFV RNase H nicht möglich war, sind die beiden Enden dieser Schleife zugeordnet und in die Substratbindung involviert. Aufgrund ihrer Flexibilität ist es

jedoch wahrscheinlich, dass auch der mittlere Teil dieser Schleifenregion an der Substratbindung beteiligt ist. Zusätzlich besitzt die PFV RNase H an dieser Stelle vier basische Reste, die gut lösungsmittelzugänglich sind und von denen drei direkt hintereinander positioniert sind (Einzelarbeit B, Abb. 1 und Abb. 6), was eine Interaktion mit dem Substrat sehr wahrscheinlich macht. Somit wirkt die C-Helix der PFV RNase H wie ein Lineal, das die basische Schleife zum Substrat orientiert und auch für die Substrataffinität des Enzyms unverzichtbar ist. Dieser Zusammenhang erklärt eventuell auch, warum die freie HIV-1 RNase H bzw. auch die freie MoMLV RH $\Delta$ C keine Aktivität zeigten. Ohne C-Helix fehlt der basischen Schleife eine geeignete Orientierung, die auch raumgreifend genug ist, um das Substrat optimal zu binden. Im Fall der HIV-1 RNase H ist die Schleifenregion nicht nur zu kurz, um räumlich das Substrat binden zu können; sie besitzt wahrscheinlich auch nicht genügend basische Reste für eine ausreichende Substrataffinität.

Vergleicht man die Struktur der HIV-1 RNase H und der PFV RNase H-Domäne im Kontext zur HIV-1 RT, deren Kristallstruktur im Komplex mit einem PPT-Hybrid-Substrat gelöst wurde (pdb: 1HYS), wird der Bereich der *basic protrusion* der PFV RNase H in der HIV-1 RT durch eine Schleifenregion aus der Verbindungs-Subdomäne der p66-Untereinheit ersetzt (Abb. 3.4). Dabei handelt es sich um die Reste R356-N363 (dunkelgrüne Schleife in Abb. 3.4) der p66-Untereinheit der HIV-1 RT. Ähnlich zu der basischen Schleife der *basic protrusion* befinden sich auch in der Schleife der Verbindungs-Subdomäne basische Reste (R356, R358). Die Bedeutung dieser Schleifenregion für die RNase H Aktivität konnte bereits durch die Mutation des Restes H361 gezeigt werden (Julias *et al.*, 2003). Da die HIV-1 RT ein Heterodimer bestehend aus einer p66- und einer p51-Untereinheit ist, gibt es neben den oben genannten Aminosäuren der Schleifenregion noch weitere Reste, die für die optimale Positionierung des Substrates und damit für die RNase H Aktivität wichtig sind. Sie werden unter dem Begriff RNase H *primer grip* (Kap. 1.6.2) zusammengefasst. Da die HIV-1 RNase H keine *basic protrusion* besitzt, außerdem in einer heterodimeren Gesamtstruktur liegt und das Substrat das aktive Zentrum der RNase H um ca. 2-4 Å verfehlt (Nowotny *et al.*, 2007; Lapkouski *et al.*, 2013), ergeben sich zwischen der HIV-1 und der PFV RNase H kaum gemeinsame Interaktionsbereiche mit dem Substrat.



**Abb. 3.4: Strukturüberlagerung der PFV und HIV-1 RNase H-Domänen mit der HIV-1 RT.** *Oben links:* Strukturüberlagerung der PFV (blau; pdb: 2LSN) und HIV-1 (grün; pdb: 1HYS) RNase H-Domänen. Die C-Helix und die basische Schleife der PFV RNase H sind pink und orange dargestellt. *Unten links:* Die Strukturüberlagerung der PFV und HIV-1 RNase H-Domänen von *oben links* befindet sich um 180° gedreht in der HIV-1 RT, die im Komplex mit dem PPT-Substrat (pdb: 1HYS) vorliegt. *Rechts:* Vergrößerung des Bereiches in der HIV-1 RT, in der sich die C-Helix und die basische Schleife der PFV RNase H befinden. An dieser Stelle ragt eine Schleifenregion (dunkelgrün) der Verbindungs-Subdomäne in der p66-Untereinheit der HIV-1 RT zum Substrat. Die RNA ist in violett und die DNA in rosa dargestellt.

Leider existiert bisher keine Struktur einer monomeren RT mit ausreichender Elektronendichte für die RNase H-Domäne, die mit Substrat komplexiert vorliegt, um dahingehend Vergleiche anzustellen. Es ist aber zu vermuten, dass die RNase H in der MoMLV RT wesentlich freier in ihrer Bewegung ist als die HIV-1 RNase H im RT-Heterodimer, da sie in der Kristallstruktur der MoMLV RT nicht aufgelöst werden konnte und die Verbindungs-Subdomäne einen längeren Schleifenbereich zur RNase H-Domäne aufweist als die HIV-1 RT (Das & Georgiadis, 2004; Cote & Roth, 2008). Trotz dieser Flexibilität der MoMLV RNase H scheint das Zusammenspiel zwischen der Verbindungs-Subdomäne und der RNase H abgestimmt zu sein, da Insertionen von 5 zusätzlichen Resten in den Schleifenregionen der Verbindungs-Subdomäne zu nicht lebensfähigen Viren führte (Puglia *et al.*, 2006). Da die RNase H-Domäne der MoMLV RT jedoch nicht so kompakt in die RT eingebettet ist wie die HIV-1 RNase H, scheint die *basic protrusion* in der MoMLV RNase H für die notwendige Substrataffinität erforderlich zu sein. Die Spezifität und eine Steigerung der Affinität zum Substrat erfahren alle retroviralen RNase H-Domänen durch die Polymerase-Domäne.

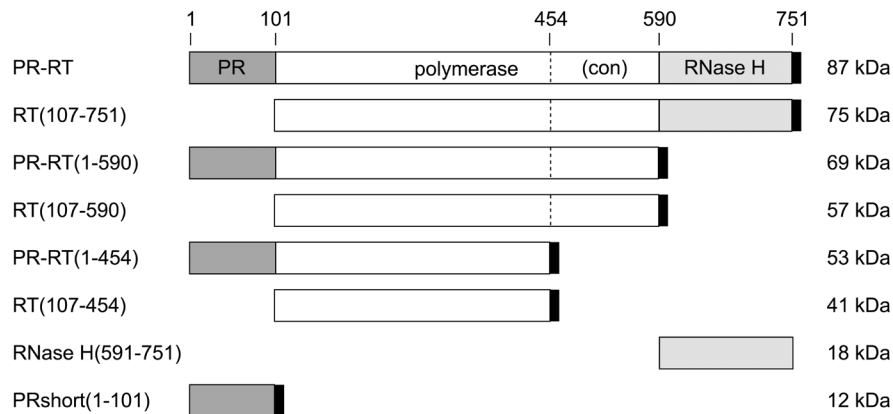
Da das Zusammenspiel zwischen RNase H-Domäne und der Verbindungs-Subdomäne in der HIV-1 RT unabdingbar für die RNase H Aktivität ist, könnte dort auch ein weiterer Angriffspunkt für die antiretrovirale Therapie liegen, der bisher vernachlässigt worden ist. Gegenwärtig zielen alle in der Therapie eingesetzten RT-Inhibitoren (RTI) auf die Polymerase-Domäne und bisherige genotypische Analysen erfassen nur etwa die ersten 300

Aminosäuren und können daher nur in diesem Bereich resistenz-assoziierte Mutationen identifizieren (Delviks-Frankenberry *et al.*, 2010). Dabei stellte sich in Untersuchungen heraus, dass auch Mutationen in der Verbindungs-Subdomäne und vereinzelt in der RNase H-Domäne der HIV-1 RT die Resistenz der RT gegen nukleosidische und nicht-nukleosidische RTI verstärken können. Die Verstärkung beruht in den meisten Fällen auf einer Verminderung der RNase H-Aktivität durch die Mutationen in den (Sub-)Domänen. Dadurch ergibt sich für die RT mehr Zeit, das Nukleotid-Analogon auszubauen bzw. von nicht-nukleosidischen RTI zu dissoziieren, bevor die RNase H die RNA-Matrize abbauen kann und sich der Komplex vom Substrat löst (Delviks-Frankenberry *et al.*, 2010). Bisher gibt es für die Verbindungs-Subdomäne und die RNase H-Domäne selbst keine Inhibitoren, die in der Klinik angewendet werden. RNase H-Inhibitoren, die auf das aktive Zentrum der RNase H abzielen, bergen auch die Gefahr, die humane RNase H1 zu hemmen, was ernste Folgen bezüglich der DNA-Replikation haben könnte (Kap. 1.7). Daher müsste man bei der Entwicklung von Inhibitoren der Verbindungs-Subdomäne und der RNase H-Domäne der RT auf eine spezifische und hochaffine Bindung sowie vollständige Inhibition der retroviralen RNase H achten, um nicht die Resistenzentwicklung der RT gegen nukleosidische und nicht-nukleosidische RTI zu fördern. Da eine 100%ige Inhibition der RNase H kaum zu realisieren ist, müssten diese Inhibitoren z.B. den spezifischen RNase H-Schnitt am PPT unterdrücken, da dadurch keine doppelsträngige DNA mehr hergestellt werden kann (Kap. 1.5). Somit könnte der Bereich in der HIV-1 RT aus der Verbindungs-Subdomäne, der für die Substratbindung relevant ist, ein wichtiger Anhaltspunkt für eine Inhibitorbindestelle sein.

### **3.5 Die Domäneneinteilung der FV PR-RT**

Bei HIV-1 und MoMLV besteht die RT aus einer Polymerase- und einer RNase H-Domäne. Während die HIV-1 RT ein Heterodimer aus den Untereinheiten p51 und p66 ist, liegt die MoMLV RT als Monomer vor. Bei FV kommt zu der Polymerase- und RNase H-Domäne eine N-terminale Protease-Domäne (PR) dazu (PR-RT), die nicht während der Virusreifung abgespalten wird (Hartl *et al.*, 2010a); eine Tatsache, die in einer früheren Publikation noch nicht erkannt wurde (Pfrepper *et al.*, 1998). Die FV PR-RT liegt ebenfalls als Monomer vor (Hartl *et al.*, 2010a). Unklar war bisher, wo genau sich die Grenzen der PR, der Polymerase- und RNase H-Domäne befinden. Ungenaue Domänenabgrenzungen können bei der Strukturaufklärung von einzelnen Domänen hinderlich sein, da die Tertiärstruktur beeinträchtigt sein kann. Dies führt auch zu Funktions- und Löslichkeitsproblemen dieser

Domänen. Eine genaue Domänenkenntnis unterstützt demnach die Planung und Interpretation von *in vivo*- und *in vitro*-Experimenten. Für die PFV RNase H (Q591-N751) und die SFV PR (M1-T101) aus Makaken (SFVmac PRshort) konnte die Domänengrenze biochemisch mit Hilfe von Aktivitätstest *in vitro* identifiziert werden, da diese separat produzierten Domänen Aktivität zeigten, während weitere Deletionen zu inaktiven Enzymen führten (Einzelarbeit A) (Hartl *et al.*, 2008). Mit Hilfe dieser Daten und unter Berücksichtigung früherer Publikationen wurden verschiedene SFVmac PR-RT Deletionsvarianten konstruiert (Abb. 3.5) und deren Enzymaktivitäten getestet. Um herauszufinden, ob die Deletion der PR-Domäne Auswirkungen auf die RNase H Aktivität hat, weil z.B. die Struktur des verbleibenden Proteins stark verändert wird, wurden für die SFVmac PR-RT wt, RT(107-751) und die RNase H (E591-H751) ein qualitativer RNase H-Test durchgeführt (Einzelarbeit C, Abb. 2). Dabei zeigten die PR-RT wt und die RT(107-751) nicht nur das gleiche Spaltungsmuster, sondern auch die gleiche Aktivität. Das bedeutet, dass die PR-Domäne keinen Einfluss auf die RNase H-Aktivität sowie -Struktur hat und, wie zu erwarten, auch für die Substratbindung und Substrataffinität nicht notwendig ist. Wie auch schon bei der PFV RNase H (Einzelarbeit A und B) zeigte die freie SFVmac RNase H-Domäne Aktivität; sie musste jedoch in sehr hoher Konzentration eingesetzt werden. Darüber hinaus waren die Schnitte nicht, wie bei der PR-RT wt und RT(107-751), vom 3'-OH-Ende der DNA bestimmt, sondern gleichmäßig auf das Substrat verteilt, weil wahrscheinlich mehrere Enzymmoleküle an das Substrat binden.



**Abb. 3.5: Schematische Übersicht über die verschiedenen SFVmac PR-RT Konstrukte.** Die Protease- (PR; dunkelgrau), Polymerase- (weiß) und RNase H-Domäne (hellgrau) sind schematisch für jede Deletionsvariante dargestellt und mittels durchgezogener Linie voneinander getrennt. Die Verbindungs-Subdomäne (con) ist durch eine gestrichelte Linie von der Polymerase-Domäne abgegrenzt. Die Bezeichnung der einzelnen Deletionsvarianten ist links und die molekulare Masse rechts zu entnehmen. Oberhalb der Konstrukte stehen die Sequenzpositionen der einzelnen (Sub-)Domänen. Die schwarzen Rechtecke stellen den C-terminalen 6xHis-Reinigungsanhang dar.

Zur Identifizierung des Anfangs der Polymerase-Domäne wurden zwei Konstrukte getestet. Der Startpunkt des ersten Proteins lag bei Q144, da in einer früheren Publikation diese

Position als Beginn der Polymerase-Domäne postuliert wurde (Pfrepper *et al.*, 1998) und H107, kurz hinter dem Ende der PR-Domäne (Hartl *et al.*, 2008; Hartl *et al.*, 2010b). Da die SFVmac PR-RT Variante mit dem Startpunkt Q144 im Gegensatz zur RT(107-751) keine Polymerisationsaktivität mehr zeigte, scheint die Region H107-N143 ganz oder teilweise wichtig für die Integrität der Polymerase-Domäne zu sein. Obwohl die MoMLV RT im Vergleich zur HIV-1 RT eine N-terminale Verlängerung von 40 Aminosäuren aufweist (Cote & Roth, 2008), ist diese nicht homolog zur Region H107-N143 der FV PR-RT. Das ist auch nicht unerwartet, da die MoMLV RT im Gegensatz zur FV PR-RT keine N-terminale PR-Domäne aufweist.

Um den Anfang der Verbindungs-Subdomäne zu definieren, wurden weitere Deletionskonstrukte hergestellt. Dabei erwiesen sich die Proteine PR-RT(1-454) und RT(107-454) als löslich, während PR-RT(1-449) und RT(107-449) unlöslich waren. Damit konnte der Beginn der Verbindungs-Subdomäne auf den Aminosäurerest E455 festgelegt werden. Weiterhin sollte untersucht werden, inwiefern Deletionen der PR, der RNase H und der Verbindungs-Subdomäne die Bindungseigenschaften und Polymerisationsfähigkeit der Polymerase-Domäne beeinflussen.

$K_D$ -Bestimmungen durch Fluoreszenzanisotropie-Gleichgewichtstitrationen (Einzelarbeit C, Tabelle 1) ergaben für die PR-RT wt und die RT(107-751) Werte von  $6 \mu\text{M} \pm 1 \mu\text{M}$  und  $3 \mu\text{M} \pm 1 \mu\text{M}$ . Somit hat die PR-Domäne keine direkte oder indirekte Auswirkung auf die Nukleinsäurebindung der RT. Wurde hingegen die RNase H-Domäne entfernt, verschlechterte sich die Affinität zum Substrat um 90 %. Deletierte man zusätzlich noch die Verbindungs-Subdomäne, verlor man nochmals an Affinität. Dieser Einfluss der Verbindungs-Subdomäne spiegelt sich auch in den Polymerisationstests wider (Einzelarbeit C, Abb. 3 und 4). SFVmac Deletionsvarianten ohne RNase H und ohne Verbindungs-Subdomäne zeigten nur in hoher Konzentration Polymerisationsprodukte. Alle anderen RT-Varianten polymerisierten auch bei niedrigen Enzymkonzentrationen, wobei keine Variante an die Aktivität der PR-RT wt heranreichte. Dies zeigt zum einen, dass jede Deletion der PR-RT die strukturelle Integrität der Polymerase-Domäne beeinflusst und zum anderen, dass die Verbindungs-Subdomäne nicht nur ein Bindeglied zwischen der Polymerase- und RNase H-Domäne ist, sondern einen relevanten Teil der Polymerase-Domäne darstellt, der sehr wichtig für die Substratbindung und Polymeraseaktivität ist. Weiterhin haben Untersuchungen zur thermischen Stabilität mit Hilfe der  $1\text{D-}^1\text{H-NMR}$ -Spektroskopie erkennen lassen, dass die Verbindungs-Subdomäne die Integrität der RT festigt (Einzelarbeit C, Abb. 5). Die RT(107-454) verliert ihre Struktur bei einer niedrigeren Temperatur als die RT(107-590). Dies kann an einer starken Signalabnahme

der RT(107-454) bereits bei 35°C erkannt werden, während der gleiche Effekt bei der RT(107-590) erst bei 40°C auftritt (Einzelarbeit C, Abb. 5C und 5D).

Im Zuge der Identifizierung der Domänengrenzen sollte auch geklärt werden, welche Domänen wichtig für die Aktivierung der PR der PR-RT sind. Die PR muss zur Aktivierung dimerisieren. Sowohl die SFVmac PRshort als auch die PR-RT liegen als Monomer vor. Eine Dimerisierung kann entweder durch 2-3 M NaCl oder im Falle der PR-RT auch durch eine bestimmte Region der genomischen RNA (*protease activating RNA motif*, PARM) hervorgerufen werden, die sich im Genbereich von *pol* befindet (Hartl *et al.*, 2011). PRshort kann durch PARM und Niedrigsalzbedingungen nicht aktiviert werden (Einzelarbeit C, Abb. 6C). Da die PR der RT-Deletionsvarianten PR-RT(1-591) und PR-RT(1-454) auch nicht durch PARM aktiviert werden konnten (Einzelarbeit C, Abb. 6C), sind die Verbindungs-Subdomäne und die RNase H wichtig für die Bindung der PR-RT an PARM und somit für die PR-Dimerisierung. Aufgrund bisheriger Literaturdaten und unserer Ergebnisse (Einzelarbeit C) wird folgender Ablauf für die Prozessierung von Pol angenommen (Hartl *et al.*, 2011) und um die Bedeutung der Verbindungs-Subdomäne und RNase H-Domäne erweitert. Zuerst bindet Pol an PARM zur Aktivierung der PR, sodass Pol vor der reversen Transkription in PR-RT und IN gespalten wird. Dadurch können sich IN-Tetramere bilden, die zur Integration der dsDNA wichtig sind. Durch die reverse Transkription verschwindet PARM und es entsteht die zur Infektion notwendige dsDNA. Dadurch ist die RNase H nicht nur, wie auch bei Orthoretroviren, unverzichtbar für die reverse Transkription, sie stellt in FV auch ein zentrales Regulationsinstrument der zeitlichen Abläufe des gesamten Replikationszyklus dar.





## 4 Abkürzungsverzeichnis

|                         |  |
|-------------------------|--|
| 1D                      | eindimensional   |
| 2D                      | zweidimensional  |
| 3D                      | dreidimensionl   |
| Å                       | Ångström   |
| AIDS                    | erworbenes Immundefizienzsyndrom (engl. <i>acquired immunodeficiency syndrome</i> )                                    |
| AZT                     | 3'-Azido-3'-desoxythymidin   |
| <i>basic protrusion</i> | Sekundärstrukturbereich der Typ 1 RNase H, der die C-Helix mit einer sich anschließender basischer Schleife beinhaltet |
| <i>B. haludorans</i>    | <i>Bacillus halodurans</i>   |
| bp                      | Basenpaare   |
| CPE                     | zytopathischer Effekt (engl. <i>cytopathic effect</i> )  |
| DNA                     | Desoxyribonukleinsäure (engl. <i>deoxyribonucleic acid</i> )   |
| dNTP                    | Desoxyribonukleosidtriphosphat   |
| ds                      | doppelsträngig (engl. <i>double stranded</i> )   |
| FV                      | Foamy Virus (engl. <i>Foamy Virus</i> )  |
| <i>E. coli</i>          | <i>Escherichia coli</i>  |
| Env                     | virale Hüllproteine  |
| Gag                     | virale Strukturproteine  |
| HBD                     | Hybridbindungsdomäne   |
| HIV                     | Humanes Immundefizienzvirus (engl. <i>Human Immunodeficiency Virus</i> )   |
| HSQC                    | <i>Heteronuclear Single Quantum Coherence</i>  |
| HTLV                    | humanes T-Zell-lymphotropes Virus (engl. <i>Human T-cell lymphotropic Virus</i> )                                      |
| IN                      | Integrase  |
| IP                      | interner Promotor  |
| LP                      | <i>leaderpeptide</i>   |
| LTR                     | lange terminale Wiederholungssequenz (engl. <i>long terminal repeat</i> )  |
| 5'-LTRP                 | 5'-LTR-Promotor  |
| MLV                     | Mausleukämievirus (engl. <i>Murine Leukemia Virus</i> )  |
| MoMLV                   | Moloney Mausleukämievirus (engl. <i>Moloney Murine Leukemia Virus</i> )  |
| mtDNA                   | mitochondriale DNA   |

|                    |   |
|--------------------|---|
| NaCl               | Natriumchlorid  |
| NMR                | Kernresonanz (engl. <i>nuclear magnetic resonance</i> )   |
| NOESY              | <i>Nuclear Overhauser Effect Spectroscopy</i>   |
| mRNA               | Boten-RNA (engl. <i>messenger RNA</i> )   |
| mtDNA              | mitochondriale DNA  |
| nt                 | Nukleotid   |
| P                  | Promotor  |
| p51                | Untereinheit der HIV-1 RT   |
| p66                | Untereinheit der HIV-1 RT   |
| PARM               | <i>protease activating RNA motif</i>  |
| PBS                | Primer-Bindestelle (engl. <i>primer binding site</i> )  |
| PFV                | Prototyp Foamy Virus (engl. <i>Prototype Foamy Virus</i> )  |
| Pol                | virale Enzyme   |
| R                  | redundante Sequenz des LTR (engl. <i>LTR repeat region</i> )  |
| PPT                | <i>polypurine tract</i>   |
| PR                 | Protease  |
| RH $\Delta$ C      | Bezeichnung einer RNase H-Domäne, deren C-Helix deletiert wurde   |
| rmsd               | <i>root-mean-square deviation</i>   |
| RNA                | Ribonukleinsäure (engl. <i>ribonucleic acid</i> )   |
| RSV                | Rous-Sarkom-Virus (engl. <i>Rous Sarcoma Virus</i> )  |
| RT                 | Reverse Transkriptase   |
| RTI                | RT-Inhibitoren  |
| SFV <sub>mac</sub> | Simian Foamy Virus aus Makaken (engl. <i>Simian Foamy Virus from macaques</i> )                                 |
| ss                 | einzelsträngig (engl. <i>single stranded</i> )  |
| SU                 | Oberflächenprotein (engl. <i>surface glycoprotein</i> )   |
| TM                 | Transmembranprotein (engl. <i>transmembrane glycoprotein</i> )  |
| tRNA               | Transfer-RNA (engl. <i>transfer RNA</i> )   |
| U3                 | unikale 3'-Region des LTR (engl. <i>LTR unique 3' region</i> )  |
| U5                 | unikale 5'-Region des LTR (engl. <i>LTR unique 5' region</i> )  |
| wt                 | Wildtyp   |
| XMRV               | Xenotropisches Mausleukämievirus verwandte Virus (engl. <i>Xenotropic Murine Leukemia Virus related Virus</i> ) |

## 5 Literaturverzeichnis

**Achong BG, Mansell PW, Epstein MA, Clifford P** (1971) An unusual virus in cultures from a human nasopharyngeal carcinoma. *J Natl Cancer Inst* **46(2)**: 299-307

**Aguilera A, Garcia-Muse T** (2012) R loops: from transcription byproducts to threats to genome stability. *Mol Cell* **46**: 115-124

**Arts EJ, Hazuda DJ** (2012) HIV-1 Antiretroviral Drug Therapy. *Cold Spring Harb Perspect Med* **2**: a007161

**Baltimore D** (1970) RNA-dependent DNA polymerase in virions of RNA tumor viruses. *Nature* **226**: 1209-1211

**Basu VP, Song M, Gao L, Rigby ST, Hanson MN und Bambara RA** (2008) Strand transfer events during HIV-1 reverse transcription. *Virus Res* **134**: 19-38

**Baunach G, Maurer B, Hahn H, Kranz M und Rethwilm A** (1993) Functional analysis of human foamy virus accessory reading frames. *J Virol* **67**: 5411-5418

**Beijerinck MW** (1898) Über ein contagium vivum fluidum als Ursache der Fleckenkrankheit der Tabaksblätter. *Verh Kon Akad Wetensch* **6**: 3-21

[Englische Übersetzung 1942: Concerning a contagium vivum fluidum as a cause of the spot-disease of tobacco leaves. *Phytopathol Classics* **7**: 33-52]

**Ben-Artzi H, Zeelon E, Le-Grice SF, Gorecki M und Panet A** (1992) Characterization of the double stranded RNA dependent RNase activity associated with recombinant reverse transcriptases. *Nucleic Acids Res* **20**: 5115-5118

**Boyer PL, Gao HQ, Frank P, Clark PK und Hughes SH** (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

**Boyer PL, Stenbak CR, Clark PK, Linial ML und Hughes SH** (2004) Characterization of the polymerase and RNase H activities of human foamy virus reverse transcriptase. *J Virol* **78**: 6112-6121

**Cerritelli SM, Crouch RJ** (2009) Ribonuclease H: the enzymes in eukaryotes. *FEBS J* **276**: 1494-1505

**Cerritelli SM, Frolova EG, Feng C, Grinberg A, Love PE und Crouch RJ** (2003) Failure to produce mitochondrial DNA results in embryonic lethality in Rnaseh1 null mice. *Mol Cell* **11**: 807-815

**Champoux JJ, Schultz SJ** (2009) Ribonuclease H: properties, substrate specificity and roles in retroviral reverse transcription. *FEBS J* **276**: 1506-1516

**Cote ML, Roth MJ** (2008) Murine leukemia virus reverse transcriptase: structural comparison with HIV-1 reverse transcriptase. *Virus Res* **134**: 186-202

**Crow YJ, Leitch A, Hayward BE, Garner A, Parmar R, Griffith E, Ali M, Semple C, Aicardi J, Babul-Hirji R, Baumann C, Baxter P, Bertini E, Chandler KE, Chitayat D, Cau D, Dery C, Fazzi E, Goizet C, King MD, Klepper J, Lacombe D, Lanzi G, Lyall H, Martinez-Frias ML, Mathieu M, McKeown C, Monier A, Oade Y, Quarrell OW, Rittey CD, Rogers RC, Sanchis A, Stephenson JB, Tacke U, Till M, Tolmie JL, Tomlin P, Voit T, Weschke B, Woods CG, Lebon P, Bonthron DT, Ponting CP und Jackson AP** (2006) Mutations in genes encoding ribonuclease H2 subunits cause Aicardi-Goutieres syndrome and mimic congenital viral brain infection. *Nat Genet* **38**: 910-916

**Das D, Georgiadis MM** (2004) The crystal structure of the monomeric reverse transcriptase from Moloney murine leukemia virus. *Structure* **12**: 819-829

**Davies JF 2nd, Hostomska Z, Hostomsky Z, Jordan SR und Matthews DA** (1991) Crystal structure of the ribonuclease H domain of HIV-1 reverse transcriptase. *Science* **252**: 88-95

**Delviks-Frankenberry KA, Nikolenko GN und Pathak VK** (2010) The "Connection" Between HIV Drug Resistance and RNase H. *Viruses* **2**: 1476-1503

**di Marzo Veronese F, Copeland TD, DeVico AL, Rahman R, Oroszlan S, Gallo RC und Sarngadharan MG** (1986) Characterization of highly immunogenic p66/p51 as the reverse transcriptase of HTLV-III/LAV. *Science* **231**: 1289-1291

**Ding J, Das K, Hsiou Y, Sarafianos SG, Clark AD, Jr, Jacobo-Molina A, Tantillo C, Hughes SH und Arnold E** (1998) Structure and functional implications of the polymerase active site region in a complex of HIV-1 RT with a double-stranded DNA template-primer and an antibody Fab fragment at 2.8 Å resolution. *J Mol Biol* **284**: 1095-1111

**Ellermann V, Bang O** (1908) Experimentelle Leukämie bei Hühnern. *Zentralbl Bakteriol Parasitenkd Infektionskr Hyg Abt Orig* **46**: 595-609

**Enders JF, Peebles TC** (1954) Propagation in tissue cultures of cytopathogenic agents from patients with measles. *Proc Soc Exp Biol Med* **86**: 277-286

**Enssle J, Fischer N, Moebes A, Mauer B, Smola U und Rethwilm A** (1997) Carboxy-terminal cleavage of the human foamy virus Gag precursor molecule is an essential step in the viral life cycle. *J Virol* **71**: 7312-7317

**Filippov V, Philippov M und Gill SS** (2001) Drosophila RNase H1 is essential for development but not for proliferation. *Mol Genet Genomics* **265**: 771-777

**Flügel RM** (2012) Humanes Spumaretrovirus. In Darai G, Handermann M, Sonntag H-G, Zöller L, eds, *Lexikon der Infektionskrankheiten des Menschen*. Springer-Verlag, Berlin Heidelberg New York, pp 426-429

**Flügel RM, Pfrepper KI** (2003) Proteolytic processing of foamy virus Gag and Pol proteins. *Curr Top Microbiol Immunol* **277**: 63-88

**Geiselhart V, Schwantes A, Bastone P, Frech M und Löchelt M** (2003) Features of the Env leader protein and the N-terminal Gag domain of feline foamy virus important for virus morphogenesis. *Virology* **310**: 235-244

- Gilboa E, Mitra SW, Goff S und Baltimore D** (1979) A detailed model of reverse transcription and tests of crucial aspects. *Cell* **18**: 93-100
- Goff SP** (2007) Retroviridae: The Retroviruses and Their Replication. In Knipe DM, Howley PM, eds, *Fields Virology*. Lippincott Williams & Wilkins, Philadelphia, pp 1999-2069
- Golomb M, Grandgenett DP** (1979) Endonuclease activity of purified RNA-directed DNA polymerase from avian myeloblastosis virus. *J Biol Chem* **254**: 1606-1613
- Hansen J, Schulze T, Mellert W und Moelling K** (1988) Identification and characterization of HIV-specific RNase H by monoclonal antibody. *EMBO J* **7**: 239-243
- Hartl MJ, Bodem J, Jochheim F, Rethwilm A, Rösch P und Wöhrl BM** (2011) Regulation of Foamy Virus Protease Activity by Viral RNA - a Novel and Unique Mechanism Among Retroviruses. *J Virol* **85**: 4462-4469
- Hartl MJ, Mayr F, Rethwilm A und Wöhrl BM** (2010a) Biophysical and enzymatic properties of the simian and prototype foamy virus reverse transcriptases. *Retrovirology* **7**: 5
- Hartl MJ, Schweimer K, Reger MH, Schwarzinger S, Bodem J, Rösch P und Wöhrl BM** (2010b) Formation of transient dimers by a retroviral protease. *Biochem J* **427**: 197-203
- Hartl MJ, Wöhrl BM, Rösch P und Schweimer K** (2008) The solution structure of the simian foamy virus protease reveals a monomeric protein. *J Mol Biol* **381**: 141-149
- Haruki M, Noguchi E, Kanaya S und Crouch RJ** (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
- Heneine W, Schweizer M, Sandstrom P und Folks T** (2003) Human infection with foamy viruses. *Curr Top Microbiol Immunol* **277**: 181-196
- Herchenröder O, Renne R, Loncar D, Cobb EK, Murthy KK, Schneider J, Mergia A und Luciw PA** (1994) Isolation, cloning, and sequencing of simian foamy viruses from chimpanzees (SFVcpz): high homology to human foamy virus (HFV). *Virology* **201**: 187-199
- Huber HE, Richardson CC** (1990) Processing of the primer for plus strand DNA synthesis by human immunodeficiency virus 1 reverse transcriptase. *J Biol Chem* **265**: 10565-10573
- Ivanofsky D.** (1892) Über die Mosaikkrankheit der Tabakspflanze. *St Petersburg Acad Imp Sci Bull* **35**: 67-70  
[Englische Übersetzung 1942: Concerning the mosaic disease of tobacco. *Phytopathol Classics* **7**: 25-30]
- Jacobo-Molina A, Ding J, Nanni RG, Clark AD Jr., Lu X, Tantillo C, Williams RL, Kamer G, Ferris AL, Clark P, Hizi A, Hughes SH und Arnold E** (1993) Crystal structure of human immunodeficiency virus type 1 reverse transcriptase complexed with double-stranded DNA at 3.0 Å resolution shows bent DNA. *Proc Natl Acad Sci U S A* **90**: 6320-6324

**Julias JG, McWilliams MJ, Sarafianos SG, Alvord WG, Arnold E und Hughes SH** (2003) Mutation of amino acids in the connection domain of human immunodeficiency virus type 1 reverse transcriptase that contact the template-primer affects RNase H activity. *J Virol* **77**: 8548-8554

**Kanaya E, Kanaya S** (1995) Kinetic analysis of Escherichia coli ribonuclease HI using oligomeric DNA/RNA substrates suggests an alternative mechanism for the interaction between the enzyme and the substrate. *Eur J Biochem* **231**: 557-562

**Kanaya S, Katsuda-Nakai C und 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

**Kanaya S, Kohara A, Miura Y, Sekiguchi A, Iwai S, Inoue H, Ohtsuka E und 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

**Kang Y, Blair WS und Cullen BR** (1998) Identification and functional characterization of a high-affinity Bel-1 DNA binding site located in the human foamy virus internal promoter. *J Virol* **72**: 504-511

**Katayanagi K, Miyagawa M, Matsushima M, Ishikawa M, Kanaya S, Nakamura H, Ikehara M, Matsuzaki T und Morikawa K** (1992) Structural details of ribonuclease H from Escherichia coli as refined to an atomic resolution. *J Mol Biol* **223**: 1029-1052

**Keck JL, Marqusee S** (1996) The putative substrate recognition loop of Escherichia coli ribonuclease H is not essential for activity. *J Biol Chem* **271**: 19883-19887

**Keck JL, Marqusee S** (1995) Substitution of a highly basic helix/loop sequence into the RNase H domain of human immunodeficiency virus reverse transcriptase restores its Mn(2+)-dependent RNase H activity. *Proc Natl Acad Sci U S A* **92**: 2740-2744

**Kim JH, Kang S, Jung SK, Yu KR, Chung SJ, Chung BH, Erikson RL, Kim BY und Kim SJ** (2012) Crystal structure of xenotropic murine leukaemia virus-related virus (XMRV) ribonuclease H. *Biosci Rep* **32**: 455-463

**Kochiwa H, Tomita M und Kanai A** (2007) Evolution of ribonuclease H genes in prokaryotes to avoid inheritance of redundant genes. *BMC Evol Biol* **7**: 128

**Kohlstaedt LA, Wang J, Friedman JM, Rice PA und Steitz TA** (1992) Crystal structure at 3.5 Å resolution of HIV-1 reverse transcriptase complexed with an inhibitor. *Science* **256**: 1783-1790

**Konvalinka J, Löchelt M, Zentgraf H, Flügel RM und Kräusslich H-G** (1995) Active foamy virus proteinase is essential for virus infectivity but not for formation of a Pol polyprotein. *J Virol* **69**: 7264-7268

**Lapkouski M, Tian L, Miller JT, Le Grice SF und Yang W** (2013) Complexes of HIV-1 RT, NNRTI and RNA/DNA hybrid reveal a structure compatible with RNA degradation. *Nat Struct Mol Biol* **20**: 230-236

- Lecellier CH, Saib A** (2000) Foamy viruses: between retroviruses and pararetroviruses. *Virology* **271**: 1-8
- Lee EG, Linial ML** (2008) The C terminus of foamy retrovirus Gag contains determinants for encapsidation of Pol protein into virions. *J Virol* **82**: 10803-10810
- Lehmann-Che J, Giron ML, Delelis O, Lochelt M, Bittoun P, Tobaly-Tapiero J, de The H und Saib A** (2005) Protease-dependent uncoating of a complex retrovirus. *J Virol* **79**: 9244-9253
- Lim D, Gregorio GG, Bingman C, Martinez-Hackert E, Hendrickson WA und Goff SP** (2006) Crystal structure of the moloney murine leukemia virus RNase H domain. *J Virol* **80**: 8379-8389
- Lim D, Orlova M und Goff SP** (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
- Lindemann D, Pietschmann T, Picard-Maureau M, Berg A, Heinkelein M, Thurow J, Knaus P, Zentgraf H und Rethwilm A** (2001) A particle-associated glycoprotein signal peptide essential for virus maturation and infectivity. *J Virol* **75**: 5762-5771
- Lindemann D, Rethwilm A** (2011) Foamy virus biology and its application for vector development. *Viruses* **3**: 561-585
- Linial M** (2007) Foamy Viruses. In Knipe DM, Howley PM, eds, *Fields Virology*. Lippincott Williams & Wilkins, Philadelphia, pp 2245-2262
- Löchelt M, Aboud M und Flügel RM** (1993) Increase in the basal transcriptional activity of the human foamy virus internal promoter by the homologous long terminal repeat promoter in cis. *Nucleic Acids Res* **21**: 4226-4230
- Löchelt M, Romen F, Bastone P, Muckenfuss H, Kirchner N, Kim YB, Truyen U, Rosler U, Battenberg M, Saib A, Flory E, Cichutek K und Munk C** (2005) The antiretroviral activity of APOBEC3 is inhibited by the foamy virus accessory Bet protein. *Proc Natl Acad Sci U S A* **102**: 7982-7987
- Loeffler F, Frosch P** (1898) Berichte (I-III) der Kommission zur Erforschung der Maul- und Klauenseuche bei dem Institut für Infektionskrankheiten in Berlin. *Centrbl Bakt Parasitenk* **23**: 371-391
- Maurer B, Bannert H, Darai G und Flügel RM** (1988) Analysis of the primary structure of the long terminal repeat and the gag and pol genes of the human spumaretrovirus. *J Virol* **62**: 1590-1597
- Mayer A.** (1882) Over de mozaïkziekte van de tabak; voorloopige mededeeling. *Tijdschr Landbouwk* **2**: 359-364  
[Englische Übersetzung: On the mosaic disease of tobacco: preliminary communication.]
- Meiring CD, Linial ML** (2001) Historical perspective of foamy virus epidemiology and infection. *Clin Microbiol Rev* **14**: 165-176

**Mizrahi V, Brooksbank RL und Nkabinde NC** (1994) Mutagenesis of the conserved aspartic acid 443, glutamic acid 478, asparagine 494, and aspartic acid 498 residues in the ribonuclease H domain of p66/p51 human immunodeficiency virus type I reverse transcriptase. Expression and biochemical analysis. *J Biol Chem* **269**: 19245-19249

**Mizrahi V, Usdin MT, Harington A und Dudding LR** (1990) Site-directed mutagenesis of the conserved Asp-443 and Asp-498 carboxy-terminal residues of HIV-1 reverse transcriptase. *Nucleic Acids Res* **18**: 5359-5363

**Moebes A, Enssle J, Bieniasz PD, Heinkelein M, Lindemann D, Bock M, McClure MO und Rethwilm A** (1997) Human foamy virus reverse transcription that occurs late in the viral replication cycle. *J Virol* **71**: 7305-7311

**Moelling K** (1974) Characterization of reverse transcriptase and RNase H from friend-murine leukemia virus. *Virology* **62**: 46-59

**Moelling K, Bolognesi DP, Bauer H, Büsen W, Plassmann HW und Hausen P** (1971) Association of retroviral reverse transcriptase with an enzyme degrading the RNA moiety of RNA-DNA hybrids. *Nat New Biol* **234**: 240-243

**Müllers E, Uhlig T, Stirrnagel K, Fiebig U, Zentgraf H und Lindemann D** (2011) Novel functions of prototype foamy virus Gag glycine- arginine-rich boxes in reverse transcription and particle morphogenesis. *J Virol* **85**: 1452-1463

**Nasimuzzaman M und Persons DA** (2012) Cell Membrane-associated Heparan Sulfate Is a Receptor for Prototype Foamy Virus in Human, Monkey, and Rodent Cells. *Mol Ther* **20**: 1158-1166, United States

**Ndongwe TP, Adedeji AO, Michailidis E, Ong YT, Hachiya A, Marchand B, Ryan EM, Rai DK, Kirby KA, Whatley AS, Burke DH, Johnson M, Ding S, Zheng YM, Liu SL, Kodama E, Delviks-Frankenberry KA, Pathak VK, Mitsuya H, Parniak MA, Singh K und Sarafianos SG** (2012) Biochemical, inhibition and inhibitor resistance studies of xenotropic murine leukemia virus-related virus reverse transcriptase. *Nucleic Acids Res* **40**: 345-359

**Nowak E, Potrzebowski W, Konarev PV, Rausch JW, Bona MK, Svergun DI, Bujnicki JM, Le Grice SF und Nowotny M** (2013) Structural analysis of monomeric retroviral reverse transcriptase in complex with an RNA/DNA hybrid. *Nucleic Acids Res*

**Nowotny M, Cerritelli SM, Ghirlando R, Gaidamakov SA, Crouch RJ und Yang W** (2008) Specific recognition of RNA/DNA hybrid and enhancement of human RNase H1 activity by HBD. *EMBO J* **27**: 1172-1181

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

**Nowotny M, Gaidamakov SA, Ghirlando R, Cerritelli SM, Crouch RJ und Yang W** (2007) Structure of human RNase H1 complexed with an RNA/DNA hybrid: insight into HIV reverse transcription. *Mol Cell* **28**: 264-276



- Ohtani N, Haruki M, Morikawa M und Kanaya S** (1999) Molecular diversities of RNases H. *J Biosci Bioeng* **88**: 12-19
- Perkovic M, Schmidt S, Marino D, Russell RA, Stauch B, Hofmann H, Kopietz F, Kloke BP, Zielonka J, Strover H, Hermle J, Lindemann D, Pathak VK, Schneider G, Lochelt M, Cichutek K und Munk C** (2009) Species-specific inhibition of APOBEC3C by the prototype foamy virus protein bet. *J Biol Chem* **284**: 5819-5826
- Peters K, Barg N, Gärtner K und Rethwilm A** (2008) Complex effects of foamy virus central purine-rich regions on viral replication. *Virology* **33**: 51-60
- Peters K, Wiktorowicz T, Heinkelein M und Rethwilm A** (2005) RNA and protein requirements for incorporation of the Pol protein into foamy virus particles. *J Virol* **79**: 7005-7013
- Pfrepper KI, Rackwitz HR, Schnölzer M, Heid H, Löchelt M und Flügel RM** (1998) Molecular characterization of proteolytic processing of the Pol proteins of human foamy virus reveals novel features of the viral protease. *J Virol* **72**: 7648-7652
- Plochmann K, Horn A, Gschmack E, Armbruster N, Krieg J, Wiktorowicz T, Weber C, Stirnagel K, Lindemann D, Rethwilm A und Scheller C** (2012) Heparan sulfate is an attachment factor for foamy virus entry. *J Virol* **86**: 10028-10035
- Puglia J, Wang T, Smith-Snyder C, Cote M, Scher M, Pelletier JN, John S, Jonsson CB und Roth MJ** (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
- Ren J, Bird LE, Chamberlain PP, Stewart-Jones GB, Stuart DI und Stammers DK** (2002) Structure of HIV-2 reverse transcriptase at 2.35-Å resolution and the mechanism of resistance to non-nucleoside inhibitors. *Proc Natl Acad Sci U S A* **99**: 14410-14415
- Rethwilm A** (2003) The replication strategy of foamy viruses. *Curr Top Microbiol Immunol* **277**: 1-26
- Rethwilm A, Erlwein O, Baunach G, Maurer B und ter Meulen V** (1991) The transcriptional transactivator of human foamy virus maps to the bel 1 genomic region. *Proc Natl Acad Sci U S A* **88**: 941-945
- Rho HM, Grandgenett DP und Green M** (1975) Sequence relatedness between the subunits of avian myeloblastosis virus reverse transcriptase. *J Biol Chem* **250**: 5278-5280
- Rigby RE, Leitch A und Jackson AP** (2008) Nucleic acid-mediated inflammatory diseases. *Bioessays* **30**: 833-842
- Rinke CS, Boyer PL, Sullivan MD, Hughes SH und Linial ML** (2002) Mutation of the catalytic domain of the foamy virus reverse transcriptase leads to loss of processivity and infectivity. *J Virol* **76**: 7560-7570

**Roberts GCK** (2003) The Determination of Equilibrium Dissociation Constants of Protein-Ligand Complexes by NMR. In Zerbe O, Mannhold R, Kubinyi H, Folkers G, eds, *BioNMR in Drug Research*. WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim, pp 309-319

**Rodgers DW, Gamblin SJ, Harris BA, Ray S, Culp JS, Hellmig B, Woolf DJ, Debouck C und Harrison SC** (1995) The structure of unliganded reverse transcriptase from the human immunodeficiency virus type 1. *Proc Natl Acad Sci U S A* **92**: 1222-1226

**Roth MJ, Tanese N und Goff SP** (1985) Purification and characterization of murine retroviral reverse transcriptase expressed in escherichia coli. *J Biol Chem* **260**: 9326-9335

**Rous P** (1911) A sarcoma of the fowl transmissible by an agent seperable from the tumor cells. *J Exp Med* **13**: 397

**Russell RA, Wiegand HL, Moore MD, Schafer A, McClure MO und Cullen BR** (2005) Foamy virus Bet proteins function as novel inhibitors of the APOBEC3 family of innate antiretroviral defense factors. *J Virol* **79**: 8724-8731

**Rustigian R, Johnston P und Reihart H** (1955) Infection of monkey kidney tissue cultures with virus-like agents. *Proc Soc Exp Biol Med* **88**: 8-16

**Sarafianos SG, Das K, Tantillo C, Clark AD Jr., Ding J, Whitcomb JM, Boyer PL, Hughes SH und Arnold E** (2001) Crystal structure of HIV-1 reverse transcriptase in complex with a polypurine tract RNA:DNA. *EMBO J* **20**: 1449-1461

**Sarafianos SG, Marchand B, Das K, Himmel DM, Parniak MA, Hughes SH und Arnold E** (2009) Structure and function of HIV-1 reverse transcriptase: molecular mechanisms of polymerization and inhibition. *J Mol Biol* **385**: 693-713

**Schatz O, Cromme FV, Gruninger-Leitch F und Le Grice SF** (1989) Point mutations in conserved amino acid residues within the C-terminal domain of HIV-1 reverse transcriptase specifically repress RNase H function. *FEBS Lett* **257**: 311-314

**Schultz SJ, Champoux JJ** (2008) RNase H activity: structure, specificity, and function in reverse transcription. *Virus Res* **134**: 86-103

**Schultz SJ, Champoux JJ** (1996) RNase H domain of Moloney murine leukemia virus reverse transcriptase retains activity but requires the polymerase domain for specificity. *J Virol* **70**: 8630-8638

**Spannaus R, Hartl MJ, Wöhrl BM, Rethwilm A und Bodem J** (2012) The prototype foamy virus protease is active independently of the integrase domain. *Retrovirology* **9**: 41

**Stahl SJ, Kaufman JD, Vikić-Topić S, Crouch RJ und Wingfield PT** (1994) Construction of an enzymatically active ribonuclease H domain of human immunodeficiency virus type 1 reverse transcriptase. *Protein Eng* **7**: 1103-1108

**Steitz TA, Steitz JA** (1993) A general two-metal-ion mechanism for catalytic RNA. *Proc Natl Acad Sci U S A* **90**: 6498-6502

- Switzer WM, Bhullar V, Shanmugam V, Cong ME, Parekh B, Lerche NW, Yee JL, Ely JJ, Boneva R, Chapman LE, Folks TM und Heneine W** (2004) Frequent simian foamy virus infection in persons occupationally exposed to nonhuman primates. *J Virol* **78**: 2780-2789
- Switzer WM, Salemi M, Shanmugam V, Gao F, Cong ME, Kuiken C, Bhullar V, Beer BE, Vallet D, Gautier-Hion A, Tooze Z, Villinger F, Holmes EC und Heneine W** (2005) Ancient co-speciation of simian foamy viruses and primates. *Nature* **434**: 376-380
- Tadokoro T, Kanaya S** (2009) Ribonuclease H: molecular diversities, substrate binding domains, and catalytic mechanism of the prokaryotic enzymes. *FEBS J* **276**: 1482-1493
- Tanese N, Goff SP** (1988) Domain structure of the Moloney murine leukemia virus reverse transcriptase: mutational analysis and separate expression of the DNA polymerase and RNase H activities. *Proc Natl Acad Sci U S A* **85**: 1777-1781
- Telesnitsky A, Blain SW und Goff SP** (1992) Defects in Moloney murine leukemia virus replication caused by a reverse transcriptase mutation modeled on the structure of Escherichia coli RNase H. *J Virol* **66**: 615-622
- Temin H, Mizutani S** (1970) RNA-dependent DNA polymerase in virions of Rous sarcoma virus. *Nature* **226**: 1211-1213
- Temin HM** (1964) Nature of the provirus of Rous sarcoma. National Cancer Institute Monograph **17**: 557-570
- Tisdale M, Schulze T, Larder BA und Moelling K** (1991) Mutations within the RNase H domain of human immunodeficiency virus type 1 reverse transcriptase abolish virus infectivity. *J Gen Virol* **72 ( Pt 1)**: 59-66
- Tobaly-Tapiero J, Bittoun P, Giron ML, Neves M, Koken M, Saib A, de Thé H** (2001) Human foamy virus capsid formation requires an interaction domain in the N terminus of gag. *J Virol* **75 (9)**: 4367-4375
- Trentin B, Rebeyrotte N und Mamoun RZ** (1998) Human T-cell leukemia virus type 1 reverse transcriptase (RT) originates from the pro and pol open reading frames and requires the presence of RT-RNase H (RH) and RT-RH-integrase proteins for its activity. *J Virol* **72**: 6504-6510
- Verma IM** (1975) Studies on reverse transcriptase of RNA tumor viruses III. Properties of purified Moloney murine leukemia virus DNA polymerase and associated RNase H. *J Virol* **15**: 843-854
- Volkman S, Wöhrl BM, Tisdale M und Moelling K** (1993) Enzymatic analysis of two HIV-1 reverse transcriptase mutants with mutations in carboxyl-terminal amino acid residues conserved among retroviral ribonucleases H. *J Biol Chem* **268**: 2674-2683
- Wilk T, Geiselhart V, Frech M, Fuller SD, Flügel RM und Löchelt M** (2001) Specific interaction of a novel foamy virus Env leader protein with the N-terminal Gag domain. *J Virol* **75**: 7995-8007

**Wöhrl BM, Moelling K** (1990) Interaction of HIV-1 ribonuclease H with polypurine tract containing RNA-DNA hybrids. *Biochemistry (N Y)* **29**: 10141-10147

**Wöhrl BM, Volkmann S und Moelling K** (1991) Mutations of a conserved residue within HIV-1 ribonuclease H affect its exo- and endonuclease activities. *J Mol Biol* **220**: 801-818

**Yang W, Hendrickson WA, Crouch RJ und Satow Y** (1990) Structure of ribonuclease H phased at 2 Å resolution by MAD analysis of the selenomethionyl protein. *Science* **249**: 1398-1405

**Yu SF, Sullivan MD und Linial ML** (1999) Evidence that the human foamy virus genome is DNA. *J Virol* **73**: 1565-1572

**Zemba M, Wilk T, Rutten T, Wagner A, Flügel RM und Löchelt M** (1998) The carboxy-terminal p3Gag domain of the human foamy virus Gag precursor is required for efficient virus infectivity. *Virology* **247**: 7-13

**Zhan X, Crouch RJ** (1997) The isolated RNase H domain of murine leukemia virus reverse transcriptase. Retention of activity with concomitant loss of specificity. *J Biol Chem* **272**: 22023-22029

**Zhou D, Chung S, Miller M, Grice SF und 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

## 6 Publikationsliste

### Einzelarbeit A

Berit Leo, Maximilian J. Hartl, Kristian Schweimer, Florian Mayr und Birgitta M. Wöhrl (2012): Insights into the structure and activity of prototype foamy virus RNase H. *Retrovirology* **9**:14.

Die Reinigung und Größenausschlusschromatographie der PFV RNase H sowie alle RNase H-Aktivitätstests wurden von mir durchgeführt und mit Hilfe von Maximilian J. Hartl ausgewertet. Florian Mayr erarbeitete die grundlegenden Testbedingungen für den fluoreszenzbasierten RNase H-Test. Die Aufnahme der NMR-Spektren wurde von Kristian Schweimer realisiert. Die Auswertung der NMR-Experimente wurde von mir und Kristian Schweimer durchgeführt. Das Projekt wurde von Birgitta M. Wöhrl betreut. Das Manuskript wurde von Birgitta M. Wöhrl, Kristian Schweimer und mir verfasst.

### Einzelarbeit B

Berit Leo, Kristian Schweimer, Paul Rösch, Maximilian J. Hartl und Birgitta M. Wöhrl (2012): The solution structure of the prototype foamy virus RNase H domain indicates an important role of the basic loop in substrate binding. *Retrovirology* **9**:73.

Die Produktion der RNase H-Mutanten sowie die Reinigungen und die Aktivitätstests derer und der wt RNase H wurden von mir durchgeführt und ausgewertet. Der Polymerisationstest der PFV PR-RT-(D599N-H724N)-Mutante und der PR-RT wt wurde von mir ausgeführt und zusammen mit Maximilian J. Hartl ausgewertet. Die NMR-Spektren zur Strukturlösung wurden von Kristian Schweimer aufgenommen. Die Auswertung der Daten der NMR-Experimente sowie die Strukturrechnung wurden von mir und Kristian Schweimer realisiert. Die NMR-Titrationsexperimente wurden von mir durchgeführt und analysiert. Die <sup>15</sup>N-Spin Relaxationsexperimente wurden von Kristian Schweimer aufgenommen und ausgewertet. Paul Rösch lieferte einen konzeptionellen Beitrag. Birgitta M. Wöhrl betreute das Projekt und verfasste zusammen mit mir und Kristian Schweimer das Manuskript.

### **Einzelarbeit C**

Anna Schneider, Daniel Peter, Jessica Schmitt, Berit Leo, Franziska Richter, Paul Rösch, Birgitta M. Wöhrl und Maximilian J. Hartl (2013): Structural requirements for enzymatic activities of foamy virus protease-reverse transcriptase. *Manuskript*.

Die Klonierung und Reinigung der SFVmac RNase H sowie die RNase H-Tests wurden von mir durchgeführt. Anna Schneider, Daniel Peter und Jessica Schmitt realisierten zusammen mit Maximilian J. Hartl die Klonierungen, Reinigungen und Polymerasetests der PR-RT-Deletionsvarianten. Die NMR-Messungen wurden von Anna Schneider und Maximilian J. Hartl durchgeführt und ausgewertet. Franziska Richter und Maximilian Hartl waren für die Protease-Aktivitätstests verantwortlich. Paul Rösch lieferte einen konzeptionellen Beitrag und trug zur Diskussion der Ergebnisse bei. Birgitta M. Wöhrl und Maximilian J. Hartl planten und betreuten die Versuche und schrieben das Manuskript.

## 7 Einzelarbeit A

Berit Leo, Maximilian J. Hartl, Kristian Schweimer, Florian Mayr und Birgitta M. Wöhl (2012): Insights into the structure and activity of prototype foamy virus RNase H. *Retrovirology* **9**:14.







RESEARCH

Open Access

# Insights into the structure and activity of prototype foamy virus RNase H

Berit Leo, Maximilian J Hartl, Kristian Schweimer, Florian Mayr and Birgitta M Wöhrl\*

## Abstract

**Background:** RNase H is an endonuclease that hydrolyzes the RNA strand in RNA/DNA hybrids. Retroviral reverse transcriptases harbor a C-terminal RNase H domain whose activity is essential for viral replication. The RNase H degrades the viral genomic RNA after the first DNA strand is synthesized. Here, we report the biophysical and enzymatic properties of the RNase H domain of prototype foamy virus (PFV) as an independently purified protein. Sequence comparisons with other retroviral RNases H indicated that PFV RNase H harbors a basic protrusion, including a basic loop and the so-called C-helix, which was suggested to be important for activity and substrate binding and is absent in the RNase H domain of human immunodeficiency virus. So far, no structure of a retroviral RNase H containing a C-helix is available.

**Results:** RNase H activity assays demonstrate that the PFV RNase H domain is active, although its activity is about 200-fold reduced as compared to the full length protease-reverse transcriptase enzyme. Fluorescence equilibrium titrations with an RNA/DNA substrate revealed a  $K_D$  for the RNase H domain in the low micromolar range which is about 4000-fold higher than that of the full-length protease-reverse transcriptase enzyme. Analysis of the RNase H cleavage pattern using a [ $^{32}$ P]-labeled substrate indicates that the independent RNase H domain cleaves the substrate non-specifically. The purified RNase H domain exhibits a well defined three-dimensional structure in solution which is stabilized in the presence of  $Mg^{2+}$  ions.

**Conclusions:** Our data demonstrate that the independent PFV RNase H domain is structured and active. The presence of the C-helix in PFV RNase H could be confirmed by assigning the protein backbone and calculating the chemical shift index using NMR spectroscopy.

**Keywords:** PFV, retroviral RNase H, C-helix, basic loop, NMR

## Background

Retroviral reverse transcription describes the formation of a double-stranded DNA using the single-stranded viral RNA genome as a template. The process is catalyzed by the viral reverse transcriptase which harbors a polymerase and an RNase H domain.

In most retroviruses reverse transcription takes place after the virus has entered the host cell. Spumaviruses, or foamy viruses (FVs), belong to a subfamily of the *retroviridae* and follow a distinct replication pattern unique among retroviruses: (a) reverse transcription occurs predominantly in the virus producing cell (b) the *pol*-gene coding for the viral enzymes is expressed from an

independently spliced mRNA and (c) the viral protease is not cleaved off from the Pol precursor protein, leading to a mature protease-reverse transcriptase (PR-RT) [1-5]. Thus the mature PR-RT of FVs harbors a protease, polymerase and RNase H domain, each possessing a distinct enzymatic activity [4].

Retroviral RNases H are domains of the RT enzymes and degrade the RNA strand of the RNA/DNA hybrid which is formed in the first step of reverse transcription. This catalytic activity is essential during reverse transcription and leads to degradation of the genomic RNA during synthesis of the first or so-called minus DNA using the RNA as a template [6,7]. Mutations that inactivate the RNase H prevent viral replication [8,9].

RNase H cleavage by retroviral RTs, though generally not sequence specific, is also not devoid of any specificity. Retroviral RNases H exhibit three types of RNA

\* Correspondence: birgitta.woehr@uni-bayreuth.de  
Universität Bayreuth, Lehrstuhl Biopolymere, Universitätsstr. 30, D-95447 Bayreuth, Germany

cleavages: DNA 3' end directed, RNA 5' end directed, and internal. Specific cleavages are required during reverse transcription, when the extended tRNA and polypurine tract (PPT) primers, which are essential for the initiation of minus and plus strand DNA synthesis, have to be removed [10-17].

The isolated RNase H domain from human immunodeficiency virus type 1 (HIV-1) is enzymatically inactive [18-21], whereas that from Moloney murine leukemia virus (MoMLV) retains RNase H activity, albeit considerably lower than that of the full length RT [22-24]. Sequence alignments and structural comparisons revealed that MoMLV RNase H contains an additional helix-loop structure, also named the "basic protrusion", which includes the so-called C-helix and a downstream basic loop element [25-27]. This structural feature is also present in the human and *Escherichia coli* (*E. coli*) RNases H, both exist as independent proteins [28,29]. In contrast, the positively charged basic protrusion, which has been suggested to be important for substrate binding and/or activity [30-32], is missing in the HIV-1 RNase H domain, which is part of the RT enzyme [33]. Obviously, positively charged residues in the connection domain fulfill this function in HIV-1 RT [34]. Deletion of the putative C-helix of MoMLV RT results in replication defective viruses [30-32]. Similar to HIV-1 RNase H, the RNase H of the bacterium *Bacillus halodurans* is also devoid of a C-helix but still harbors a basic loop element. This RNase H carries an additional N-terminal extension, the so-called RNA-hybrid binding domain (RHBD), which is important for substrate binding [35].

Here, we analyzed the enzymatic and structural properties of the RNase H domain of prototype FV (PFV) because sequence alignments with other RNases H indicated the presence of a C-helix. We wanted to receive more information on this structural element, since HIV-1 RNase H does not possess a C-helix and the published structure of MoMLV RNase H could only be solved using a deletion mutant lacking the C-helix [25], whereas the crystal structure of the MoMLV RT in complex with duplex DNA (PDB: 1RW3) did not permit the building of a detailed model for the RNase H domain, since the electron density of this domain was not sufficiently well-ordered [36].

Activity assays and NMR spectroscopy of PFV RNase H show that it exhibits enzymatic activity and a well-defined three dimensional structure which harbors the basic protrusion including the C-helix.

## Results and Discussion

### Purification of the PFV RNase H domain

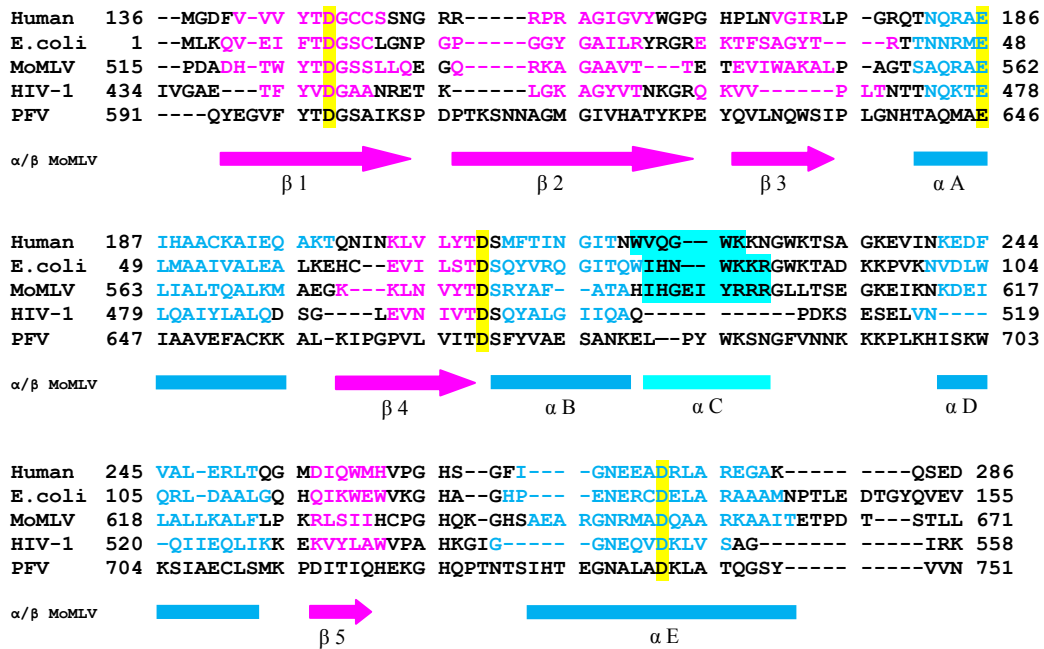
Secondary structure predictions and sequence alignments of the PFV RNase H domain with the human and *E. coli* RNases H as well as with the retroviral RNases H

from HIV-1 and MoMLV indicated the existence of a C-helix (Figure 1) in PFV and MoMLV RNases H, whereas the RNase H of HIV-1 does not possess a C-helix. The putative C-helix appears to be important for the activity of the free RNase H domain of MoMLV [25,30-32]. So far there is no structure of a retroviral RNase H domain available that actually shows the C-helix. Thus we wanted to prove experimentally by activity assays as well as NMR spectroscopy that PFV RNase H harbors a C-helix.

To analyze the properties of the RNase H domain of PFV we cloned the corresponding gene in an *E. coli* expression vector. A publication by Pfrepper *et al.* [37] suggested the existence of a cleavage site specific for FV protease (PR) between the polymerase and RNase H domains of FV RT at the sequence E<sup>593</sup>GVE<sup>596</sup>↓Y<sup>597</sup>TDG<sup>600</sup>. (The amino acid numbering always refers to the full length PFV PR-RT protein).

In contrast, we have shown recently that the only cleavage in the Pol precursor protein takes place between the RNase H and integrase domains. The mature PR-RT of PFV is monomeric, and the RNase H is not cleaved off [4,5]. Furthermore, secondary structure predictions indicated that the putative cleavage site suggested by Pfrepper *et al.* [37] is located in a  $\beta$ -strand (Figure 1). Cleavage in a secondary structure element can have detrimental effects on the structural integrity of the protein and thus this cleavage appears unlikely.

To find out more about the structure and function of PFV RNase H, we expressed two different proteins, the first one starting with Y<sup>597</sup>TDG as suggested by Pfrepper *et al.* [37] as the N-terminal starting point of the PFV RNase H domain and the second one starting six amino acids further upstream with the sequence Q<sup>591</sup>YEGVFYTDG. This starting point was chosen, to avoid cleavage in the predicted  $\beta$ -strand (Figure 1). The two proteins were expressed with an N-terminal GB1 (immunoglobulin binding domain B1 of the streptococcal protein G) solubility tag. The resulting fusion proteins harbor a hexa-histidine tag at the N-terminus and a tobacco etch virus (TEV) protease cleavage site between GB1 and the corresponding RNase H domain. However, the short version of the RNase H (Y<sup>597</sup>-N<sup>751</sup>) proved unstable during expression and purification, indicating that the putative first  $\beta$ -strand of the RNase H is essential for the proper three-dimensional structure of the protein. Only the longer RNase H version, RNase H-(Q<sup>591</sup>-N<sup>751</sup>), could be purified successfully and was used for further analysis. After TEV cleavage and removal of the GB1 tag, the purity of the RNase H-(Q<sup>591</sup>-N<sup>751</sup>) was analyzed on a 19% SDS polyacrylamide gel (Figure 2A). To determine whether the RNase H is monomeric, we performed size exclusion chromatography of RNase H-(Q<sup>591</sup>-N<sup>751</sup>) (Figure 2B). The



**Figure 1** Sequence alignment of different RNase H domains. The RNases H from human origin, *E. coli*, MoMLV RT and HIV-1 RT are shown in comparison to the PFV RNase H. The numbers represent the amino acid numbers of the corresponding full length enzymes. The alignment was performed with the program CLC Protein Workbench 5 (version 5.5.2.0). Arrows below in pink indicate the β-strands in MoMLV RNase H, the corresponding amino acids in the other RNases H are highlighted by pink letters. Blue bars indicate α-helices in MoMLV RNase H, α-helical stretches of the other RNases H are highlighted by blue letters. The C-helices are highlighted by bright cyan boxes. The active site residues are highlighted in yellow.

chromatogram shows a single peak corresponding to a molecular mass of 19 kDa (theoretical value: 18 kDa), demonstrating that the RNase H domain is monomeric.

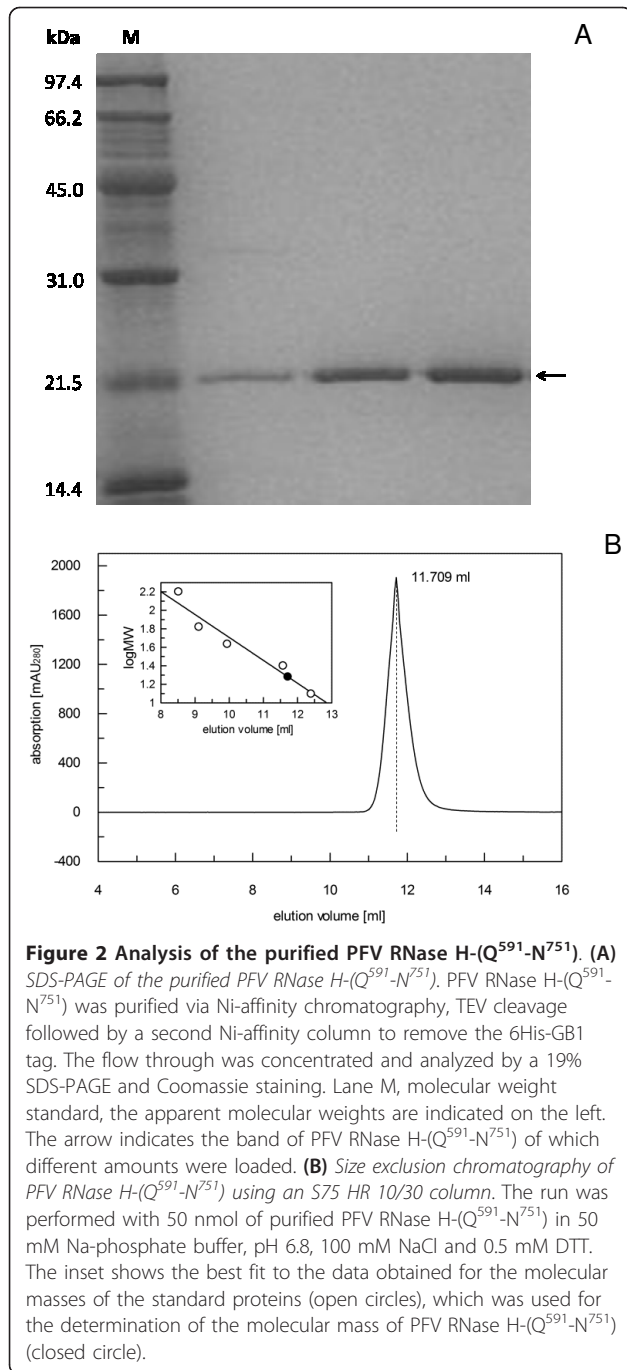
### Enzyme activity

Mutational analysis performed with MoMLV RNase H indicated that the C-helix, which is part of the basic protrusion (L<sup>681</sup>-K<sup>698</sup> in PFV RNase H) is important for the activity of the separate RNase H domain [22-24,38]. The putative basic protrusion of PFV RNase H harbors five lysine residues (Figure 1).

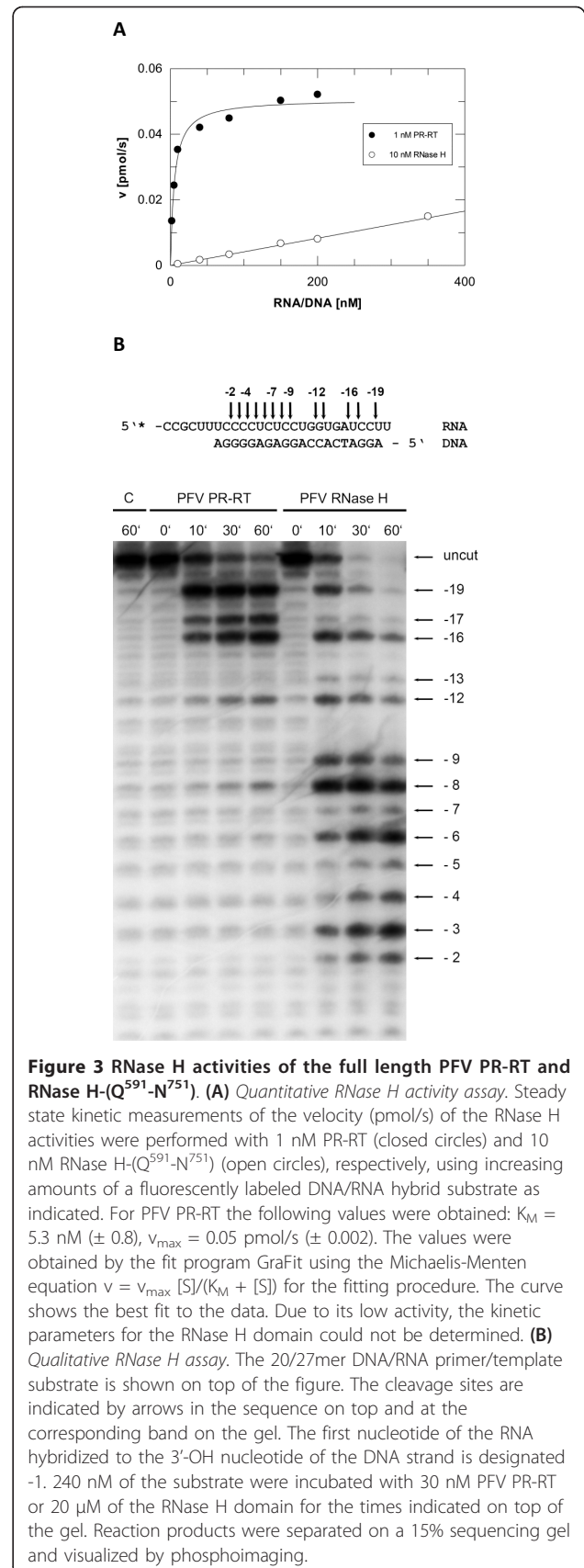
Thus, we determined the cleavage activity of the free RNase H-(Q<sup>591</sup>-N<sup>751</sup>) in comparison with the RNase H domain of the full length PFV PR-RT. We chose a fluorescence assay for the detection of the activities and kinetic parameters since this assay proved to be very sensitive (Figure 3A). Our results indicate that the free PFV RNase H-(Q<sup>591</sup>-N<sup>751</sup>) domain is still active, although to a much lesser extent than the full length PR-RT. However, due to the low activity of PFV RNase H-(Q<sup>591</sup>-N<sup>751</sup>) we were only able to determine  $K_M$  and  $v_{max}$  values for the full length PFV PR-RT ( $K_M = 5.3$  nM ( $\pm 0.8$ ),  $v_{max} = 0.05$  pmol/s ( $\pm 0.002$ )). Even when a 10-fold higher enzyme concentration of PFV RNase H-(Q<sup>591</sup>-N<sup>751</sup>) (Figure 3A) was used, the activities were more than 20 fold lower than for the PR-RT and

saturation could not be achieved. Therefore, neither the fitting procedure nor graphical methods (Lineweaver-Burk, Eadie-Hofstee, Hanes) for  $K_M$  and  $v_{max}$  determination gave reliable results.

To further characterize the activity of RNase H-(Q<sup>591</sup>-N<sup>751</sup>), a qualitative RNase H assay was performed with a non-specific 20/27 mer DNA/RNA hybrid. The RNA was labeled with [<sup>32</sup>P] at the 5' end to visualize the cleavage products. To obtain cleavage of the substrate, we had to use almost 700 fold more RNase H-(Q<sup>591</sup>-N<sup>751</sup>) than PR-RT (Figure 3B). This confirms our results with the fluorescence assay described above, which also indicated a lower activity of RNase H-(Q<sup>591</sup>-N<sup>751</sup>) (Figure 3A). This is also reflected by the much higher dissociation constant ( $K_D$ ) determined (see below) for the RNase H domain (Figure 4). The cleavage pattern of the PR-RT confirms our previous data using a different substrate [4]. The cleavages performed by PR-RT are directed by the 3' end of the DNA primer, which binds to the active site of the polymerase domain [10]. Consequently, primary cleavage sites can be detected 16-19 nucleotides away from the 3' OH primer terminus of the hybrid. After 60 minutes, shorter cleavage products can be found up to position -8. In contrast, due to the lack of a polymerase active site, the cleavage pattern of the RNase H domain is different. After just 10 minutes,



the RNase H domain exhibits a more equal distribution of cleavage sites spanning from -2 to -19. Nevertheless, certain sites appear to be preferred. Due to the high enzyme concentration used, we cannot exclude that this is derived from the binding of more than one enzyme molecule to the substrate. Our results are comparable to those obtained with the separate RNase H domain of MoMLV, which is also relatively non-specific even when



specific substrates, i.e. containing the polypurine tract are used [23,24]

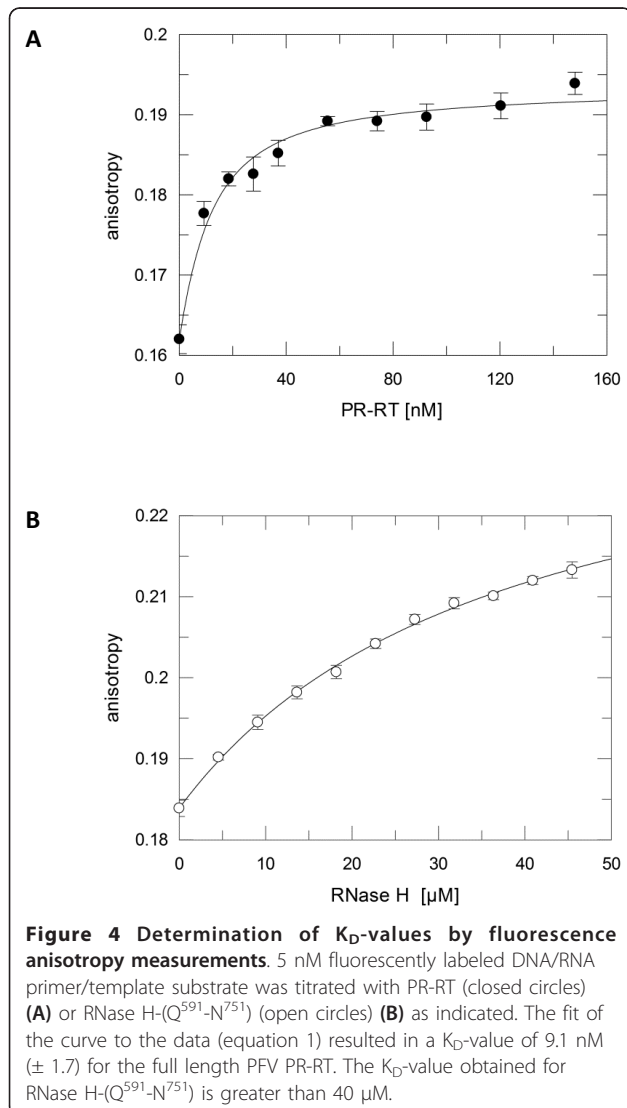
To get more information on the kinetic parameters, we set out to determine the dissociation constant ( $K_D$ ) for substrate binding. In the full length PR-RT enzyme, binding of the RNase H domain is amplified by the polymerase domain which also binds to the substrate and, due to its larger size, covers a larger portion of it. To analyze whether the higher RNase H activity of the full length PR-RT can, at least in part, be explained by a higher binding affinity of PR-RT, we performed fluorescence equilibrium titrations with a fluorescently labeled DNA/RNA primer/template substrate. The  $K_D$  value of 9.1 nM ( $\pm 1.7$ ) we obtained for the full length PFV PR-RT (Figure 4A) is comparable to previously published results [4]. However, again due to the fact that we were not able to reach saturation, the  $K_D$ -value for the RNase

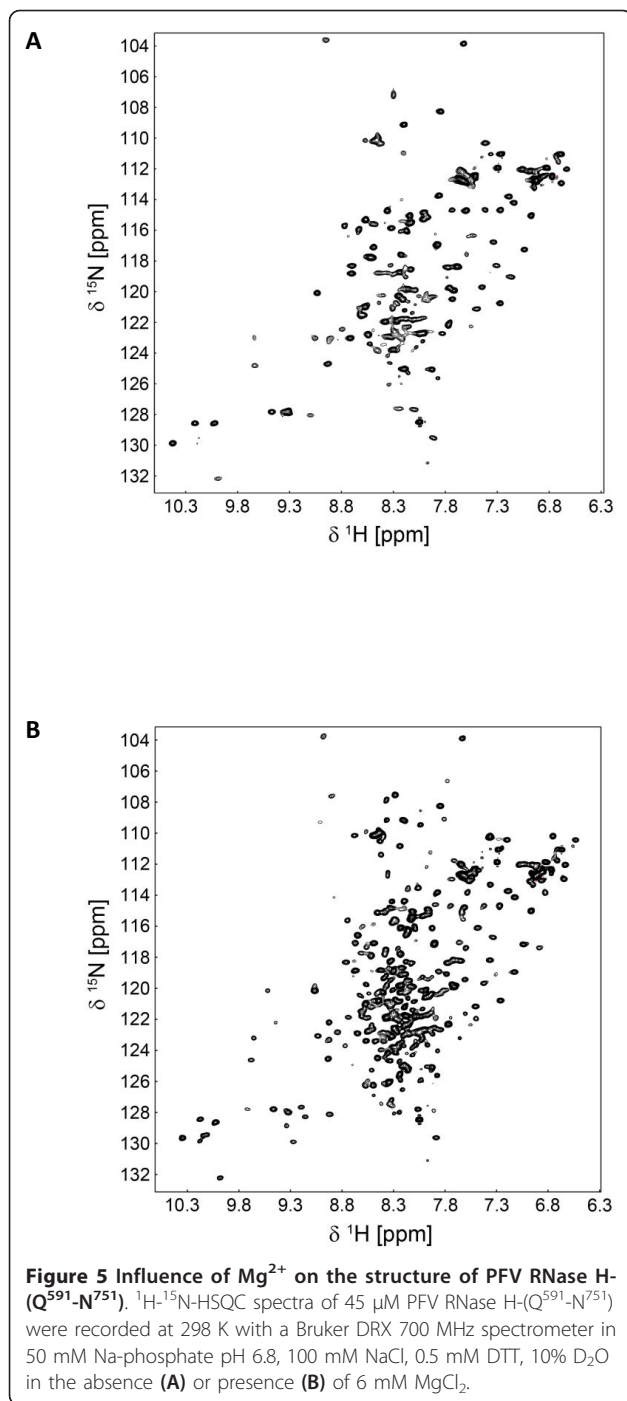
H domain could not be determined precisely and is greater than 40  $\mu$ M (Figure 4B) showing a binding affinity which is about 4000-fold lower than that of PFV PR-RT. These results imply a large contribution of the polymerase domain to substrate binding. Nevertheless, the  $K_D$ -value gives a rough estimate of the  $K_M$ -value, due to the fact that  $K_M = (k_{-1}/k_1) + (k_2/k_1)$ , or if  $k_2$  can be neglected  $K_M = k_{-1}/k_1 = K_D$ , indicating that the  $K_M$  for the RNase H domain is greater than 40  $\mu$ M.

#### NMR analyses

Proper folding of the RNase H domain was analyzed by NMR spectroscopy using uniformly  $^{15}$ N-labeled protein. The activity of retroviral RNases H has been shown to be dependent on  $Mg^{2+}$ , therefore NMR-spectra in the absence and presence of  $Mg^{2+}$  ions were recorded (Figure 5). An NMR spectrum correlating resonance frequencies of amide protons and directly bonded  $^{15}$ N-labeled nitrogen atoms (2D [ $^1$ H- $^{15}$ N] HSQC, *heteronuclear single quantum correlation*) allows the individual detection of peptide backbone amide signals. Thus, each signal represents a single amino acid of the peptide chain, and the spectrum serves as a fingerprint of the folding state of a protein. Adoption of a defined structure increases the dispersion of signals of amide protons in the HSQC spectrum to 6.9 - 9.7 ppm, relative to a random coil structure, which only exhibits an amide proton signal range from 7.8 - 8.7 ppm.

The spectra measured with the RNase H domain reveal that already in the absence of  $Mg^{2+}$  well-dispersed signals (signal range: 6.9 - 9.7 ppm) typical for a protein with a defined tertiary structure are obtained (Figure 5A). However, the number of signals is not sufficient for a protein with 165 amino acids, and the signal intensities show significant differences. This is due to non-uniform NMR relaxation properties displaying a variety of dynamic processes not typical for a well-folded stable domain. In the presence of  $Mg^{2+}$  the number of resonances increased remarkably, and the signal intensities show a more uniform distribution which is characteristic for a folded globular domain of 18 kDa (Figure 5B). These results indicate that for the structural integrity of the RNase H  $Mg^{2+}$  is essential. It has been shown for other RNases H, that two  $Mg^{2+}$  ions are bound in the catalytic center and are important for catalysis of the RNA cleavage by a two metal ion mechanism [29,39,40]. Although the presence of  $Mg^{2+}$  is apparently not necessary for substrate binding by RTs [41-43], the  $Mg^{2+}$  ions appear to stabilize the overall structure of the RNase H domain and make the protein more rigid since the  $Mg^{2+}$  ions are coordinated by three highly conserved aspartate residues and one glutamate residue in RNases H of known structure. In the case of PFV RNase H the corresponding residues are D<sup>599</sup>, E<sup>646</sup>, D<sup>669</sup>, D<sup>740</sup> (Figure 1).





NMR data analysis of triple resonance experiments of uniformly <sup>15</sup>N, <sup>13</sup>C labeled RNase H-(Q<sup>591</sup>-N<sup>751</sup>) allowed the assignment of the backbone resonances for all residues except, D<sup>607</sup>, N<sup>639</sup>, N<sup>692</sup>-K<sup>695</sup>, Q<sup>725</sup>-N<sup>728</sup>, S<sup>730</sup>, I<sup>731</sup>, E<sup>734</sup>. The assigned [<sup>1</sup>H, <sup>15</sup>N] HSQC spectrum is depicted in Figure 6A. Residues which constitute the C-helix (E<sup>680</sup> - S<sup>686</sup>), as determined by the chemical shift index (CSI) method (see below), are marked by boxes.

The proline at position 682 is not visible in the HSQC spectrum due to the lack of a backbone N-H bond.

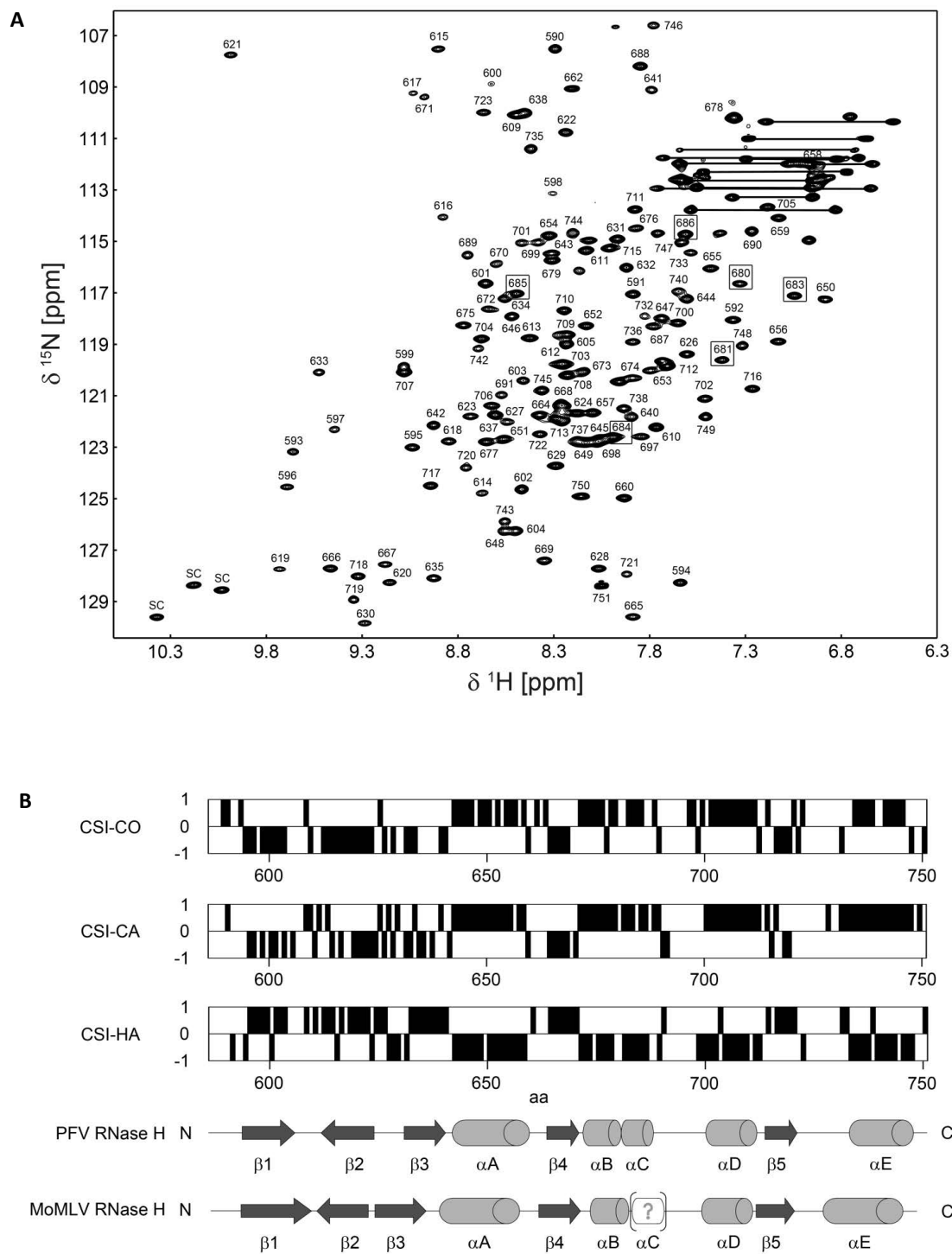
To determine the secondary structure elements and to prove the existence of the C-helix experimentally, we compared the observed NMR resonance frequencies with tabulated values found in random coil peptides by the CSI method (Figure 6B). The systematic deviations of backbone chemical shifts (resonance frequencies relative to reference standard) from random coil values in folded proteins form a valuable basis for the determination of sequence specific secondary structures and allow the experimental verification of postulated secondary structure elements. Helical secondary structures result in upfield chemical shift changes (lower resonance frequencies, lower ppm-values) of the α-protons (HA; negative CSI value) and downfield chemical shift changes (higher resonance frequencies, higher ppm-values) of the α- and carbonyl carbons (CA, CO; positive CSI value). β-strands exhibit the opposite direction of chemical shift changes compared to α-helical secondary structures.

In the RNase H domain, five sequence stretches (G<sup>594</sup>-I<sup>603</sup>, N<sup>612</sup>-Y<sup>623</sup>, N<sup>631</sup>-I<sup>635</sup>, V<sup>664</sup>-T<sup>668</sup>, I<sup>716</sup>-E<sup>721</sup>) show upfield chemical shifts of α- and carbonyl carbons and downfield chemical shifts of their α-protons, relative to random coil values and thus can be identified as β-strands (Figure 6B). The five sequence regions A<sup>642</sup>-L<sup>658</sup>, F<sup>671</sup>-K<sup>679</sup>, E<sup>680</sup>-S<sup>686</sup>, H<sup>699</sup>-S<sup>711</sup> and T<sup>733</sup>-G<sup>746</sup> exhibit the opposite direction of chemical shift changes compared to random coil values indicating the presence of α-helical secondary structures. The resulting overall sequence specific secondary structure resembles that of the RNase H domain of MoMLV (Figure 6B) and is consistent with other RNase H domains (Figure 1). The C-helix of PFV RNase H spans the α-helical region from E<sup>680</sup>-S<sup>686</sup> (Figure 6B).

In contrast, for MoMLV the C-helix was only predicted by sequence alignments and structure comparisons with other RNases H since no structure of the wild-type MoMLV RNase H could be obtained. The crystal structure of the MoMLV RNase H domain (PDB: 2HB5) could only be solved with a mutant protein lacking the C-helix [25]; thus no precise information on the orientation of the C-helix and its involvement in substrate binding could be made. Here we show that the solution structure analysis of PFV RNase H reveals the C-helix and adjacent regions. This is the first structural analysis of a retroviral RNase H where the C-helix can be detected.

## Conclusion

The detection of activity with the free PFV RNase H domain complements our NMR data which indicate that the protein is properly folded and harbors the basic



**Figure 6 NMR backbone assignment.** (A)  $[^1\text{H}, ^{15}\text{N}]$  spectrum of the uniformly  $^{15}\text{N}, ^{13}\text{C}$  labeled PFV RNase H-( $Q^{591}-N^{751}$ ). Numbers refer to amino acid positions of the corresponding residues in the full length PFV PR-RT. NH<sub>2</sub> side chains are connected by lines, the tryptophan side chains are marked by 'sc'. Several amino acids of the C-helix are marked by boxes. (B) Secondary chemical shift indices for  $H_{\alpha}$ ,  $C_{\alpha}$ , CO nuclei, and the consensus of PFV RNase H-( $Q^{591}-N^{751}$ ). Deduced  $\beta$ -strands are represented by arrows, and helical regions by cylinders. Precise amino acid numbering of the secondary structure elements is given in the text.

protrusion including the C-helix. Our results suggest that similar to MoMLV the independent monomeric PFV RNase H domain is active due to the presence of the basic protrusion.

Furthermore, the data obtained here with PFV RNase H-(Q<sup>591</sup>-N<sup>751</sup>) together with previously published data [4,5] provide additional evidence that argues against the predicted PR cleavage between the polymerase and RNase H domain taking place [37].

These data expand the knowledge obtained with free MoMLV RNase H, which can also be expressed separately as an active enzyme, but with lower activity and less substrate specificity [22-24], and can now be used for further structural analyses by NMR since it has not been possible so far to obtain the structure of a retroviral RNase H domain harboring a C-helix.

## Methods

### Cloning, expression and protein purification

The gene coding for PFV RNase H starting either with the codon for Y<sup>597</sup> or Q<sup>591</sup> was amplified by PCR with a 5' primer including an *NcoI* site and a 3' primer with a *BamHI* restriction site for further cloning. The fragment was cloned into the vector pET-GB1a (G. Stier, EMBL, Heidelberg, Germany) via *NcoI* and *BamHI*. The expressed RNase H domains (ca. 18 kDa) were thereby fused to the 56 amino acids long B1 domain of *Streptococcal* protein G (GB1, 8740 Da). Both fusion proteins (ca. 27 kDa) harbor a hexa-histidine tag at the N-terminus and a tobacco etch virus (TEV) protease cleavage site between the GB1 and RNase H domains. Due to the *NcoI* cloning site, the cleaved RNase H variants harbor an N-terminal extension of four additional amino acids (GAMG) upstream of the actual start of the RNase H amino acid sequence.

Gene expression was done in *E. coli* Rosetta DE3 (Novagen) in Luria broth [44] in the presence of 34 µg/ml chloramphenicol and 30 µg/ml kanamycin. For NMR studies M9 minimal medium [44] with 1× MEM vitamin solution (Gibco, Karlsruhe, Germany) supplemented with 1.5 g/l (<sup>15</sup>NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> (Cambridge Isotope Laboratories, Inc., Andover, MA, USA) and 4 g/l <sup>13</sup>C-glucose (Euriso-Top, GIF-SUR-YVETTE, France) as the sole nitrogen and carbon sources were used. Cells were grown at 37°C to an optical density at 600 nm of 0.8. Expression was induced by adding 1 mM IPTG followed by incubation at 16°C over night. Cells were lysed using a Microfluidizer (MFTI Corporation, Newton, MA, USA) in a buffer containing 50 mM Na-phosphate pH 6.8, 1 M NaCl, 10 mM imidazole, 0.5 mM dithiothreitol (DTT).

The protein was purified by immobilized nickel affinity chromatography (5 ml HisTrap HP column, GE Healthcare, Munich, Germany) applying an imidazole

step gradient for elution. Fractions containing GB1-RNase H were dialyzed against 50 mM Na-phosphate buffer pH 6.8, 300 mM NaCl, 0.5 mM DTT and cleaved with 1 mg TEV protease per 7 mg fusion protein for 21-24 h at 4°C. The TEV protease and the cleaved off GB1 tag were removed by a second Ni-affinity chromatography. The free PFV RNase H-(Q<sup>591</sup>-N<sup>751</sup>) was collected in the flow through. For HSQC-analysis in the presence of MgCl<sub>2</sub> the purification was performed with 6 mM MgCl<sub>2</sub> throughout the procedure.

Size exclusion chromatography of PFV RNase H-(Q<sup>591</sup>-N<sup>751</sup>) and purification of PFV PR-RT was performed as described previously [4,45].

### Quantitative RNase H activity assay

RNase H activity was measured quantitatively to determine steady state kinetic parameters by a fluorescence assay as described previously [4]. The RNA-strand of the hybrid was fluorescently labeled at the 3' end with 6-carboxy-fluorescein (6-FAM) (RNA-5'-CCGAUCGCU-CUCCUGGUGAUCCUUUCC-6-FAM), whereas the DNA primer harbored a dabcyI quencher at the 5' end (DNA-5'-DabcyI-GGAAAGGATCACCAGGAGAG) (biomers.net GmbH, Ulm Germany). To 1 nM of PFV PR-RT or 10 nM of PFV RNase H-(Q<sup>591</sup>-N<sup>751</sup>) increasing amounts of the RNA/DNA substrate were added as indicated (Figure 3A), in a buffer containing 50 mM Tris/HCl pH 8.0, 80 mM KCl, 6 mM MgCl<sub>2</sub> and 0.5 mM DTT. An increase in fluorescence emission at 520 nm (excitation 495 nm) can be detected after cleavage of the RNA in the hybrid, which leads to dissociation of a small 6-FAM labeled RNA fragment from the dabcyI quencher. Fluorescence was measured on a Fluorolog-Tau-3 spectrofluorometer (HORIBA Jobin Yvon GmbH, Unterhaching, Germany). Initial cleavage rates were calculated using the linear slope of the reaction progress curve where less than 5% of substrate was cleaved. K<sub>M</sub> and v<sub>max</sub> values and their standard errors were obtained by fitting the data to the Michaelis-Menten equation using the fitting program GraFit 5.0 (Erithacus software).

### Qualitative RNase H assay

The RNA template (5'-CCGCUUCCCCUCUCCUG-GUGAUCCUU) (biomers.net GmbH, Ulm Germany) was purified on a denaturing 20% polyacrylamide/8 M urea gel. Ca. 30 pmol of the purified RNA was 5' end-labeled with 30 µCi γ[<sup>32</sup>P]-ATP (Hartmann Analytic, Braunschweig, Germany) and 20 U T4 polynucleotide kinase (New England Biolabs, Frankfurt, Germany) as described [43]. The DNA primer (5'-AGGATCACCAG-GAGAGGGGA) (biomers.net GmbH, Ulm Germany) was hybridized to the RNA using a 1.2 fold molar excess of the DNA in 50 mM Tris/HCl, pH 7.0, 80 mM KCl by



heating the sample to 90°C for 1 minute, followed by transfer to a heating block at 70°C and slow cooling to room temperature.

Qualitative RNase H reactions were performed at 25°C in a total volume of 25 µl in 50 mM Tris/HCl, pH 7.0, 80 mM KCl, 6 mM MgCl<sub>2</sub>, 0.5 mM DTT with 30 nM PR-RT or 20 µM RNase H-(Q<sup>591</sup>-N<sup>751</sup>), respectively. After preincubation of the sample for 5 min at 25°C, reactions were started by the addition of 240 nM RNA/DNA hybrid. Aliquots were taken after 10 min, 30 min and 60 min and stopped by the addition of an equal volume of urea loading buffer (50 mM EDTA, 0.1% xylene cyanol, 0.1% bromophenol blue, 8 M urea in 1 × TBE (Tris/Borate/EDTA)). Cleavage products were separated by denaturing gel electrophoresis using 15% polyacrylamide/8 M urea gels. After drying of the gel, the reaction products were visualized using a FLA 3000 phosphoimaging device (rayrest, Straubenhardt, Germany)

#### Fluorescence anisotropy measurements

Fluorescence equilibrium titrations were performed with the full length PR-RT and the RNase H domain employing a Fluorolog-Tau-3 spectrofluorometer (HORIBA Jobin Yvon GmbH, Unterhaching, Germany) to determine the dissociation constant ( $K_D$ ) for an RNA/DNA template/primer substrate labeled with the fluorescent dye DY647 as described previously [43]. For the PR-RT a 24/40-mer DNA/RNA primer/template substrate (DNA-5'-ATCACCAGGAGAGGGGAAAGCGGA; RNA-5'-DY647-CUAAUUCGCUUCCCCUCUC-CUGGUGAUCCUUCCAUC) (biomers.net GmbH, Ulm, Germany) was used as described [43]. However, for the RNase H domain a shorter 12/18-mer primer/template was chosen to avoid binding of more than one protein molecule to the hybrid region. The following sequences were used for the 12mer DNA primer 5'-ATCAGGTGTCGC and the fluorescently labeled 18mer RNA template 5'-DY647-AUUAUGCGACACCUGAU (biomers.net GmbH, Ulm, Germany). Titrations were performed at 25°C in 1 ml of buffer consisting of 50 mM Tris/HCl pH 8.0, 80 mM KCl, 10 mM EDTA and 0.5 mM DTT. Upon excitation of the substrate at 652 nm, emission can be detected at 673 nm. The slit widths were set at 8 nm for PR-RT, and at 7 nm (excitation) and 6 nm (emission) for the RNase H-(Q<sup>591</sup>-N<sup>751</sup>) domain. 5 nM of substrate were used for PR-RT and 50 nM for the RNase H-(Q<sup>591</sup>-N<sup>751</sup>), respectively. Figure 4 shows representative titration experiments. After equilibration of the sample for 3 minutes, at least six data points with an integration time of 1 s were collected for each titration point. The standard deviation for each data point is represented by the error bars in Figure 4.  $K_D$ -values and their standard errors were calculated by

non-linear curve fitting of the anisotropy data to a two component binding equation (equation 1) using the software GraFit 5.0 (Erithacus software).

equation 1:

$$A = \frac{A_S + P \cdot (R \cdot A_{ES} - A_S)}{1 - P + R \cdot P}$$

with

$$P = \frac{[ES]}{[S]_0} = \frac{K_D + [E]_0 + [S]_0 - \sqrt{(K_D + [E]_0 + [S]_0)^2 - 4 \cdot [S]_0 \cdot [E]_0}}{2 \cdot [S]_0}$$

where A = anisotropy measured;  $A_S$  = anisotropy of unbound substrate DNA/RNA;  $A_{ES}$  = anisotropy of the complex; [ES] = concentration of enzyme/substrate complex;  $[E]_0$  = total concentration of enzyme (PR-RT or RNase H-(Q<sup>591</sup>-N<sup>751</sup>));  $[S]_0$  = total concentration of substrate DNA/RNA;  $K_D$  = dissociation constant; R = ratio of fluorescence intensities of the bound and free substrate.

#### NMR analyses

1 mM uniformly <sup>15</sup>N labeled RNase H-(Q<sup>591</sup>-N<sup>751</sup>) was analyzed in 5 mM Na-phosphate pH 6.8 or 7.0, 100 mM NaCl, 0.5 mM DTT, 10% D<sub>2</sub>O (v/v) in the absence or presence of 6 mM MgCl<sub>2</sub>. For NMR resonance assignment standard double and triple resonance through-bond correlation experiments [46,47] were recorded using 1 mM <sup>15</sup>N, <sup>13</sup>C labeled RNase H-(Q<sup>591</sup>-N<sup>751</sup>), in 5 mM Na-phosphate pH 7.0, 100 mM NaCl, 6 mM MgCl<sub>2</sub>, 0.5 mM DTT, 10% D<sub>2</sub>O (v/v) at 25°C on a Bruker Avance 700 MHz spectrometer equipped with a cryogenically cooled probe. In-house protocols were used to process the NMR data and the program NMRview were utilized for analysis (B.A. Johnson, Merck, Whitehouse Station, NJ, USA).

#### Data deposition

The PFV RNase H-(Q<sup>591</sup>-N<sup>751</sup>) assignments have been deposited in the BioMagResBank, accession number: 17745

#### Abbreviations

E. coli: *Escherichia coli*; FV: foamy virus; GB1: immunoglobulin binding domain B1 of streptococcal protein G; HIV-1: human immunodeficiency virus type 1; IN: integrase; IPTG: isopropyl-thiogalactoside; PFV: prototype foamy virus; MoMLV: Moloney murine leukemia virus; PPT: polypurine tract; PR: protease; RHBD: RNA-hybrid binding domain; RNase H: ribonuclease H; RT: reverse transcriptase; TEV: tobacco etch virus;

#### Acknowledgements

We thank the DFG (Wo630/7-3) and the University of Bayreuth for financial support.

#### Authors' contributions

BMW conceived and coordinated the study. BL and FM performed the experiments. BL, BMW, MJH and KS analyzed the data. BMW, BL and KS wrote the paper. All authors read and approved the manuscript.

### Competing interests

The authors declare that they have no competing interests.

Received: 26 October 2011 Accepted: 10 February 2012

Published: 10 February 2012

### References

- Linial M: **Foamy viruses**. In *Fields Virology*. Edited by: Knipe DM, Howley PM. Lippincott Williams 2007:2245-2262.
- Rethwilm A: **Foamy viruses**. In *Topley & Wilson's Microbiology and Microbial Infections - Virology*. Edited by: ter Meulen V, Mahy BWJ. London: Hodder Arnold; 2005:1304-1321.
- Rethwilm A: **Molecular biology of foamy viruses**. *Med Microbiol Immunol* 2010, **199**:197-207.
- Hartl MJ, Mayr F, Rethwilm A, Wöhrl BM: **Biophysical and enzymatic properties of the simian and prototype foamy virus reverse transcriptases**. *Retrovirology* 2010, **7**:5.
- Hartl MJ, Bodem J, Jochheim F, Rethwilm A, Rösch P, Wöhrl BM: **Regulation of foamy virus protease activity by viral RNA - a novel and unique mechanism among retroviruses**. *J Virol* 2011, **85**:4462-4469.
- Moelling K, Bolognesi DP, Bauer H, Büsen W, Plassmann HW, Hausen P: **Association of retroviral reverse transcriptase with an enzyme degrading the RNA moiety of RNA-DNA hybrids**. *Nat New Biol* 1971, **234**:240-243.
- Gilboa E, Mitra SW, Goff S, Baltimore D: **A detailed model of reverse transcription and tests of crucial aspects**. *Cell* 1979, **18**:93-100.
- Repaske R, Hartley JW, Kavlick MF, O'Neill RR, Austin JB: **Inhibition of RNase H activity and viral replication by single mutations in the 3' region of Moloney murine leukemia virus reverse transcriptase**. *J Virol* 1989, **63**:1460-1464.
- Tisdale M, Schulze T, Larder BA, Moelling K: **Mutations within the RNase H domain of human immunodeficiency virus type 1 reverse transcriptase abolish virus infectivity**. *J Gen Virol* 1991, **72**:59-66.
- Schultz SJ, Champoux JJ: **RNase H activity: Structure, specificity, and function in reverse transcription**. *Virus Res* 2008, **134**:86-103.
- Pullen KA, Rattray AJ, Champoux JJ: **The sequence features important for plus strand priming by human immunodeficiency virus type 1 reverse transcriptase**. *J Biol Chem* 1993, **268**:6221-6227.
- Huber HE, Richardson CC: **Processing of the primer for plus strand DNA synthesis by human immunodeficiency virus 1 reverse transcriptase**. *J Biol Chem* 1990, **265**:10565-10573.
- Luo GX, Sharmeen L, Taylor J: **Specificities involved in the initiation of retroviral plus-strand DNA**. *J Virol* 1990, **64**:592-597.
- Wöhrl BM, Moelling K: **Interaction of HIV-1 ribonuclease H with polypurine tract containing RNA-DNA hybrids**. *Biochemistry (N Y)* 1990, **29**:10141-10147.
- Furfine ES, Reardon JE: **Human immunodeficiency virus reverse transcriptase ribonuclease H: Specificity of tRNA<sup>lys3</sup>-primer excision**. *Biochemistry (N Y)* 1991, **30**:7041-7046.
- Pullen KA, Ishimoto LK, Champoux JJ: **Incomplete removal of the RNA primer for minus-strand DNA synthesis by human immunodeficiency virus type 1 reverse transcriptase**. *J Virol* 1992, **66**:367-373.
- Smith JS, Roth MJ: **Specificity of human immunodeficiency virus-1 reverse transcriptase-associated ribonuclease H in removal of the minus-strand primer, tRNA(lys3)**. *J Biol Chem* 1992, **267**:15071-15079.
- Evans DB, Brawn K, Deibel MR Jr, Tarpley WG, Sharma SK: **A recombinant ribonuclease H domain of HIV-1 reverse transcriptase that is enzymatically active**. *J Biol Chem* 1991, **266**:20583-20585.
- Hostomsky Z, Hostomska Z, Hudson GO, Moomaw EW, Nodes BR: **Reconstitution in vitro of RNase H activity by using purified N-terminal and C-terminal domains of human immunodeficiency virus type 1 reverse transcriptase**. *Proc Natl Acad Sci USA* 1991, **88**:1148-1152.
- Smith JS, Roth MJ: **Purification and characterization of an active human immunodeficiency virus type 1 RNase H domain**. *J Virol* 1993, **67**:4037-4049.
- Smith JS, Gritsman K, Roth MJ: **Contributions of DNA polymerase subdomains to the RNase H activity of human immunodeficiency virus type 1 reverse transcriptase**. *J Virol* 1994, **68**:5721-5729.
- Tanese N, Goff SP: **Domain structure of the Moloney murine leukemia virus reverse transcriptase: Mutational analysis and separate expression of the DNA polymerase and RNase H activities**. *Proc Natl Acad Sci USA* 1988, **85**:1777-1781.
- Schultz SJ, Champoux JJ: **RNase H domain of Moloney murine leukemia virus reverse transcriptase retains activity but requires the polymerase domain for specificity**. *J Virol* 1996, **70**:8630-8638.
- Zhan X, Crouch RJ: **The isolated RNase H domain of murine leukemia virus reverse transcriptase. retention of activity with concomitant loss of specificity**. *J Biol Chem* 1997, **272**:22023-22029.
- Lim D, Gregorio GG, Bingman C, Martinez-Hackert E, Hendrickson WA, Goff SP: **Crystal structure of the Moloney murine leukemia virus RNase H domain**. *J Virol* 2006, **80**:8379-8389.
- Iwai S, Wakasa M, Ohtsuka E, Kanaya S, Kidera A, Nakamura H: **Interaction of the basic protrusion of *Escherichia coli* ribonuclease HI with its substrate**. *J Mol Biol* 1996, **263**:699-706.
- Katayanagi K, Miyagawa M, Matsushima M, Ishikawa M, Kanaya S, Nakamura H, Ikehara M, Matsuzaki T, Morikawa K: **Structural details of ribonuclease H from *Escherichia coli* as refined to an atomic resolution**. *J Mol Biol* 1992, **223**:1029-1052.
- Yang W, Hendrickson WA, Crouch RJ, Satow Y: **Structure of ribonuclease H phased at 2 Å resolution by MAD analysis of the selenomethionyl protein**. *Science* 1990, **249**:1398-1405.
- Nowotny M, Gaidamakov SA, Ghirlando R, Cerritelli SM, Crouch RJ, Yang W: **Structure of human RNase H1 complexed with an RNA/DNA hybrid: Insight into HIV reverse transcription**. *Mol Cell* 2007, **28**:264-276.
- Kanaya S, Katsuda-Nakai C, Ikehara M: **Importance of the positive charge cluster in *Escherichia coli* ribonuclease HI for the effective binding of the substrate**. *J Biol Chem* 1991, **266**:11621-11627.
- Telesnitsky A, Blain SW, Goff SP: **Defects in Moloney murine leukemia virus replication caused by a reverse transcriptase mutation modeled on the structure of *Escherichia coli* RNase H**. *J Virol* 1992, **66**:615-622.
- Lim D, Orlova M, Goff SP: **Mutations of the RNase H C helix of the Moloney murine leukemia virus reverse transcriptase reveal defects in polypurine tract recognition**. *J Virol* 2002, **76**:8360-8373.
- Davies JF, Hostomska Z, Hostomsky Z, Jordan SR, Matthews DA: **Crystal structure of the ribonuclease H domain of HIV-1 reverse transcriptase**. *Science* 1991, **252**:88-95.
- Julias JG, McWilliams MJ, Sarafianos SG, Alvord WG, Arnold E, Hughes SH: **Mutation of amino acids in the connection domain of human immunodeficiency virus type 1 reverse transcriptase that contact the template-primer affects RNase H activity**. *J Virol* 2003, **77**:8548-8554.
- Nowotny M, Gaidamakov SA, Crouch RJ, Yang W: **Crystal structures of RNase H bound to an RNA/DNA hybrid: Substrate specificity and metal-dependent catalysis**. *Cell* 2005, **121**:1005-1016.
- Das D, Georgiadis MM: **The crystal structure of the monomeric reverse transcriptase from Moloney murine leukemia virus**. *Structure* 2004, **12**:819-829.
- Pfepper KJ, Rackwitz HR, Schnölzer M, Heid H, Löchelt M, Flügel RM: **Molecular characterization of proteolytic processing of the pol proteins of human foamy virus reveals novel features of the viral protease**. *J Virol* 1998, **72**:7648-7652.
- Boyer PL, Gao HQ, Frank P, Clark PK, Hughes SH: **The basic loop of the RNase H domain of MLV RT is important both for RNase H and for polymerase activity**. *Virology* 2001, **282**:206-213.
- Yang W, Lee JY, Nowotny M: **Making and breaking nucleic acids: Two-Mg<sup>2+</sup>-ion catalysis and substrate specificity**. *Mol Cell* 2006, **22**:5-13.
- Steitz TA, Steitz JA: **A general two-metal-ion mechanism for catalytic RNA**. *Proc Natl Acad Sci USA* 1993, **90**:6498-6502.
- Huber HE, McCoy JM, Seehra JS, Richardson CC: **Human immunodeficiency virus 1 reverse transcriptase. Template binding, processivity, strand displacement synthesis, and template switching**. *J Biol Chem* 1989, **264**:4669-4678.
- DeStefano JJ, Bambara RA, Fay PJ: **Parameters that influence the binding of human immunodeficiency virus reverse transcriptase to nucleic acid structures**. *Biochemistry (N Y)* 1993, **32**:6908-6915.
- Hartl MJ, Kretzschmar B, Frohn A, Nowrouzi A, Rethwilm A, Wöhrl BM: **AZT resistance of simian foamy virus reverse transcriptase is based on the excision of AZTMP in the presence of ATP**. *Nucleic Acids Res* 2008, **36**:1009-1016.
- Sambrook J, Fritsch EF, Maniatis T: **Molecular cloning - A laboratory manual**. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY; 1994.
- Hartl MJ, Wöhrl BM, Rösch P, Schweimer K: **The solution structure of the simian foamy virus protease reveals a monomeric protein**. *J Mol Biol* 2008, **381**:141-149.

46. Bax A, Grzesiek A: Methodological advances in protein NMR. *Acc Chem Res* 1993, **26**:131-138.
47. Sattler M, Schleucher J, Griesinger C: Herteronuclear multidimensional NMR experiments for the structure determination of proteins in solution empolying pulsed field gradients. *Prog NMR Spectrosc* 1999, **34**:93-158.

doi:10.1186/1742-4690-9-14

**Cite this article as:** Leo *et al.*: Insights into the structure and activity of prototype foamy virus RNase H. *Retrovirology* 2012 **9**:14.

**Submit your next manuscript to BioMed Central  
and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

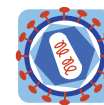




## 8 Einzelarbeit B

Berit Leo, Kristian Schweimer, Paul Rösch, Maximilian J. Hartl und Birgitta M. Wöhrl (2012): The solution structure of the prototype foamy virus RNase H domain indicates an important role of the basic loop in substrate binding. *Retrovirology* **9**:73.





RESEARCH

Open Access

# The solution structure of the prototype foamy virus RNase H domain indicates an important role of the basic loop in substrate binding

Berit Leo, Kristian Schweimer, Paul Rösch, Maximilian J Hartl and Birgitta M Wöhrl\*

## Abstract

**Background:** The ribonuclease H (RNase H) domains of retroviral reverse transcriptases play an essential role in the replication cycle of retroviruses. During reverse transcription of the viral genomic RNA, an RNA/DNA hybrid is created whose RNA strand needs to be hydrolyzed by the RNase H to enable synthesis of the second DNA strand by the DNA polymerase function of the reverse transcriptase. Here, we report the solution structure of the separately purified RNase H domain from prototype foamy virus (PFV) revealing the so-called C-helix and the adjacent basic loop, which both were suggested to be important in substrate binding and activity.

**Results:** The solution structure of PFV RNase H shows that it contains a mixed five-stranded  $\beta$ -sheet, which is sandwiched by four  $\alpha$ -helices (A-D), including the C-helix, on one side and one  $\alpha$ -helix (helix E) on the opposite side. NMR titration experiments demonstrate that upon substrate addition signal changes can be detected predominantly in the basic loop as well as in the C-helix. All these regions are oriented towards the bound substrate. In addition, signal intensities corresponding to residues in the B-helix and the active site decrease, while only minor or no changes of the overall structure of the RNase H are detectable upon substrate binding. Dynamic studies confirm the monomeric state of the RNase H domain. Structure comparisons with HIV-1 RNase H, which lacks the basic protrusion, indicate that the basic loop is relevant for substrate interaction, while the C-helix appears to fulfill mainly structural functions, i.e. positioning the basic loop in the correct orientation for substrate binding.

**Conclusions:** The structural data of PFV RNase H demonstrate the importance of the basic loop, which contains four positively charged lysines, in substrate binding and the function of the C-helix in positioning of the loop. In the dimeric full length HIV-1 RT, the function of the basic loop is carried out by a different loop, which also harbors basic residues, derived from the connection domain of the p66 subunit. Our results suggest that RNases H which are also active as separate domains might need a functional basic loop for proper substrate binding.

**Keywords:** Foamy virus, Retroviral RNase H, C-helix, Basic loop, NMR, Basic protrusion, Solution structure, Substrate binding

## Background

RTs of retroviruses are multifunctional enzymes, which in addition to a polymerase domain harbor a C-terminal RNase H domain whose activity is essential during reverse transcription. The RNase H function is necessary for degrading the RNA in the emerging RNA/DNA hybrid [1] to allow synthesis of the second DNA strand. In contrast to orthoretroviruses, spumaretroviruses, also

called foamy viruses (FVs), possess a mature protease-reverse transcriptase (PR-RT) in the virion, whereas all other retroviruses cleave off the PR domain from the Pol precursor [2,3]. Similarly to the RT from Moloney murine leukemia virus (MoMLV) [4], the PR-RT of the prototype foamy virus (PFV) is also monomeric [3].

We have shown previously, that the isolated RNase H domain of PFV exhibits activity, albeit to a much lower extent than the mature full length PR-RT [5]. This behavior is comparable to that of the separate RNase H domain of MoMLV [6-8]. Furthermore, NMR spectroscopy of PFV RNase H revealed the existence of the so-

\* Correspondence: birgitta.woehrl@uni-bayreuth.de  
Universität Bayreuth, Lehrstuhl Biopolymere, Universitätsstr. 30, D-95447  
Bayreuth, Germany

called “basic protrusion” which consists of an  $\alpha$ -helix, referred to as the “C-helix” followed by a basic loop [5]. The basic protrusion of RNases H has been implicated in substrate binding and activity [9-11] and has also been found in the retroviral RNase H domain of MoMLV [12] and in the recently solved x-ray structure of xenotropic murine leukemia virus-related virus (XMRV) [13]. XMRV is highly related to MoMLV and appears to have arisen through recombination events during passaging of human tumors in mice [14].

A basic protrusion can also be found in the human and *Escherichia coli* (*E. coli*) RNases H, whereas the RNase H from *Bacillus halodurans* (*B. halodurans*) contains only the basic loop but no C-helix. However, an N-terminal RNA/DNA-hybrid binding domain (RHBD) compensates for the lack of the helix [15]. In contrast, HIV-1 RNase H, which is not active when expressed in isolation [16-20], does not harbor the basic protrusion, rather, positively charged residues in the connection domain assume this function.

Here, we show the solution structure of PFV RNase H and titration experiments with an RNA/DNA substrate indicating an important role of the basic loop in substrate binding.

## Results and discussion

To determine the solution structure of the wild type (wt) PFV RNase H we purified the recombinant  $^{15}\text{N}$  or  $^{15}\text{N}$ ,  $^{13}\text{C}$  labeled protein PFV RNase H-(Q<sup>591</sup>-N<sup>751</sup>) from *E. coli*. We have shown previously, that the purified domain exhibits RNase H activity and possesses a defined tertiary structure including the C-helix [5].

### Structure of the isolated PFV RNase H

To verify the oligomerization state of the PFV RNase H domain  $^{15}\text{N}$  spin relaxation experiments were performed. Comparison of  $^{15}\text{N}$  transverse ( $R_2$ ) and longitudinal ( $R_1$ ) relaxation rates at two different concentrations (1 mM and 0.4 mM) reveals deviations of less than 6 % for  $R_2$  and 3 % for  $R_1$ , indicating that no aggregation is observed under the conditions applied for NMR spectroscopy. From the relaxation rates, a rotational correlation time ( $t_c$ ) of 12.4 ns was determined for an isotropic model. In addition, the data were fitted to an axial symmetric as well as a fully asymmetric rotational diffusion tensor using the coordinates from the lowest energy structure, but no further statistical improvement was obtained. The higher  $t_c$  of PFV RNase H as compared to the isolated RNase H domain of HIV-RT (9.7 ns, [21]) reflects the larger size of the PFV RNase H (165 aa vs. 138 aa, 18 kDa vs. 15.3 kDa), and confirms the monomeric character of the domain already demonstrated by size exclusion chromatography [5].

Sequence comparisons (Figure 1) indicated that similar to MoMLV and XMRV RNase H, PFV RNase H contains a basic protrusion, including the so-called C-helix followed by a basic loop structure. We have shown previously that the purified, separate PFV RNase H domain is weakly active [5]. Structural analyses by NMR spectroscopy revealed the presence of a structured C-helix [5]. Thus, we set out to determine the solution structure of PFV RNase H (Figure 2).

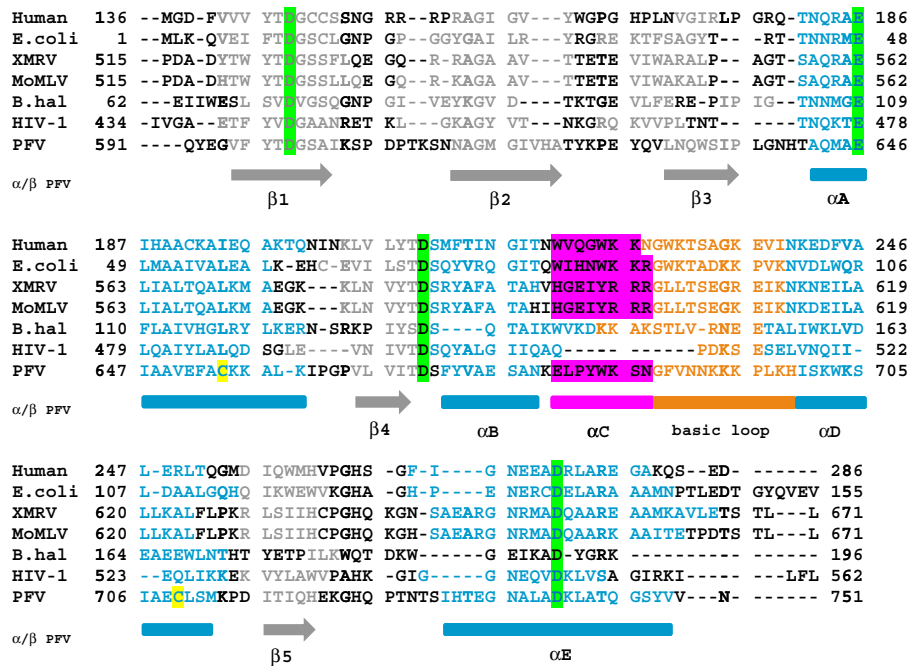
Using multidimensional heteronuclear edited NOESY experiments, 1686 distance restraints together with 66 restraints for hydrogen bonds and 225 dihedral angle restraints based on chemical shift analyses could be derived (Table 1). The final structure calculation resulted in an ensemble of structures with no distance restraint violation larger than 0.1 Å and no dihedral restraint violation larger than 2.4° together with good stereochemical properties (Table 1). The structures superimpose well in the secondary structure regions (Figure 2A). For several loop regions no distance restraints could be derived because of unassigned residues [5]. Their NMR signals are most likely broadened beyond detection due to exchange processes. Therefore the large coordinate variability reflects the dynamic character of these regions.

The structure of PFV RNase H shows the typical tertiary fold of an RNase H domain, which is formed by a five-stranded mixed  $\beta$ -sheet flanked by five  $\alpha$ -helices. The long carboxy-terminal helix E packs on one side of the  $\beta$ -sheet, and the helices A, B, C and D are located on the other side of the  $\beta$ -sheet (Figure 2B). Due to a lack of structural restraints accessible by NMR, no attempts were made to characterize the active site. Despite the absence of certain restraints in this region, the active site residues (D599, E646, D669, D740) are closely enough together to allow coordination of magnesium ions.

A characteristic feature of PFV RNase H is the presence of helix C which directly follows helix B after a kink. Helix C precedes a basic loop containing three neighboring lysines. Helix C as well as the adjacent loop is also found in XMRV RNase H. However, in XMRV RNase H the consecutive basic residues (3 x Arg) are part of helix C [13], whereas in PFV RNase H four lysines (KKKPLK) are located in the adjacent basic loop (Figure 1). The slightly different position of the basic residues might cause different distances to the active site and might thus contribute to possible differences in cleavage activities of the two enzymes [5,13,22]. The orientation of helix C is determined by numerous hydrophobic contacts to helix D. This suggests that helix C acts as a molecular ruler to orient the basic loop in order to optimize substrate binding and specificity.

The structure of PFV RNase H further revealed that two cysteine residues, Cys654 and Cys709 (Figure 1), located

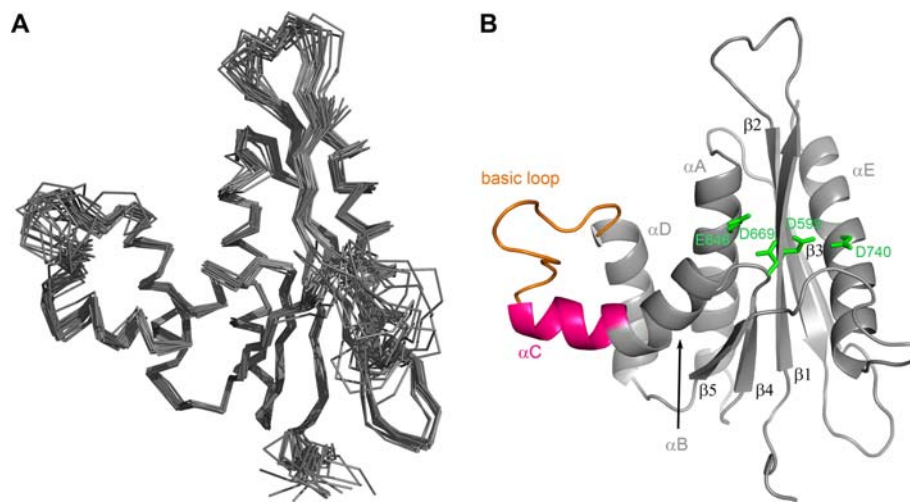




**Figure 1** Sequence alignment of different RNase H domains. The RNases H from human origin, *E. coli*, XMRV, MoMLV, *B. haludorans*, HIV-1 and PFV are shown. The numbers represent the amino acid numbers of the corresponding full length enzymes. The alignment was performed with the program CLC Protein Workbench 5 (version 5.3). The arrows and boxes below indicate the secondary structures determined for PFV RNase H;  $\alpha$ -helices are shown in blue,  $\beta$ -strands in gray. The active site residues are highlighted in bright green; the cysteines (C654, C709) of PFV RNase H are marked in yellow. The C-helices and basic loops of the enzymes are highlighted in magenta and orange, respectively.

in helix A and D, respectively, are facing each other, thus allowing the potential formation of a disulfide bridge (Figure 3A). To find out whether this disulfide bridge is important for RNase H function, we mutated Cys654 to serine to abolish disulfide formation. However, analyses of the RNase H activities of the mutant RNase H-(C654S)

and the wt enzyme revealed no differences in activity (Figure 3B), indicating that creation of the disulfide bridge is not relevant for RNase H activity. Sequence comparisons of PFV RNase H with other foamy virus RNases H (feline, bovine) showed, that the cysteines are not conserved. Only the closely related RNase H from simian FV



**Figure 2** Solution structure of PFV RNase H. (A) Superposition of the 19 lowest energy structures. (B) Ribbon diagram of the lowest energy structure. The C-helix is highlighted in magenta, and the basic loop in orange. The active site residues D599, E646, D669 and D740 are highlighted in bright green.

**Table 1 Solution structure statistics**

| <b>Experimentally derived restraints</b>                             |                     |
|--|---------------------|
| distance restraints  |                     |
|  | NOE 1686            |
|  | intraresidual 527   |
|  | sequential 343      |
|  | medium range 245    |
|  | long range 505      |
|  | ambiguous 66        |
|  | hydrogen bonds 62   |
| dihedral restraints 225  |                     |
| <b>Restraint violation</b>   |                     |
| average distance restraint violation (Å)                             | 0.0035 +/- 0.0008   |
| maximum distance restraint violation (Å)                             | < 0.1               |
| average dihedral restraint violation (°)                             | 0.10 +/- 0.04       |
| maximum dihedral restraint violation (°)                             | 2.4                 |
| <b>deviation from ideal geometry</b>                                 |                     |
| bond length (Å)  | 0.00064 +/- 0.00005 |
| bond angle (Å)   | 0.12 +/- 0.006      |
| <b>Coordinate precision<sup>a,b</sup></b>                            |                     |
| backbone heavy atoms (Rmsd) (Å) (all / defined secondary structure ) | 1.05 / 0.46         |
| all heavy atoms (Rmsd) (Å)   | 1.58 / 0.92         |
| <b>Ramachandran plot statistics<sup>c</sup> (%)</b>                  |                     |
|  | 89.7/9.6/0.3/0.4    |

<sup>a</sup> The precision of the coordinates is defined as the average atomic root mean square difference between the accepted simulated annealing structures and the corresponding mean structure calculated for the given sequence regions.

<sup>b</sup> calculated for residues 593–747 (all) or 594–604, 612–621, 629–637, 642–659, 664–668, 671–686, 700–712, 717–720, 731–747 (defined secondary structure).

<sup>c</sup> Ramachandran plot statistics are determined by PROCHECK and noted by most favored/additionally allowed/generously allowed/disallowed.

from macaques (SFVmac) retains the cysteines (data not shown). Other RNases H also do not contain two cysteines facing each other (Figure 1).

### Substrate binding

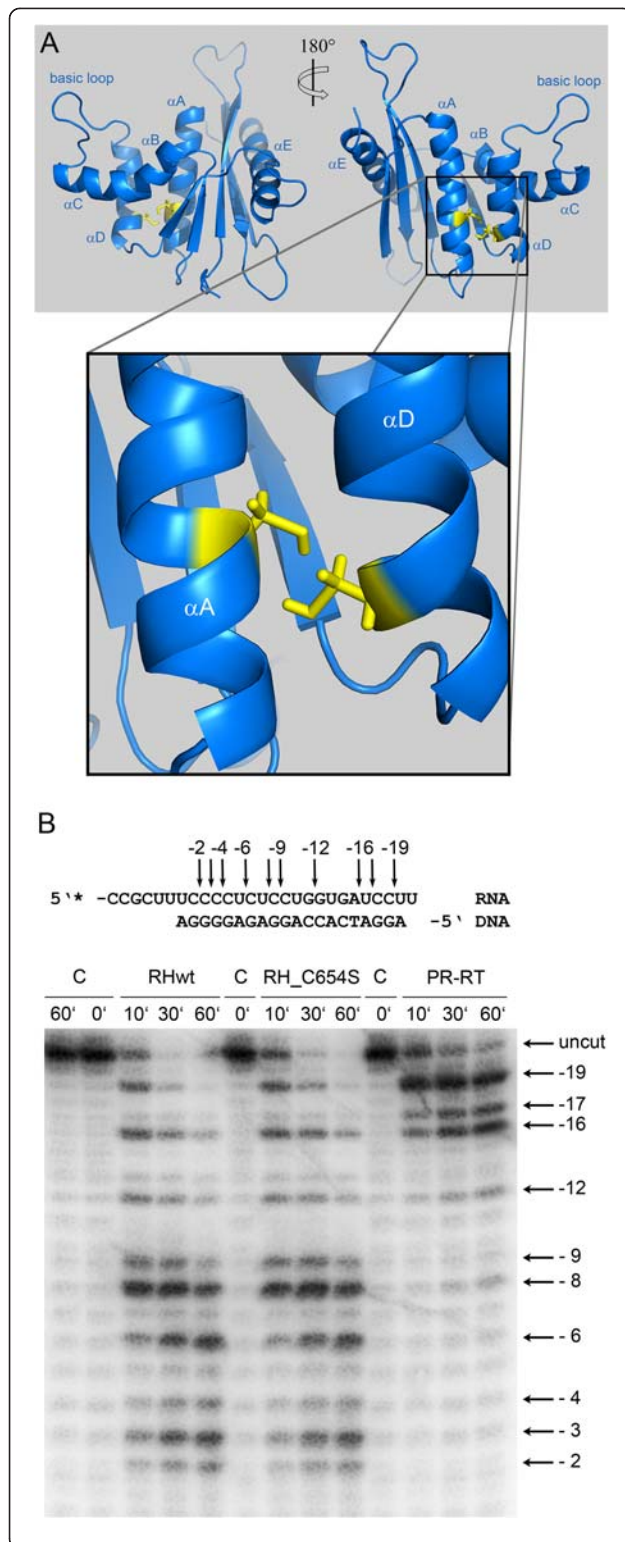
To analyze the interaction of PFV RNase H with an RNA/DNA substrate, we had to abolish RNase H activity to avoid cleavage of the substrate during long-term NMR analyses. Thus, we mutated the active site residue Asp599 and, in addition, residue His724, located in the loop region between  $\beta$ -strand 5 and helix E to asparagine. The corresponding histidine residue in HIV-1 RNase H had been shown to severely reduce RNase H activity [23,24]. The RNase H activity of the double mutant D599N-H742N was then analyzed qualitatively with the same 10/10-mer RNA/DNA substrate, which we wanted to use further in NMR analyses. Our results demonstrate that the RNase H activity of the mutant is seriously impaired (Figure 4A). However, only minor changes were observed in the recorded NMR spectra, indicating that the overall tertiary structure of the RNase H domain is unaffected by the amino acid exchanges (data not shown).

Furthermore, we introduced the two amino acid exchanges into the full length PR-RT to determine the impact of the mutations in DNA polymerization. Analysis of the polymerization activities on a homopolymeric substrate resulted in similar specific activities of 8.2 U and 8.8 U (per  $\mu$ g of enzyme) for the wt and mutant PR-RT-(D599N-H742N), respectively, revealing that the mutations do not impair polymerase activity and overall structure of PFV PR-RT.

To determine the binding affinity of the mutant RNase H-(D599N-H742N) we used the same 10/10-mer substrate, now labeled with the fluorescence marker DY647 to allow fluorescence anisotropy measurements [5]. Despite of its low RNase H cleavage activity, the mutant revealed comparable binding affinities for the RNA/DNA substrate ( $K_D \approx 31 \mu\text{M} \pm 5$ ) as the wt enzyme ( $K_D \approx 23 \mu\text{M} \pm 3$ ) (Figure 4B). This characteristic made the double mutant D599N-H742N suitable for NMR titration experiments.

### NMR titration experiments

<sup>1</sup>H-<sup>15</sup>N-HSQC spectra of 50  $\mu\text{M}$  mutant PFV RNase H-(D599N-H742N) were recorded using increasing molar

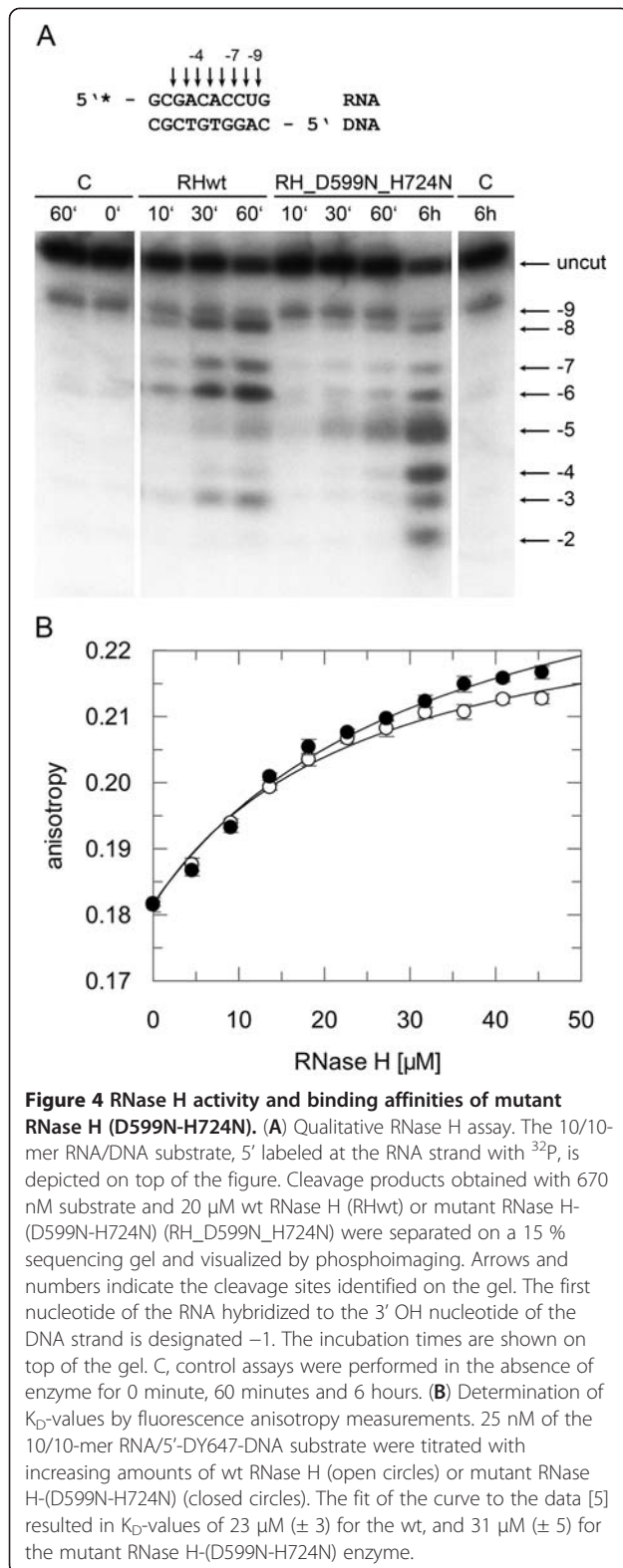


**Figure 3 Orientation of Cys654 and Cys709 in PFV RNase H. (A)**

Enlargement of  $\alpha$ -helices A and D of PFV RNase H. The blow-up shows the structural elements  $\alpha$ A and  $\alpha$ D, harboring residues C654 and C709, respectively. **(B)** Qualitative RNase H activity assay. Comparison of RNase H mutant C654S (RH\_C654S) with the wt RNase H (RHwt) and the full length PR-RT. Reaction products were separated on a denaturing sequencing gel and visualized by phosphorimaging. The 27/20-mer RNA/DNA substrate,  $^{32}$ P-labeled at the 5' end of the RNA, is shown on top of the gel. Arrows and numbers indicate the cleavage sites identified on the sequencing gel.

ratios of RNA/DNA:protein of 0, 1, 2, 3, 4 and 5. Upon substrate addition a decrease of the signal intensities of several residues (G600, A614, T641, Q643, D669, F671, A674, E675, S676, E680, K685, S686, N687, F689, V690, N691, H699, I700) could be observed. A signal change was considered relevant, if the signal decreased to 50 % or less of the original value. The residues potentially involved in substrate binding are highlighted in orange in Figure 5. Apart from the active site residues, helix B and C as well as the basic loop appear to be involved in substrate binding. The similarity of PFV RNase H and human RNase H allowed the generation of a simple model of PFV RNase H in complex with substrate. Thus, we superimposed the PFV RNase H structure with that of the human RNase H/substrate complex (PDB: 2QK9) to determine whether the residues, which show significant effects in the NMR titration experiments, are located spatially close to the substrate (Figure 5A). The alignment indicates a potential role of helix C in positioning the basic loop close to the DNA strand, thus supporting substrate binding, whereas the active site residues are close to the RNA strand, which has to be cleaved during catalysis.

Further information on substrate binding is provided by the alignment of the complex shown in Figure 5A with HIV-1 RT (PDB: 1HYS) in complex with a polypurine tract (PPT) containing RNA/DNA hybrid [25]. Only the PPT hybrid substrate modeled onto the PFV RNase H is shown (Figure 5B). The modeling reveals that the PPT RNA of the hybrid is not as close to the active site as the RNA strand from the other substrate depicted in Figure 5A. This is not surprising since the PPT-RNA/DNA hybrid, which is not cleaved by the RNase H during reverse transcription, adopts a peculiar structure in the complex with HIV-1 RT [25]. In the PPT region, which is rich in adenines, an unpaired base on the DNA strand takes the base pairing out of register leading to two offset base pairs. Then, an unpaired base on the RNA strand re-establishes the normal register. This structural deviation of PPT containing substrates was suggested to play a role in the resistance of the PPT to RNase H cleavage [25].



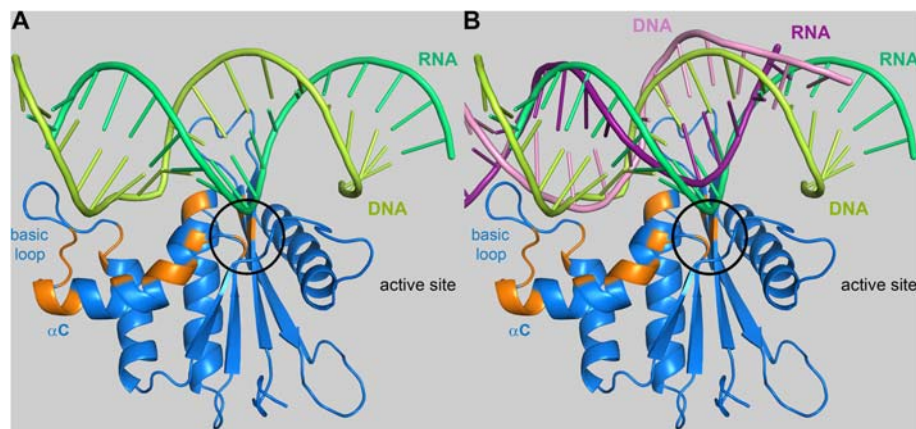
A model of XMRV RNase H complexed with substrate, published recently, suggests that several aromatic residues (Y586, H594 and Y598) located in helices B and C are capable of hydrogen-bond or stacking interactions with the substrate [13]. The corresponding residues in PFV RNase H are Y672 (located in helix B), K679 (located between helix B and C) and Y683 (located in helix C). Since these residues are solvent exposed in PFV RNase H, substrate interaction might also be possible. These results suggest that monomeric RNase H domains that are active in isolation might need the basic protrusion to confer sufficient affinity for the substrate to permit cleavage.

To further analyze the structural features of PFV RNase H we performed structural comparisons with the RNases H of *B. halodurans* [15], human origin [26], XMRV [13] as well as HIV-1 [20] (Figure 6). The structure of PFV RNase H most closely resembles those of XMRV and HIV-1, followed by human RNase H, while the orientation of the basic loop in *B. halodurans* RNase H is strikingly different (Figure 6C). However, similar to HIV-1 RNase H (Figure 6D), *B. halodurans* RNase H also lacks the C-helix, but harbors an additional N-terminal extension instead, which was shown to be involved in substrate binding [15]. Superposition of PFV RNase H with HIV-1 RNase H (Figure 6D) demonstrates that due to the lack of the C-helix in HIV-1 the adjacent loop is too short and is not positioned correctly for substrate binding.

To obtain a clearer picture of the orientation of these loops, we superimposed the structure of the full length HIV-1 RT in complex with the PPT containing hybrid (PDB: 1HYS) with PFV RNase H (Figure 6E). The enlargement of the region where the basic protrusion of PFV RNase H is located clearly shows that the basic loop is close to the substrate, whereas the corresponding loop of HIV-1 RNase H is much too short to contact the nucleic acid. Rather, a different loop derived from the connection domain (R356 – N363) of the p66 subunit of the heterodimeric HIV-1 RT appears to fulfill this function [27]. Similarly to the basic loop of PFV RNase H, which contains several lysine residues, this loop of HIV-1 RT also contains basic residues (R356, R358).

## Conclusions

In summary, we were able to identify regions in the RNase H domain - and by structural alignments also in the connection domain of HIV-1 RT - that interact with the substrate. Our data suggest that the basic loop plays an important role in substrate binding and that the major role of the C-helix is to orientate the basic loop properly towards the substrate. Furthermore, regions in the B-helix, which are also present in HIV-1 RNase H, have been identified to be involved in substrate binding.



**Figure 5 Modeling of an RNA/DNA substrate onto PFV RNase H.** Ribbon diagram of mutant RNase H-(D599N-H724N). The regions highlighted in orange exhibited significant signal changes upon substrate addition. **(A)** Based on the structure of the human RNase H/substrate complex (PDB: 2QK9) a structural alignment of the PFV and human RNase H was performed to model the RNA/DNA substrate of the human RNase H onto PFV RNase H (Rmsd: 2.35 Å). The RNA strand is shown in dark green, the DNA strand in light green. **(B)** Alignment of the complex shown in **(A)** with the HIV-1 RT/RNA-DNA complex (PDB: 1HYS) harboring the PPT RNA-DNA. The PPT substrate is depicted in dark (RNA) and light purple (DNA). The location of the active site residues of PFV RNase H is indicated by a black circle.

Together, this information might lead to a better understanding of the mechanism of action of HIV RNase H inhibitors and can contribute to the design and development of more specific and more potent inhibitors.

## Methods

### Mutant plasmid construction

The plasmid pET-GB1a-RNase H-(Q<sup>591</sup>-N<sup>751</sup>) coding for the gene of the wt PFV RNase H-(Q<sup>591</sup>-N<sup>751</sup>) as described previously [5] was used as a template for site-directed mutagenesis to obtain the mutant proteins RNase H-(C654S) and RNase H-(D599N-H724N) using the following primers:

- (i) mutation C654S:  
5'-GCTGCAGTTGAATTTGCCAGTAAAAAAG-  
CTTTAAAAATACCTGG;  
5'-CCAGGTATTTTTAAAGCTTTTTTACTGGC-  
AAATTCAACTGCAGC;
- (ii) mutation D599N:  
5'-  
GAAGGAGTGTTTTATACTAATGGCTCGGCC-  
ATCAAAAAGTCC  
5'-  
GGACTTTTGATGGCCGAGCCATTAGTATAA-  
AACACTCCTTC
- (iii) mutation H724N:  
5'-CTATTC AACATGAAAAAGGGAATCAGCCT-  
ACAAATACCAGTAT  
5'-  
ATACTGGTATTTGTAGGCTGATTCCTTTTT-  
CATGTTGAATAG.

Point mutations were introduced according to the QuikChange kit (Stratagene, Waldbronn, Germany) and confirmed by sequencing (LGC Genomics, Berlin, Germany). Mutations coding for the amino acid exchanges D599N and H724N were introduced into the wt full length PR-RT using the same primers as for the mutation of the isolated RNase H domain.

### Gene expression and protein purification

Protein purifications of PFV RNase H wt, RNase H-(C654S) and RNase H-(D599N-H724N) were performed as described previously [5]. PFV PR-RT(D599N-H724N) as well as <sup>15</sup>N-labeled mutant RNase H-(D599N-H724N) and <sup>15</sup>N,<sup>13</sup>C labeled wt RNase H used in NMR studies were also purified as described previously [3,5].

### Qualitative RNase H assay

Reactions were performed as described previously [5] with 20 μM wt RNase H, RNase H-(C654S) and RNase H-(D599N-H724N), respectively, using a 10/10-mer or a 27/20-mer RNA/DNA substrate, whose RNA was 5' end labeled with <sup>32</sup>P [5]. The sequence of the 10-mer RNA was 5'-GCGACACCUG, the 10-mer DNA sequence was complementary to it. Sequences of the 27/20-mer RNA/DNA hybrid have been published previously [5]. After preincubation of the sample for 5 minutes at 25°C, reactions were started by the addition of ca. 670 nM 10/10-mer or 240 nM 27/20-mer RNA/DNA hybrid. Aliquots were taken as indicated in the figures and processed further as described previously [5].

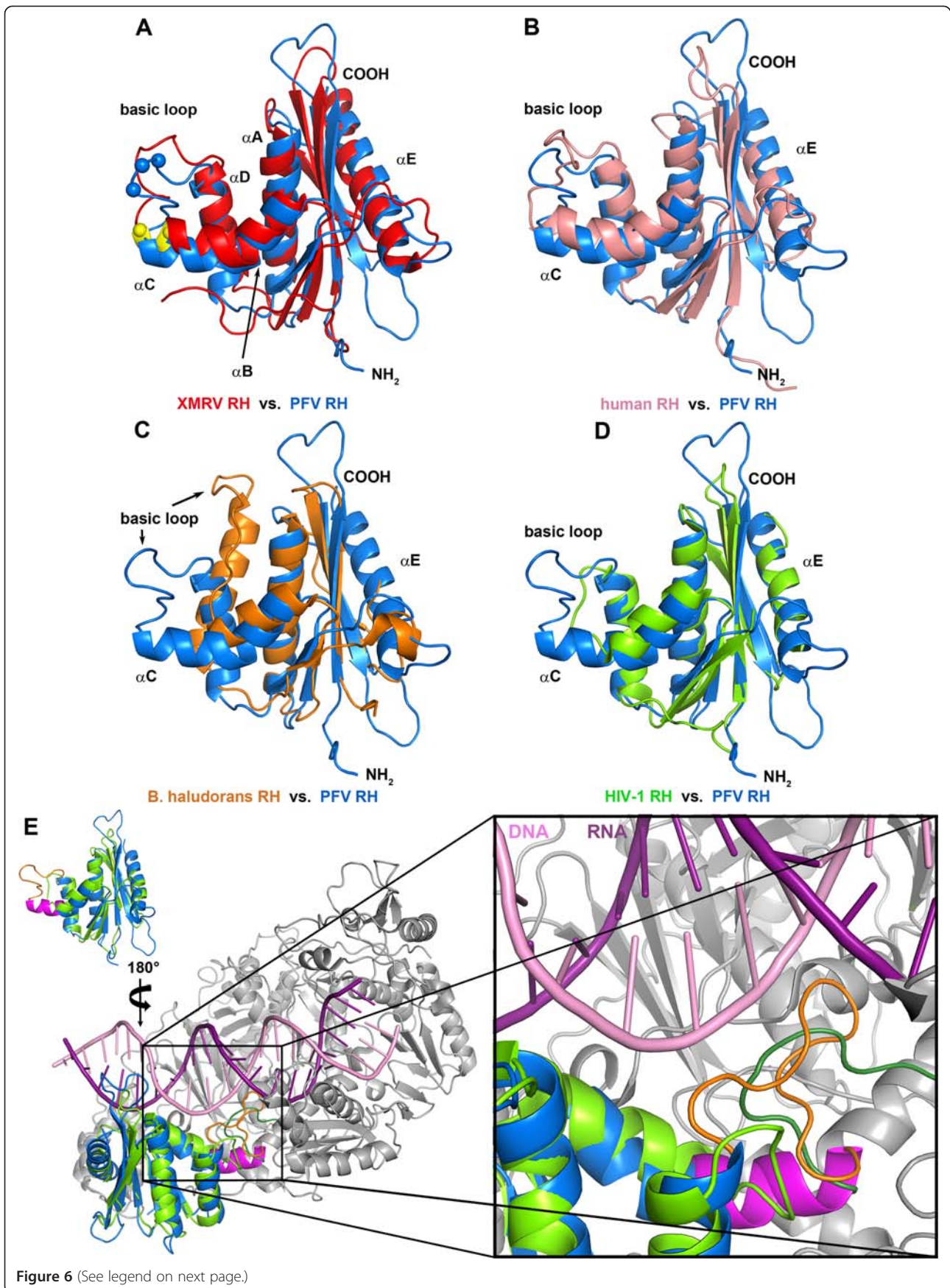


Figure 6 (See legend on next page.)

(See figure on previous page.)

**Figure 6 Structural overlay of PFV RNase H with different RNase H structures and HIV-1 RT. (A – D)** Superposition of the lowest energy structure of PFV RNase H with different RNases H. **(A)** XMRV (PDB: 3V1O, Rmsd: 2.10 Å), the C- $\alpha$  positions of the three positively charged arginine (XMRV) or lysine (PFV) residues in the basic protrusion are highlighted as yellow (XMRV) or blue (PFV) spheres, **(B)** human (PDB: 2QK9, Rmsd: 2.35 Å), **(C)** *B. halodurans* (PDB: 1ZBF, Rmsd: 2.37 Å), and **(D)** HIV-1 (PDB: 1HRH, Rmsd: 1.90 Å). **(E)** Orientation of PFV and HIV-1 RNase H in the full length HIV-1 RT. Superposition of the RNase H domains from PFV (blue) (PDB: 2LSN) and HIV-1 (green) (PDB: 1HYS) in the full length HIV-1 RT (gray) (PDB: 1HYS) in complex with a DNA/RNA substrate (Rmsd: 1.85 Å). With regard to **(D)** the RNase H domains are turned vertically by 180°. The enlargement shows the region harboring the C-helix (magenta) and the basic loop (orange) of PFV RNase H and the corresponding loop (dark green) derived from the connection domain of the p66 subunit of the heterodimeric HIV-1 p66/p51 RT. The RNA strand of nucleic acid substrate is colored in dark purple, the DNA strand is in light purple.

### Fluorescence anisotropy measurements

Fluorescence equilibrium titrations were performed as described [5] to determine the dissociation constant ( $K_D$ ) of wt RNase H and mutant RNase H-(D599N-H724N) employing 25 nM of the 10/10-mer RNA/DNA substrate used above for the qualitative RNase H activity assay, however with the fluorescence label DY647 attached at the 5' end of the DNA (DNA-5'-DY647-CAGGTGTCGC) (biomers.net GmbH, Ulm, Germany). The excitation wavelength was 652 nm and emission was detected at 673 nm. The slit widths were set at 8 nm and 7 nm for excitation and emission, respectively. The standard deviation for each data point is represented by error bars in Figure 4B. Calculation of the  $K_D$ -values was performed as described previously [5].

### NMR analyses

NMR samples containing 0.4 – 1.0 mM uniformly  $^{15}\text{N}$  or  $^{13}\text{C}$ ,  $^{15}\text{N}$  labeled RNase H were analyzed in 5 mM Na-phosphate, pH 7.0, 100 mM NaCl, 0.5 mM DTT, 10 %  $\text{D}_2\text{O}$  (v/v), 6 mM  $\text{MgCl}_2$ . All NMR experiments were conducted at 298 K on a Bruker Avance 600 MHz or a Bruker Avance 700 MHz (equipped with a cryogenic probe) spectrometer. In addition to the previously described experiments for resonance assignments [5]  $^{15}\text{N}$ - and  $^{13}\text{C}$ -edited NOESY experiments (mixing time 120 ms) were acquired for obtaining distance restraints. Standard  $^{15}\text{N}$  spin relaxation experiments [28,29] were conducted at 600 MHz  $^1\text{H}$  NMR frequency for obtaining transverse and longitudinal  $^{15}\text{N}$  relaxation rates.

For NMR titration experiments a 1.2 fold molar excess of the non-labeled 10-mer DNA was hybridized to the complementary non-labeled 10-mer RNA in 50 mM Na-phosphate, pH 7.0, 100 mM NaCl by heating the sample to 95°C for 2 minutes, followed by transfer to a heating block at 70°C and slow cooling to room temperature. 50  $\mu\text{M}$   $^{15}\text{N}$ -labeled mutant RNase H-(D599N-H724N) was dissolved in 50 mM Na-phosphate, pH 7.0, 100 mM NaCl, 6 mM  $\text{MgCl}_2$ , 2 mM DTT, 10 %  $\text{D}_2\text{O}$  (v/v). Different molar ratios of RNA/DNA:protein of 0, 1, 2, 3, 4 and 5 were employed.

### Structure calculation

Distance restraints for structure calculation were derived from  $^{15}\text{N}$ -edited NOESY and  $^{13}\text{C}$ -edited NOESY spectra.

NOESY cross peaks were classified according to their relative intensities and converted to distance restraints with upper limits of 3.0 Å (strong), 4.0 Å (medium), 5.0 Å (weak), and 6.0 Å (very weak). For ambiguous distance restraints the  $r^{-6}$  summation over all assigned possibilities defined the upper limit. Experimental NOESY spectra were validated semi-quantitatively against back-calculated spectra to confirm the assignment and to avoid bias of upper distance restraints by spin diffusion. Dihedral restraints were taken from analysis of chemical shifts by the TALOS software package [30]. Hydrogen bonds were included for backbone amide protons in regular secondary structure, when the amide proton did not show a water exchange cross peak in the  $^{15}\text{N}$ -edited NOESY spectrum.

The structure calculations were performed with the program XPLOR-NIH 1.2.1 ([31]) using a three-step simulated annealing protocol with floating assignment of prochiral groups including a conformational database potential. The 19 structures (out of total 120 structures) showing the lowest values of the target function excluding the database potential were further analyzed with XPLOR, PROCHECK [32] and pyMOL [33].

### Polymerization assay

RNA-dependent DNA polymerase activities of PFV PR-RT and PFV PR-RT(D599N-H724N) were quantitated at 37°C on a poly(rA)/oligo(dT)<sub>15</sub> substrate (0.2 U/ml) (Roche Diagnostics GmbH, Mannheim, Germany) in a reaction volume of 24  $\mu\text{l}$  and 50 mM Tris/HCl, pH 8.0, 80 mM KCl, 6 mM  $\text{MgCl}_2$ , 0.5 mM DTT with 150  $\mu\text{M}$  unlabeled TTP and 1.95  $\mu\text{Ci}$  of [ $^3\text{H}$ ]TTP (120 Ci/mmol; MP Biomedicals Inc., Irvine, CA, USA) [34]. Samples were preincubated for 5 min at 37°C. The reaction was started by the addition of 12 nM enzyme. After 5 min, 7.5  $\mu\text{l}$  aliquots were taken out and spotted on DEAE filter paper. Filters were washed twice with 2 x SSC (300 mM NaCl, 30 mM Na-citrate, pH 7.0) for 10 minutes, twice with 96 % ethanol for 10 min and dried. Incorporation of radiolabeled precursor into polydeoxynucleotide was determined by liquid scintillation counting. Under the conditions used 1 unit (U) of enzyme catalyzes the incorporation of 1 nmol TTP into poly(rA)/oligo(dT)<sub>15</sub> in 10 min at 37°C.

## Structure alignments

Structure alignments shown in Figure 5 were performed using the program WinCoot [35].

## Accession codes

The structure coordinates of the wt PFV RNase H (Q<sup>591</sup>-N<sup>751</sup>) domain were deposited in the Protein Data Bank under the accession code 2LSN. Chemical shift assignments were deposited in the BioMagResBank, accession number: 17745 [5].

## Abbreviations

*B. halodurans*: *Bacillus halodurans*; DTT: Dithiothreitol; *E. coli*: *Escherichia coli*; FV: Foamy virus; GB1: Immunoglobulin binding domain B1 of streptococcal protein G; HIV-1: Human immunodeficiency virus type 1; PFV: Prototype foamy virus; MoMLV: Moloney murine leukemia virus; PPT: Polypurine tract; PR: Protease; RHBD: RNA-hybrid binding domain; RNase H: Ribonuclease H; RT: Reverse transcriptase; SFVmac: Simian foamy virus from macaques; TEV: Tobacco etch virus; XMRV: Xenotropic murine leukemia virus-related virus.

## Competing interests

The authors declare that they have no competing interests.

## Authors' contributions

BMW conceived and coordinated the study. BL performed the experiments. BL, BMW, MJH and KS analyzed the data. BMW, BL and KS wrote the paper. PR provided conceptual input. All authors read and approved the manuscript.

## Acknowledgement

We thank the DFG (Wo630/7-3), the Bavarian International Graduate School of Science (BIGSS) and the University of Bayreuth for financial support. We thank Prof. Dr. Paul Rösch for continuous support and Ramona Heissmann for excellent technical assistance.

Received: 14 June 2012 Accepted: 10 August 2012

Published: 10 September 2012

## References

1. Moelling K, Bolognesi DP, Bauer H, Büsen W, Plassmann HW, Hausen P: Association of retroviral reverse transcriptase with an enzyme degrading the RNA moiety of RNA-DNA hybrids. *Nat New Biol* 1971, **234**:240-243.
2. Hartl MJ, Bodem J, Jochheim F, Rethwilm A, Rösch P, Wöhrl BM: Regulation of foamy virus protease activity by viral RNA - a novel and unique mechanism among retroviruses. *J Virol* 2011, **85**:4462-4469.
3. Hartl MJ, Mayr F, Rethwilm A, Wöhrl BM: Biophysical and enzymatic properties of the simian and prototype foamy virus reverse transcriptases. *Retrovirology* 2010, **7**:5.
4. Das D, Georgiadis MM: The crystal structure of the monomeric reverse transcriptase from moloney murine leukemia virus. *Structure* 2004, **12**:819-829.
5. Leo B, Hartl MJ, Schweimer K, Mayr F, Wöhrl BM: Insights into the structure and activity of the prototype foamy virus RNase H. *Retrovirology* 2012, **9**:14.
6. Tanese N, Goff SP: Domain structure of the moloney murine leukemia virus reverse transcriptase: Mutational analysis and separate expression of the DNA polymerase and RNase H activities. *Proc Natl Acad Sci U S A* 1988, **85**:1777-1781.
7. Schultz SJ, Champoux JJ: RNase H domain of moloney murine leukemia virus reverse transcriptase retains activity but requires the polymerase domain for specificity. *J Virol* 1996, **70**:8630-8638.
8. Zhan X, Crouch RJ: The isolated RNase H domain of murine leukemia virus reverse transcriptase. Retention of activity with concomitant loss of specificity. *J Biol Chem* 1997, **272**:22023-22029.
9. Kanaya S, Katsuda-Nakai C, Ikehara M: Importance of the positive charge cluster in *Escherichia coli* ribonuclease HI for the effective binding of the substrate. *J Biol Chem* 1991, **266**:11621-11627.
10. Telesnitsky A, Blain SW, Goff SP: Defects in moloney murine leukemia virus replication caused by a reverse transcriptase mutation modeled on the structure of *Escherichia coli* RNase H. *J Virol* 1992, **66**:615-622.
11. Lim D, Orlova M, Goff SP: Mutations of the RNase H C helix of the moloney murine leukemia virus reverse transcriptase reveal defects in polypurine tract recognition. *J Virol* 2002, **76**:8360-8373.
12. Lim D, Gregorio GG, Bingman C, Martinez-Hackert E, Hendrickson WA, Goff SP: Crystal structure of the moloney murine leukemia virus RNase H domain. *J Virol* 2006, **80**:8379-8389.
13. Zhou D, Chung S, Miller M, Grice SF, Wlodawer A: Crystal structures of the reverse transcriptase-associated ribonuclease H domain of xenotropic murine leukemia-virus related virus. *J Struct Biol* 2012, **177**:638-645.
14. Paprotka T, Delviks-Frankenberry KA, Cingoz O, Martinez A, Kung HJ, Tepper CG, Hu WS, Fivash MJ Jr, Coffin JM, Pathak VK: Recombinant origin of the retrovirus XMRV. *Science* 2011, **333**:97-101.
15. Nowotny M, Gaidamakov SA, Crouch RJ, Yang W: Crystal structures of RNase H bound to an RNA/DNA hybrid: Substrate specificity and metal-dependent catalysis. *Cell* 2005, **121**:1005-1016.
16. Evans DB, Brawn K, Deibel MR Jr, Tarpley WG, Sharma SK: A recombinant ribonuclease H domain of HIV-1 reverse transcriptase that is enzymatically active. *J Biol Chem* 1991, **266**:20583-20585.
17. Hostomsky Z, Hostomska Z, Hudson G, Moomaw EW, Nides BR: Reconstitution in vitro of RNase H activity by using purified N- terminal and C-terminal domains of human immunodeficiency virus type 1 reverse transcriptase. *Proc Natl Acad Sci U S A* 1991, **88**:1148-1152.
18. Smith JS, Roth MJ: Purification and characterization of an active human immunodeficiency virus type 1 RNase H domain. *J Virol* 1993, **67**:4037-4049.
19. Smith JS, Gritsman K, Roth MJ: Contributions of DNA polymerase subdomains to the RNase H activity of human immunodeficiency virus type 1 reverse transcriptase. *J Virol* 1994, **68**:5721-5729.
20. Davies JF 2nd, Hostomska Z, Hostomsky Z, Jordan SR, Matthews DA: Crystal structure of the ribonuclease H domain of HIV-1 reverse transcriptase. *Science* 1991, **252**:88-95.
21. Mueller GA, Pari K, DeRose EF, Kirby TW, London RE: Backbone dynamics of the RNase H domain of HIV-1 reverse transcriptase. *Biochemistry* 2004, **43**:9332-9342.
22. Kirby KA, Marchand B, Ong YT, Ndongwe TP, Hachiya A, Michailidis E, Leslie MD, Sietsema DV, Fetterly TL, Dorst CA, Singh K, Wang Z, Parniak MA, Sarafianos SG: Structural and inhibition studies of the RNase H function of xenotropic murine leukemia virus-related virus reverse transcriptase. *Antimicrob Agents Chemother* 2012, **56**:2048-2061.
23. Wöhrl BM, Volkmann S, Moelling K: Mutations of a conserved residue within HIV-1 ribonuclease H affect its exo- and endonuclease activities. *J Mol Biol* 1991, **220**:801-818.
24. Tisdale M, Schulze T, Larder BA, Moelling K: Mutations within the RNase H domain of human immunodeficiency virus type 1 reverse transcriptase abolish virus infectivity. *J Gen Virol* 1991, **72**:59-66.
25. Sarafianos SG, Das K, Tantillo C, Clark AD 2nd, Ding J, Whitcomb JM, Boyer PL, Hughes SH, Arnold E: Crystal structure of HIV-1 reverse transcriptase in complex with a polypurine tract RNA:DNA. *EMBO J* 2001, **20**:1449-1461.
26. Nowotny M, Gaidamakov SA, Ghirlando R, Cerritelli SM, Crouch RJ, Yang W: Structure of human RNase H1 complexed with an RNA/DNA hybrid: Insight into HIV reverse transcription. *Mol Cell* 2007, **28**:264-276.
27. Julius JG, McWilliams MJ, Sarafianos SG, Alvord WG, Arnold E, Hughes SH: Mutation of amino acids in the connection domain of human immunodeficiency virus type 1 reverse transcriptase that contact the template-primer affects RNase H activity. *J Virol* 2003, **77**:8548-8554.
28. Kay LE, Torchia DA, Bax A: Backbone dynamics of proteins as studied by <sup>15</sup>N inverse detected heteronuclear NMR spectroscopy: Application to staphylococcal nuclease. *Biochemistry* 1989, **28**:8972-8979.
29. Farrow NA, Muhandiram R, Singer AU, Pascal SM, Kay CM, Gish G, Shoelson SE, Pawson T, Forman-Kay JD, Kay LE: Backbone dynamics of a free and phosphopeptide-complexed src homology 2 domain studied by <sup>15</sup>N NMR relaxation. *Biochemistry* 1994, **33**:5984-6003.
30. Cornilescu G, Delaglio F, Bax A: Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J Biomol NMR* 1999, **13**:289-302.
31. Schwieters CD, Kuszewski JJ, Tjandra N, Clore GM: The Xplor-NIH NMR molecular structure determination package. *J Magn Reson* 2003, **160**:66-74.



32. Laskowski RA, Rullmann JAC, MacArthur MW, Kaptein R, Thornton JM: **AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR.** *J Biomol NMR* 1996, **8**:477–486.
33. Schrödinger L: *The PyMOL molecular graphics system, version 1.3.* Schrödinger, LLC: Mannheim, Germany; 2010.
34. Hartl MJ, Kretzschmar B, Frohn A, Nowrouzi A, Rethwilm A, Wöhrl BM: **AZT resistance of simian foamy virus reverse transcriptase is based on the excision of AZTMP in the presence of ATP.** *Nucleic Acids Res* 2008, **36**:1009–1016.
35. Emsley P, Cowtan K: **Coot: Model-building tools for molecular graphics.** *Acta Crystallogr D* 2004, **60**:2126–2132.

doi:10.1186/1742-4690-9-73

**Cite this article as:** Leo *et al.*: The solution structure of the prototype foamy virus RNase H domain indicates an important role of the basic loop in substrate binding. *Retrovirology* 2012 **9**:73.

**Submit your next manuscript to BioMed Central and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)





## 9 Einzelarbeit C

Anna Schneider, Daniel Peter, Jessica Schmitt, Berit Leo, Franziska Richter, Paul Rösch, Birgitta M. Wöhrl und Maximilian J. Hartl (2013): Structural requirements for enzymatic activities of foamy virus protease-reverse transcriptase. *Manuskript*.



# **Structural Requirements for Enzymatic Activities of Foamy Virus Protease-Reverse Transcriptase**

**Anna Schneider, Daniel Peter, Jessica Schmitt, Berit Leo, Franziska Richter, Paul Rösch, Birgitta M. Wöhrl\*, Maximilian J. Hartl\***

Universität Bayreuth, Lehrstuhl Biopolymere, Universitätsstr. 30, D-95447 Bayreuth,  
Germany

Corresponding authors, mailing address for Birgitta M. Wöhrl: Universität Bayreuth,  
Lehrstuhl Biopolymere, Universitätsstr. 30, D-95447 Bayreuth, Germany;

Phone: +49 921 55-3542; Fax: +49 921 55-3544; E-mail: birgitta.woehrl@uni-bayreuth.de

mailing address for Maximilian J. Hartl: Universität Bayreuth, Lehrstuhl Biopolymere,  
Universitätsstr. 30, D-95447 Bayreuth, Germany;

Phone: +49 921 55-3869; Fax: +49 921 55-3544; E-mail: maximilian.hartl@uni-bayreuth.de

**Key words: foamy virus, reverse transcriptase, retroviral protease, substrate binding, domain structure, enzyme activity...**

## Abstract

**Background:** Reverse transcriptases (RTs) of retroviruses are required for virus replication and harbor a polymerase and an RNase H domain. The polymerase uses the viral RNA template to reverse transcribe it into double stranded DNA, while the RNase H degrades the RNA strand during this process. Structural analyses and sequence alignments revealed that the polymerase domain is sub-divided into fingers, palm, thumb and the connection subdomain which links the polymerase to the C-terminal RNase H domain. The RT of foamy viruses is distinct from that of other retroviruses, since, in addition to the polymerase and RNase H domains, it harbors the protease (PR) domain at its N-terminus. Thus, the mature enzyme is the PR-RT. Due to these differences and to the fact that homology of foamy virus RTs to orthoretroviral RTs is low, it is difficult to define the precise boundaries and functions of the (sub-)domains.

**Results:** To understand better the structural implications on the catalytic activity of the PR-RT enzyme we tested the properties of truncated PR-RTs. Both, the RNase H domain and the connection subdomain contribute substantially to polymerase integrity and stability as well as to polymerase activity and substrate binding. The 42 amino acid region C-terminal of the PR is important for polymerase stability and activity. *In vitro*, the cleavage efficiencies of FV PR for the Gag and Pol cleavage site are comparable, even though in virus particles only the Pol site is cleaved to completion. Furthermore, we found that PR activation *via* binding of PR-RT to the viral RNA requires the presence of the full length PR-RT including the RNase H domain.

**Conclusions:** The data presented here identify precise (sub-)domain boundaries for FV PR-RT and a role for the 42 amino acid region, following the PR sequence, for polymerase function. We show that the RNase H domain is required for PR activation. Although *in vivo* Gag cleavage by PR is less efficient, we found similar cleavage efficiencies for Gag and Pol *in vitro*, although. This implies that additional factors are involved in controlling PR activity and that virus maturation needs to be strictly regulated.

## Introduction

Foamy viruses (FVs) are retroviruses but constitute a separate subfamily, the so-called *spumaretrovirinae*. In contrast to members of the only other subfamily *orthoretrovirinae*, they synthesize separate Gag and Pol proteins instead of a Gag-Pol fusion protein [1]. Pol harbors the viral enzymes protease (PR), reverse transcriptase (RT) and integrase (IN). As opposed to other retroviruses, FVs process Pol into mature PR-RT and IN proteins, i.e. PR is not cleaved off from the RT [2,3]. Nevertheless, FVs have to strictly control Pol cleavage during the viral life cycle to avoid untimely processing before virus assembly, since this would lead to non-infectious virus particles. The second FV PR cleavage site is located close to the C-terminus of the Gag p71 polyprotein separating the FV capsid protein p68 from p3, a 27 amino acid (aa) extension of unknown function [4]. While Pol processing in infectious virus particles is complete, Gag cleavage remains partial. In addition, FV p3 deletion mutants are able to replicate, while FVs with an inactive Gag p68/p3 cleavage site are not, demonstrating the necessity for a tight control of PR activity during FV assembly and maturation [5-8].

We have shown previously, that the isolated PR domain from simian FV from macaques (SFVmac) is a monomer of ca. 10 kDa which has to dimerize in order to be active [3,9,10]. In the mature PR-RT, the PR is located N-terminally of the RT. PR-RT is also monomeric and dimerization of PR in the PR-RT context is achieved *via* specific binding of at least two PR-RT molecules to a region in the *pol* gene of the viral RNA, the so-called *protease activating RNA motif* (PARM), thus enabling the PR domains to form active homodimers [3,11,12].

So far, crystal structures of the RTs from human immunodeficiency virus type 1 and type 2 (HIV-1, HIV-2) [13-15] and of Moloney murine leukemia virus (MoMLV) [16] are available. The functional forms of HIV-1 and HIV-2 RTs are heterodimers. The larger subunit harbors an N-terminal polymerase domain and a C-terminal RNase H domain. The sequence of the small subunit is identical to the N-terminal region of the large subunit, but lacks the RNase H domain [17-20]. In contrast, MoMLV RT is a monomeric enzyme [16] and thus resembles FV RTs [11]. However, since the overall similarity of FV RTs to MoMLV RT is below 25% (data not shown) and FV PR is not cleaved off from the RT, precise subdomain assignments are difficult. Nevertheless, it is widely assumed that the organization of the FV PR-RT is comparable to the PR and RT of other retroviruses. RTs consist of a C-terminal RNase H and an N-terminal polymerase domain, which is divided into the subdomains fingers, thumb, palm and connection. The latter links the polymerase domain to the N-terminus of the RNase H.

Determination of the three-dimensional structures of SFVmac PR and PFV RNase H confirmed that, despite their low sequence homologies, high structural similarities in these

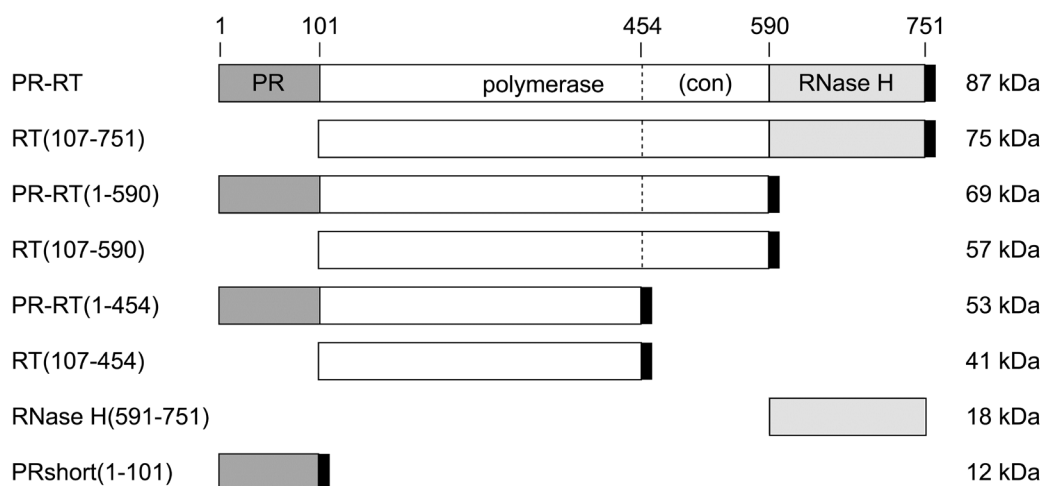
regions are conserved throughout the family of retroviruses [9,21]. On the other hand, the two structures also demonstrated that the assumed FV PR-RT domain borders predicted by sequence alignments were incorrect [2]. Inaccurate determination of domain and subdomain borders can cause serious problems in the design and interpretation of *in vivo* as well as *in vitro* experiments. Moreover, precise domain definitions are required for protein structure determination. Therefore, in the absence of detailed structural data of the complete FV PR-RT, knowledge about its exact domain and subdomain organization is urgently needed. An important question regarding the domain organization of FV PR-RT is the function of the region aa 102-143, located between the C-terminus of the PR domain as defined by Hartl *et al.* [9] and the N-terminus of the RT according to Pfrepper *et al.* [2]. Interestingly, monomeric MoMLV RT also carries an N-terminal extension of 39 aa, which neither exhibits homology to the HIV-1 RT sequence nor to the region aa 102-143 of FV PR-RT. To date it has not been clarified whether in FVs this peptide stretch is necessary for proper FV RT function, i.e. whether it is part of the RT domain and important for accurate positioning of the PR to achieve PR dimerization in the PR-RT bound to PARM, or whether it solely represents a flexible linker connecting the two domains. Structural and functional data on SFVmac PR demonstrated that it is not an essential part of the PR domain [9]. Furthermore, based on mere sequence alignments with HIV-1 RT, identification of the precise junction between the polymerase and RNase H domains from prototype FV (PFV) and SFVmac has proved difficult [2]. The RNase H of FV is not cleaved off from PR-RT, whereas in the heterodimeric HIV-1 RT p66/p51 the RNase H domain is removed in the p51 subunit, thus defining the RNase H domain junction [21-24].

To learn more about the structural features and enzymatic requirements of FV PR-RTs, we used the SFVmac PR-RT and designed C- and N-terminal PR-RT deletion mutants, which we subjected to enzymatic and structural investigations. Here, we identify exact domain and subdomain boundaries of SFVmac PR-RT as well as the intrinsic requirements for each of its enzymatic activities and for the structural integrity of the domains. Finally, we demonstrate that SFVmac PR does not discriminate between the Gag and Pol cleavage sites *in vitro*, although in the virus Pol is completely cleaved into PR-RT and IN, whereas the Gag p68/p3 cleavage is incomplete. Consequently, we hypothesize that in the viral context, complete cleavage is prevented by steric hindrance and/or additional factors. Such a mechanism could be required to control the correct sequence of events essential for virus maturation.



## Results

*Protein variants.* To analyze the domain structure and the intrinsic requirements of SFVmac PR-RT enzyme activities we constructed several C- and N-terminally deleted proteins and analyzed their properties and activities. Figure 1 gives an overview of the protein variants tested. Previous studies have already shown that the first 101 aa are sufficient for PR activity in the presence of 3 M NaCl (Figure 1, PRshort) [9]. However, from those data it was not clear, whether the region following the PR domain is required for polymerase and/or RNase H activities. We tested two different starting points for the polymerase: either after the PR domain at aa 107, or at aa 144 as suggested by Pfrepper et al. [2]. Enzymes starting at aa 144 lacked polymerization activity (data not shown), indicating a putative role of the region aa 107-143 for integrity and functionality of the polymerase.

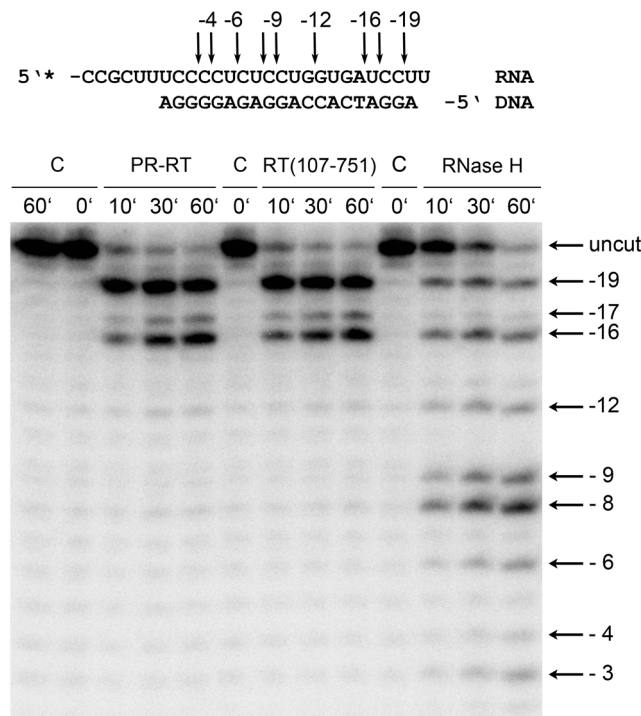


**Figure 1: Schematic overview of the SFVmac PR-RT constructs and domain structure.** SFVmac PR-RT comprises PR, polymerase and RNase H domains. The connection (con) subdomain of the polymerase is shown as well. Boundaries between domains are represented as continuous and boundaries between subdomains as dashed lines. The exact amino acid number of the C-terminal end of each domain is depicted on top of the figure. The denominations of the used protein constructs are given on the left and the corresponding molecular masses on the right. The C-terminal 6 x His tag is depicted as a black rectangle.

Similarly, in order to define the minimal length of the C-terminus required to sustain polymerase activity and PR activation via PARM-RNA, several C-terminal deletion mutants were constructed. These mutants either lacked only the RNase H or the RNase H plus the connection subdomain. Based on sequence alignments with MoMLV or HIV-1 RT, aa 449 and aa 454, respectively, were tested as starting points for the connection domain. However, PR-RT(1-449) and RT(107-449) were completely insoluble after expression in *E. coli*. In contrast, PR-RT(1-454) and RT(107-454) could be obtained as soluble proteins.

*RNase H activities.* In a first set of experiments we determined whether the presence of the PR domain at the N-terminus has an effect on RNase H activities. Therefore, wild type (wt) PR-

RT, RT(107-751) lacking the PR domain and the isolated RNase H domain were analyzed in a qualitative RNase H assay. Although the separate RNase H domain from PFV still exhibits RNase H cleavage activity, it is severely reduced as compared to the wt PR-RT [21]. To achieve cleavage with the independent SFVmac RNase H domain, we used 20  $\mu$ M of RNase H as compared to 30 nM for the wt and for RT(107-751) (Figure 2).



**Figure 2: Qualitative RNase H activities of different SFVmac PR-RT deletion variants.** The 20/27 mer DNA/RNA primer/template substrate, 5' end-labeled at the RNA strand with  $^{32}$ P, is presented on top of the figure. Each enzyme (30 nM wt PR-RT, 30 nM RT(107-751) and 20  $\mu$ M RNase H) was incubated with 250 nM substrate at 25 °C for the times shown on top of the gel. Cleavage products were separated on a 15% sequencing gel and visualized by phosphoimaging. Arrows and numbers indicate the cleavage sites identified on the gel. The first nucleotide of the RNA hybridized to the 3'-OH nucleotide of the DNA strand is designated -1. C: control reactions were performed in the absence of enzyme for 0 min and 60 min.

wt and RT(107-751) exhibited comparable RNase H activities with identical cleavage sites (specified in the sequence on top of the gel). In contrast, the isolated RNase H domain cleaved the substrate indiscriminately in hybrid region. This is due to non-specific binding of one or more RNase H molecules, whereas binding and cleavage of the wt and RT(107-751) are directed by the 3'OH end of the DNA-primer. Binding of the RT polymerase active site to the 3'OH orients the RNase H at a distance of ca. 17-19 nucleotides (N) away from the 3'OH end of the primer and endonucleolytic cleavage can take place in the RNA strand [24,25].

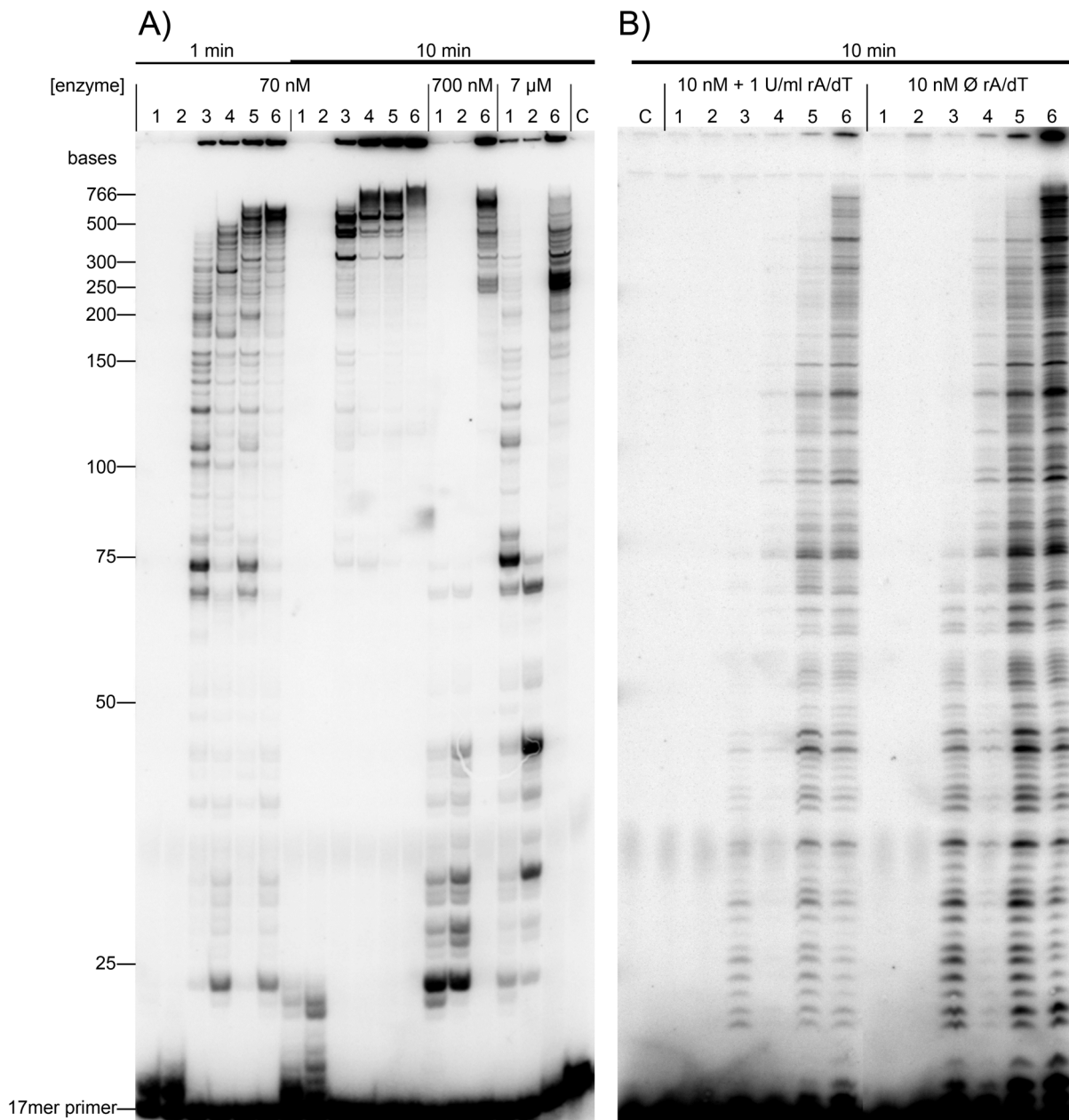
Quantitative RNase H activity analysis with a fluorescently labeled substrate further indicated that the RNase H activities of the wt and the N-terminal deletion mutant RT(107-751) were comparable, whereas that of the isolated RNase H domain was roughly three orders of magnitude lower (Table 1), owing to a reduced affinity for nucleic acid substrates [21].

**Table 1: Kinetic parameters of polymerization and RNase H activities.**

| Enzyme           | RNase H activity<br>( $\frac{\text{pmol}}{10 \text{ min} \times \text{pmol enzyme}}$ ) | Polymerization<br>activity<br>( $\frac{\text{U}}{\text{nmol enzyme}}$ ) | Dissociation constant<br>$K_d$ (nM) |
|------------------|--|---|-------------------------------------|
| PR-RT            | 6.1  | 505   | 6 ( $\pm 1$ )                       |
| RT(107-751)      | 4.7  | 244   | 3 ( $\pm 1$ )                       |
| PR-RT(1-590)     | -  | 206   | 54 ( $\pm 7$ )                      |
| RT(107-590)      | -  | 147   | 37 ( $\pm 4$ )                      |
| PR-RT(1-454)     | -  | $243 \times 10^{-3}$  | 200 ( $\pm 6$ )                     |
| RT(107-454)      | -  | $224 \times 10^{-3}$  | 213 ( $\pm 13$ )                    |
| RNase H(591-751) | $2.1 \times 10^{-3}$   | -   | -                                   |

*Binding affinities.* Since changes in substrate affinities can have an impact on RNase H and polymerase enzyme activities we investigated the binding properties of the proteins with a fluorescently labeled DNA/DNA substrate. We performed enzyme titrations and used the change in fluorescence anisotropy to determine  $K_D$ -values (Table 1, Figure S1). Both, wt and RT(107-751) possessed similar substrate binding affinities with  $K_D$ -values of 6 nM and 3 nM, respectively, indicating that the PR domain is not involved in binding of nucleic acid substrates, neither directly nor indirectly (Table 1, Figure S1A). However, deletion of the RNase H domain, either in the presence or absence of the PR domain, resulted in a ca. 90 % reduction of affinity with PR-RT(1-590) or RT(107-590) (Table 1, Figure S1B). Further deletion of the connection subdomain (= PR-RT(1-454) and RT(107-454)) led to an additional decrease of affinity (Table 1, Figure S1C). Thus, the RNase H domain and the connection subdomain - but not the PR domain - of SFVmac PR-RT are important for high affinity substrate binding.

*Polymerization and processivity.* To analyze the polymerization properties of the enzymes we performed qualitative polymerization assays on a single stranded heteropolymeric M13 substrate. RT(107-454) and PR-RT(1-454) exhibited very little polymerization activity and even after 10 min only very short elongation products ( $< 8$  N) were visible (Fig. 3A, lanes 2 and 3). After one min most elongations stopped at +1. A tenfold increase of the enzyme concentration yielded detectable amounts of extension products of up to 75 N, another tenfold increase of enzyme resulted in even longer products. All other enzymes tested were much more active and synthesized extension products of 300 – 700 N, already at an enzyme concentration of 70 nM.



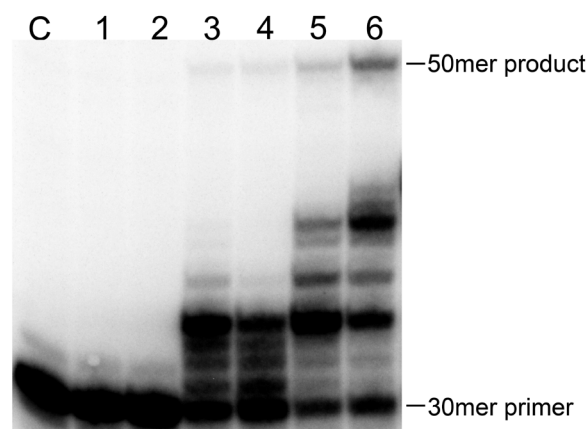
**Figure 3: Polymerization properties.** (A) Time dependent primer extension using varying enzyme concentrations. DNA dependent polymerization reactions with 10 nM single stranded M13 DNA were performed at 37 °C. Reactions were started by the addition of enzyme (lanes 1: RT(107-454), lanes 2: PR-RT(1-454), lanes 3: RT(107-590), lanes 4: RT(107-751), lanes 5: PR-RT(1-590), lanes 6: wt PR-RT) and stopped at the time points depicted on top. DNA size markers are shown on the left. C: control reactions without enzyme. (B) Processivity assay. Primer elongations with 10 nM of enzyme are shown in presence (rA/dT) and absence (Ø rA/dT) of the competitor poly(rA)/oligo(dT)<sub>15</sub>, respectively. Labeling of lanes as in (A).

However, none of the PR-RT deletion mutants exhibited polymerization activity comparable to the wt enzyme (Figure 3A, lanes 6), suggesting that deletions either at the C- or N-terminus disturb the structural integrity of the proteins. This was also noticeable with RT(107-751) only lacking the PR domain (Figure 3A, lanes 4), although this domain is not involved in nucleic acid binding. Most enzymes produced strong stop bands below the 25 N marker, i.e.

within the first 8 N of extension, suggesting that they have to overcome an initiation phase in order to polymerize processively.

The processivity of a DNA polymerase is defined as the length of polymerization products synthesized during one round of binding and polymerization without any dissociation/re-association step. To further examine processivity, we performed polymerization assays at very low enzyme concentrations (10 nM) on the M13 substrate in the presence of excess competitor substrate to prevent re-association (Figure 3B, left panel). At these enzyme concentrations, polymerization products were not visible in the presence of competitor with RT(107-454) and PR-RT(1-454) (lanes 1 and 2). Both, RT(107-590) and RT(107-751) (Figure 3B, lanes 3 and 4) were less processive than PR-RT(1-590) and the wt PR-RT (Figure 3B, lanes 5 and 6), so that again, deletion of the PR domain appears to lead to some minor structural changes in the RT that influence its stability and/or polymerization behavior. More importantly, in the absence of competitor, even for the wt protein, a major band was visible after the extension of only +1 N or +2 N (Figure 3B, right panel). Quantification of the wt PR-RT polymerization products revealed that around 50 % are +1 N and +2 N extension products, supporting our notion that the enzymes have to overcome a polymerization initiation state before transition into the processive phase takes place.

*Initiation of polymerization.* To study the initiation of polymerization in more detail, we used a 30/50 mer DNA/DNA primer/template substrate, which allowed a more precise separation of the short extension products created during the initiation phase (Figure 4). As before, RT(107-454) and PR-RT(1-454), both lacking the connection domain, produced no extension products under the conditions tested in this assay (Figure 4, lanes 1 and 2).



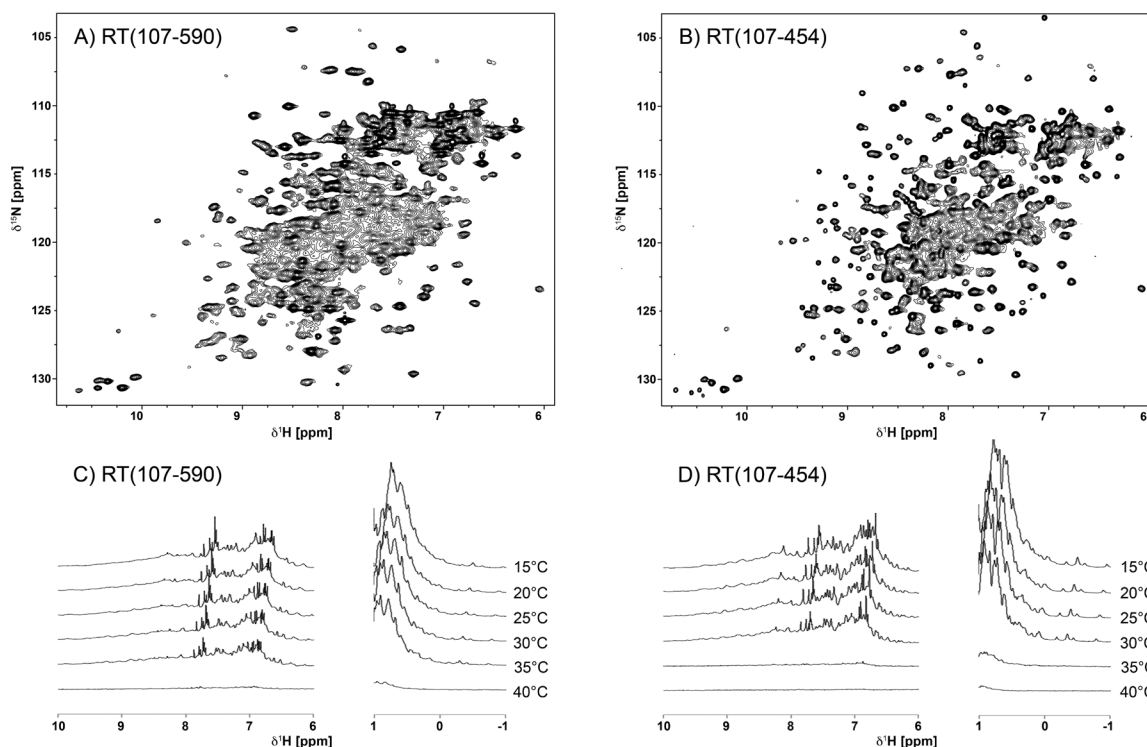
**Figure 4: Efficiency of the initiation of polymerization.** Reactions were carried out at 37 °C for 15 min with 20 nM primer/template and 10 nM RT on a 30/50 mer DNA/DNA primer/template. Lane 1: RT(107-454), lane 2: PR-RT(1-454), lane 3: RT(107-590), lane 4: RT(107-751), lane 5: PR-RT(1-590), lane 6: wt PR-RT, C: control reaction without enzyme.

Although all other enzymes exhibited polymerization activity, mainly short extension products of +1 to +4 N were synthesized, with a particularly strong band at +4 N. The wt protein managed to overcome this strong pause site and to extend a considerable proportion of substrate to the end of the template. Remarkably, there were no further stops after the extension of 10 N (Figure 4, lane 6), in good agreement with the results for the wt enzyme (Figure 3A, lanes 6) [11]. This suggests that, after incorporation of a few nucleotides, the enzyme needs to undergo a conformational change to switch to the elongation phase, which finally results in processive polymerization.

*Structural integrity and thermal stability studied by NMR spectroscopy.* By analyzing the enzymatic activities and binding affinities of the deletion mutants we were able to determine the domain and subdomain borders. We could demonstrate that the two smallest constructs RT(107-454) and RT(107-590) were still able to polymerize, although with reduced activities (Table 1, Figure 3). To get some structure data on these enzymes, we recorded [ $^{15}\text{N}$ ,  $^1\text{H}$ ] HSCQ NMR spectra of the  $^{15}\text{N}$  labeled RTs (Figure 5A, B). The spectra of the two enzymes revealed well dispersed signals between 6.5 and 9.5 ppm, indicative of folded proteins with a defined tertiary structure. However, due to the size of the proteins the signals overlap and signal intensities are inhomogeneous.

To determine the stability of the two enzymes, we performed 1D- $^1\text{H}$  NMR experiments at different temperatures ranging from 15 °C up to 40 °C (Figure 5C, D). As expected, RT(107-590) still harboring the connection subdomain was more stable than RT(107-454) (Figure 5C, D). A strong signal decrease could be observed with RT(107-454) when increasing the temperature from 30 °C to 35 °C, whereas with RT(107-590) a similar decrease was only visible at 40 °C. This signal loss is probably due to protein unfolding and irreversible aggregation. Figure 5 also suggests that the size and instability of the proteins at elevated temperatures impedes structure determination by NMR.

These results are in good agreement with the data presented in Table 1. In summary, they demonstrate that the connection subdomain plays an important role in high affinity substrate binding and polymerization and, in addition, contributes significantly to the stability of FV RTs.



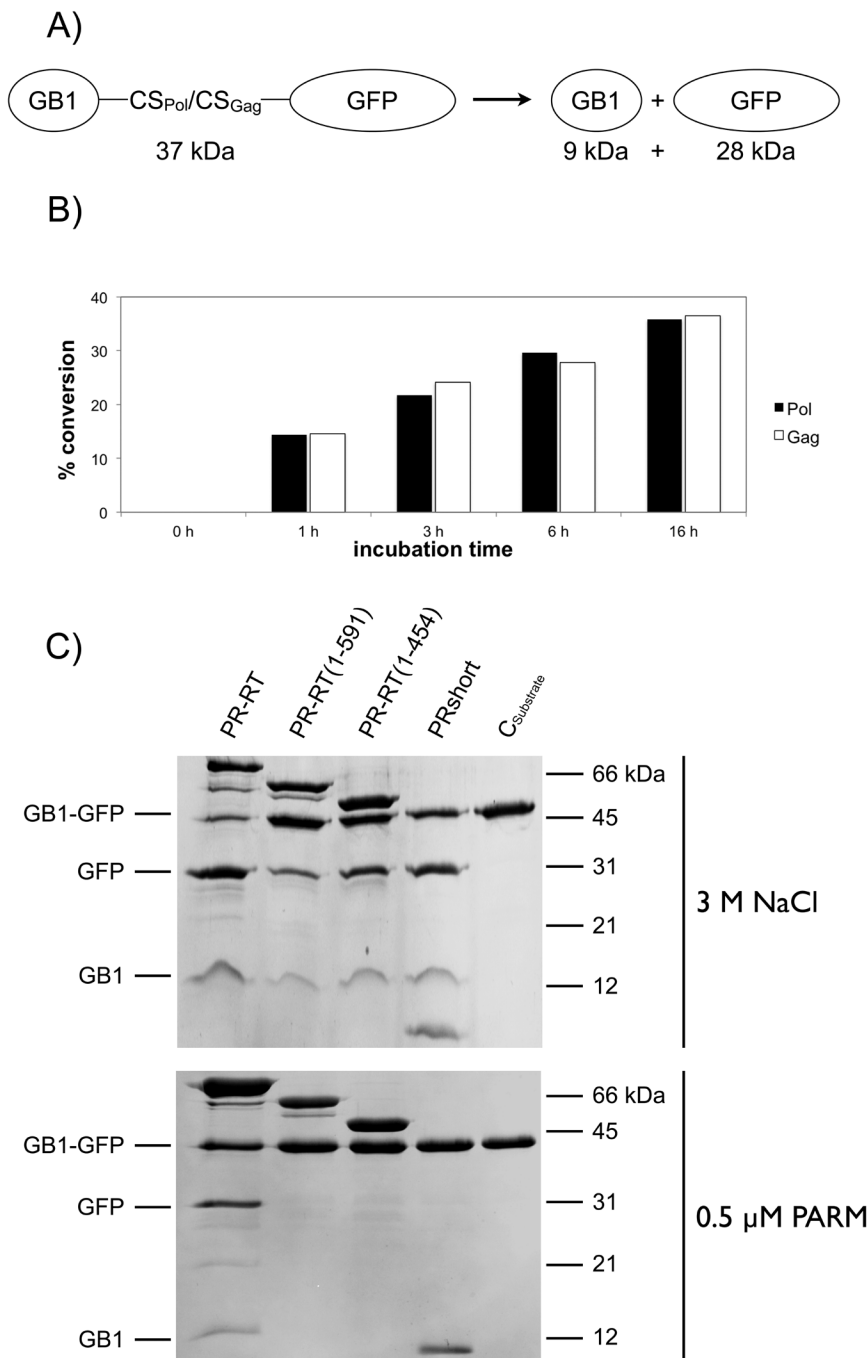
**Figure 5: Tertiary structure analysis and thermal stability of SFVmac RT(107-590) and RT(107-454).** [ $^{15}\text{N}$ ,  $^1\text{H}$ ] HSQC spectra of 120  $\mu\text{M}$  RT(107-590) (A) and 100  $\mu\text{M}$  RT(107-454) (B) at 25  $^\circ\text{C}$  in 50 mM  $\text{K}_2\text{HPO}_4/\text{KH}_2\text{PO}_4$  pH 7.4, 100 mM KCl, 10%  $\text{D}_2\text{O}$  and in 50 mM HEPES pH 7.4, 300 mM NaCl and 10%  $\text{D}_2\text{O}$ . Temperature dependent  $^1\text{H}$  spectra of 50  $\mu\text{M}$  RT(107-590) (C) and RT(107-454) (D) in 50 mM  $\text{K}_2\text{HPO}_4/\text{KH}_2\text{PO}_4$  pH 7.4 100 mM KCl, 10%  $\text{D}_2\text{O}$ . All measurements were performed with a Bruker DRX 800 MHz spectrometer.

*Protease activity.* Regulation of PR activity in FVs is distinct from that of other known retroviruses [11]. PR activation is achieved by binding of at least two PR-RT molecules to the PARM on the viral RNA, which results in dimerization of two PR domains. Only then cleavage can occur. *In vitro*, dimerization in the absence of PARM can also be attained using buffers containing 2 to 3 M NaCl. The high ionic strength of the buffer appears to promote hydrophobic interactions of the PR domains.

*In vivo*, two PR cleavage sites have been determined: one in Pol, leading to PR-RT and IN proteins, the other one at the C-terminus of the Gag p71 precursor giving rise to the products p68 and p3. However, analysis of the protein composition of virus particles demonstrated that in FVs Gag cleavage is only partial, whereas Pol is processed completely into PR-RT and IN [5-8].

Therefore, we wanted to determine, whether *in vitro* one of the two cleavage sites is preferred and cleaved with higher efficiency by wt PR-RT. Two substrates were used, containing either the natural Gag or Pol cleavage site between the proteins GB1 and GFP (Figure 6A). A time course experiment with the two substrates in the presence of PARM RNA showed no

difference in cleavage efficiencies for the two substrates suggesting that other regulatory factors interfere with Gag cleavage (Figure 6B).



**Figure 6: Protease activity assays. (A)** Scheme of the polymerization assays. The substrates containing either the SFVmac Pol or Gag cleavage site (CS) between GB1 and GFP are cleaved into the products GB1 and GFP by an active SFVmac PR. **(B)** Time dependent decrease of substrate GB1-CS<sub>Pol</sub>-GFP (black) or GB1-CS<sub>Gag</sub>-GFP (gray) catalyzed by SFVmac PR-RT in the presence of 0.5  $\mu$ M PARM and 100 mM NaCl. **(C)** Activity assay with different SFVmac PR-RTs and the GB1-CS<sub>Pol</sub>-GFP substrate. 10  $\mu$ M of GB1-CS<sub>Pol</sub>-GFP were incubated with 5  $\mu$ M of the enzyme depicted on top in the presence of 3 M NaCl (upper panel) or 0.5  $\mu$ M PARM and 100 mM NaCl (lower panel). Reaction products were separated by 19% SDS-PAGE. C<sub>Substrate</sub>: uncleaved substrate. The sizes of the standard proteins are shown on the right.



Finally, we wanted to identify regions of the PR-RT important for PR activation by PARM. Since no differences in substrate preference were detected, we used the Pol cleavage site containing construct for analysis. The PR activities of the C-terminal deletion mutants PR-RT(1-590), lacking the RNase H domain, and PR-RT(1-454), additionally lacking the connection subdomain were tested in the presence of 3 M NaCl or in the presence of PARM (Figure 6C) and compared to the wt enzyme and PRshort. Analysis at 3 M NaCl showed the PR domain of the deletion mutants to be functional and comparable to the wt PR-RT or the isolated PR domain (Figure 6C, upper panel). However, PR activity of the two deletion mutants as well as PRshort was completely abolished in the presence of PARM and low NaCl concentrations (Figure 6C, lower panel). These results show that the structural integrity of the complete PR-RT enzyme is essential for proper high affinity binding to PARM in order to initiate PR cleavage.

## Discussion

FVs are the most ancient retroviruses known [26]. Their unique replication strategy and particle organization sets them apart from all other retroviruses [27-29]. However, compared to other retroviruses like HIV-1, detailed and thorough molecular and structural data on FV proteins are scarce and, owing to low sequence homologies, transfer of, e.g. structural data from well-characterized systems to FVs is challenging and often unreliable. Strikingly, in FVs the PR is not cleaved off from the RT resulting in a mature PR-RT enzyme [2,3], a protein unknown in other retroviruses.

Several attempts have been made to identify and characterize the domain and subdomain organization of FV PR-RT, but were hampered because the predicted domain junctions were not precise [2]. Moreover, PR-RT variants were insoluble, and the influence of deletions on the enzyme's stability and integrity restricted the use of deletions variants in *in vivo* experiments [30]. Similar experience was gained with an independent FV PR domain with an extended C-terminus, PR(1-143), which resulted in PR oligomerization and thus artificial PR activation [2,9].

Here we studied proteolytic cleavage, DNA synthesis and RNA hydrolysis to identify domain and subdomain junctions aiming at the construction of PR-RT variants for *in vivo* and structural biology work (Figure 1, Figure S2).

The structure of the RNase H domain of PFV [21,24] was used here to define the C-terminus of the connection subdomain. The independent RNase H domain (aa 591-751) exhibits significantly lower RNase H activity than the wt or the PR deletion mutant RT(107-751) (Figure 2, Table 1). Deletion of the RNase H domain resulted in a tenfold higher  $K_d$ -value (Table 1, Figure S1) for primer/template binding and in a decrease in enzymatic activity and processivity of PR-RT(1-590) and RT(107-590) (Figure 3).

Moreover, we were able to define the beginning of the connection domain at aa E455 since C-terminal deletion mutants ending at aa L449 proved insoluble. The consequence of removing the connection subdomain in addition to the RNase H domain was a further decrease in primer/template binding (ca. 100-fold, Table 1, Figure S1) and a dramatic decline in polymerization capability. Elimination of the connection subdomain had a serious impact on the structural integrity of the polymerase domain (Figure 5), suggesting that the proper protein fold was destabilized. In particular, the initiation process of polymerization and the transition to elongation appeared severely impaired (Figures 3 and 4). Our studies imply that the transition from initiation to elongation is a critical step during DNA synthesis, even for the wt enzyme (Figures 3 and 4), and that – comparable to HIV-1 RT - the connection subdomain

fulfills a more complex role than just linking the polymerase to the RNase H domain. Nevertheless, RT(107-454) exhibits a well-defined three-dimensional structure (Figure 5B) and residual enzymatic activity (Table 1, Figure 3). Thus, it can be defined as the smallest active FV polymerase and can be employed in future investigations, i.e. to study structural features of azidothymidine (AZT) resistant SFVmac PR-RTs by NMR or X-ray crystallography, because all the resistance mutations are located in this residual polymerase domain [31,32].

The N-terminal deletions demonstrate that the region aa 107-143, located adjacent to the C-terminus of the PR, is not just a linker connecting the PR and RT but an intrinsic part of the RT domain, necessary for RT solubility and integrity and possibly for proper PR positioning. Deletion of the PR domain (aa 1-106) does not substantially influence substrate binding or polymerization activities (Table 1, Figure 3). In the structure of the MoMLV RT, which also harbors an N-terminal extension, aa 23-40 of the extension are suggested to be part of the palm subdomain and form an  $\alpha$ -helix. aa 1-23 are missing in the recombinant enzyme used for structure determination, but deletion of these residues did not impair RT polymerization activities [16,33,34]. However, the N-terminal regions of FV RT (aa 102-143) and MoMLV RT do not exhibit homologies and are absent in HIV-1 RT. From our data, we propose that in FV PR-RT the region the aa 107-143 folds back onto the RT domain and thus locks the PR into a certain orientation which facilitates PR dimerization upon binding of the RT to the PARM RNA.

Our results on PR activity show, that the entire RT including the RNase H domain is needed for proper binding to the PARM RNA and thus PR activation. Upon deletion of the RNase H domain (PR-RT(1-590) and PR-RT(1-454)), PR activation via PARM binding was completely abrogated although the PR domain was intact, as demonstrated by activation of PR *via* 3 M NaCl (Figure 6C). PR inactivation by RNase H deletion is probably due to a substantial decrease in the affinity for nucleic acids of those mutants (Table 1).

Surprisingly, although in viruses PR cleavage of Gag but not Pol, is incomplete [5-8], our experiments using the natural Gag and Pol cleavage sites reveal the same cleavage efficiency for the two substrates *in vitro*. Here, we solely measured the affinity of the PR for the cleavage sites, since the adjacent amino acids and domains are identical in our substrates (Figure 6A). Obviously, *in vivo* during virus maturation, the Gag p68/p3 cleavage site is less accessible than the Pol RT/IN site, thus leading to delayed and therefore only partial Gag cleavage. Bound pre-genomic viral RNA or the Gag p71 folding topology might be the cause for steric hindrance. We postulate, that this slow-down is a regulatory feature to control the

precise order of events during virus assembly and maturation since the virus has to guarantee that IN is cleaved off from Pol before reverse transcription starts in order to be able to create active IN tetramers [35]. These are required in the next round of infection to integrate the double stranded DNA into the host genome. Premature start of the reverse transcription process will result in destruction of the PARM secondary structure required for PR activity and thus non-infectious virus. However, at least some of the Gag molecules have to be cleaved to obtain infectious virus particles [5,36].

**Conclusions:** Here, we present precise domain borders of FV PR-RT as well as the intrinsic requirements for its enzymatic activities. We identified a minimal polymerase domain RT(107-454), which includes the region following the C-terminus of the PR.

We demonstrate an essential role for region aa 107-143 for polymerase but not protease activity, whereas the connection subdomain is important for substrate affinity, protein stability and proper function of the enzyme. Furthermore, our data on PR activation show that the full length PR-RT including the RNase H domain is required for PR activity and that the affinities of the PR domain for the Gag and Pol cleavage sites are comparable. Based on this finding we conclude that steric hindrance due to Gag p71 folding or cellular or viral factors leads to a delayed Gag cleavage, important for regulating the order of events during virus maturation. This model implies that the correct sequence of events during viral replication is of utmost importance for the virus to guarantee the production of infectious viral particles.

## Methods

*Cloning, expression and protein purification.* SFVmac PR-RT [32] and PRshort [9] were prepared according to published purification protocols. SFVmac RNase H cloning, gene expression and protein purification was performed as published previously for PFV RNase H [24].

The genes coding for the other PR-RT deletion variants were cloned into the vector pET101/D-TOPO (Invitrogen, Darmstadt, Germany). The resulting proteins contained a 6 x His tag at the C-terminus directly extending the protein sequences. Protein synthesis of the PR-RT deletion variants was performed in *Escherichia coli* (*E. coli*) Rosetta DE3 pLysSRARE (Novagen, Darmstadt, Germany) at 37 °C for 3 h (RT(107-751)), at 37 °C for 4 h (PR-RT(1-590)), at 25 °C for 4 h (PR-RT(1-454) and RT(107-454)) or 25 °C for 16 h (RT(107-590)). Expressions were induced by the addition of 1 mM isopropylthiogalactoside (IPTG) at an optical density at 600 nm of 0.8–1.0. Purification via HisTrap and heparin affinity chromatography was performed as published previously [32]. <sup>15</sup>N labeled proteins were purified accordingly, however, the bacterial strain was grown in M9 minimal medium [37] supplemented with trace solution TS2 [38] and with (<sup>15</sup>NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> as the only N-source.

*RNase H activity assays.* The qualitative RNase H assay was performed as described previously with minor changes [21,24]. In a buffer containing 50 mM Tris/HCl, 80 mM KCl, 6 mM MgCl<sub>2</sub>, 0.5 mM DTT and 0.5 mM EDTA, pH 7.0, 250 nM radioactively labeled DNA/RNA hybrid was incubated with 30 nM SFVmac PR-RT, 30 nM SFVmac RT(107-751) or 20 μM SFVmac RNase H at 25 °C for the time periods indicated in the figure. Cleavage products were separated by denaturing gel electrophoresis (15 % polyacrylamide, 8 M urea) and quantified by densitometry using a phosphoimaging device (CR 35 Bio, raytest, Straubenhardt, Germany) and the software program AIDA (Image Analyzer v. 4.50).

Quantitative activity measurements were carried out with a RNA strand fluorescently labeled at the 3' end (5'-CCGAUCGCUCUCCUGGUGAUCCUUUCC-6-FAM) and hybridized to a DNA primer, which harbored a dabcyyl quencher at its 5' end (5'-Dabcyyl-GGAAAGGATCACCAGGAGAG [24]. 40 nM substrate was digested with 1nM SFVmac PR-RT, 1nM SFVmac RT(107-751) or 1,15 μM SFVmac RNase H in 50 mM Tris/HCl, 80 mM KCl, 6 mM MgCl<sub>2</sub>, and 0.5 mM DTT, pH 8.0. The linear slope of the reaction progress curves were used to calculate the activity of the enzymes in pmol of cleaved substrate per 10 min per pmol enzyme.

*Fluorescence equilibrium titrations.* Dissociation constants ( $K_d$ ) for DNA binding of the RTs were determined by fluorescence equilibrium titrations with a 24/40-mer DNA/DNA

primer/template substrate (primer 5'-ATCACCAGGAGAGGGGAAAGCGGA, template 5'-DY647-CTAATTCGCTTTCCCCTCTCCTGGTGATCCTTTCCATCC;) (biomers.net GmbH, Ulm, Germany) at 25 °C as described previously [32].  $K_d$ -values were calculated by non-linear curve fitting to the anisotropy data with a two component binding equation [24].

*Polymerization activity assays.* Quantification of the RNA-dependent DNA polymerase activity was performed using a poly(rA)/oligo(dT)<sub>15</sub> substrate (Roche Diagnostics GmbH, Mannheim, Germany) in a standard assay [32]. For qualitative analysis of the DNA-dependent DNA polymerase activity, a heteropolymeric single stranded M13mp18 DNA (New England BioLabs, Frankfurt, Germany) was used as substrate. Reactions were performed as described previously [32]. Incubation times and enzyme concentrations were varied as indicated in the figures. Elongation products were separated by denaturing gel electrophoresis (10 % polyacrylamide, 8 M urea) and quantified as described above for the RNase H assays.

*Processivity assay.* 11 nM of a radioactively labeled 17-mer DNA primer (5'-P<sup>32</sup>-GTAAAACGACGGCCATG) (biomers.net; Ulm, Germany) [32] annealed to M13mp18 DNA (New England BioLabs, Frankfurt, Germany) were pre-incubated with 10 nM of enzyme in 50 mM Tris/HCl, 80 mM KCl, 6 mM MgCl<sub>2</sub>, 0.5 mM DTT, pH 7.7 for 2 min at 37 °C. Reactions in a total volume of 10 µl were started by the addition of 250 µM dNTPs in presence or absence of 1 U/ml poly(rA)/oligo(dT)<sub>15</sub> as a competitor for the DNA substrate. The reaction was stopped after 10 min at 37 °C by mixing the sample with 10 µl of stopping buffer (8 M urea, 0.1% xylene cyanol, 0.1% bromophenol blue in 1x TBE). Polymerization products were separated by denaturing gel electrophoresis using 10 % polyacrylamide/8 M urea gels and quantified as described above.

*Initiation of polymerization.* In a total volume of 10 µl, 20 nM radioactively labeled 30-mer DNA primer hybridized to a 50-mer DNA template (primer 5'-<sup>32</sup>P-GCTGTAATGGCGTCCCTGTTCGGGCGCCTC, template 5'-GCTGTGGAAAATCTCATGCAGAGGGCGCCCGAACAGGGACGCCATTACAGC) (IBA, Göttingen, Germany) [32] were incubated with 250 µM dNTPs for 2 min at 37 °C in 50 mM Tris/HCl, 80 mM KCl, 6 mM MgCl<sub>2</sub> and 0.5 mM DTT, pH 7.7. The polymerization reaction was initiated by the addition of enzyme to a final concentration of 10 nM. Subsequently, samples were incubated for 15 min at 37 °C and reactions stopped by adding 10 µl of stopping buffer (8 M urea, 0.1 % xylene cyanol, 0.1 % bromophenol blue in 1x TBE). Elongation products were analyzed by denaturing gel electrophoresis using 10 % polyacrylamide/8 M urea gels and quantified as described above.

*NMR analysis.* [ $^{15}\text{N}$ ,  $^1\text{H}$ ] HSQC spectra were recorded for  $^{15}\text{N}$ -labeled RT(107-590) and RT(107-454) at a sample temperature of 25 °C. The samples contained 120  $\mu\text{M}$  RT(107-590) in 50 mM  $\text{K}_2\text{HPO}_4/\text{KH}_2\text{PO}_4$  pH 7.4, 100 mM KCl and 10 %  $\text{D}_2\text{O}$ , or 100  $\mu\text{M}$  RT(107-454) in 50 mM HEPES pH 7.4, 300 mM NaCl and 10 %  $\text{D}_2\text{O}$ . Temperature dependent 1D NMR spectra were measured with 50  $\mu\text{M}$  RT(107-590) or RT(107-454) in 50 mM  $\text{K}_2\text{HPO}_4/\text{KH}_2\text{PO}_4$  pH 7.4, 100 mM KCl and 10 %  $\text{D}_2\text{O}$ . All NMR experiments were recorded on a Bruker Avance 800 MHz spectrometer equipped with a cryo-probe. In-house protocols were used to process the NMR data.

*PR activity assay.* PR activity assays were performed as described previously [9,11], using a GB1-GFP cleavage substrate that harbors either the SFVmac Pol (ATQGSYVVH↓CNTTP) or the Gag (GENTDRSVN↓TVTAT) cleavage site between the immunoglobulin binding domain B1 of the streptococcal protein G (GB1) and the green fluorescent protein (GFP). In brief, 10  $\mu\text{M}$  GB1-GFP substrate was incubated with 5  $\mu\text{M}$  enzyme in 50 mM  $\text{Na}_2\text{HPO}_4/\text{NaH}_2\text{PO}_4$ , 0.5 mM DTT, pH 6.4, 3 M NaCl or alternatively 100 mM NaCl in combination with 0.5  $\mu\text{M}$  PARM RNA for 16 h at 37 °C. Cleavage products were analyzed by 19% SDS-polyacrylamide gel electrophoresis (PAGE) and quantified densitometrically [10].

## **Abbreviations**

FV, foamy virus; HEPES, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; HIV-1, human immunodeficiency virus type 1; HIV-2 human immunodeficiency virus type 2; IN, integrase; MoMLV, Moloney murine leukemia virus; N, nucleotide; NMR, nuclear magnetic resonance; PARM, protease activating RNA motif; PR, protease; RNase H, ribonuclease H; RT, reverse transcriptase; SFVmac, simian foamy virus from macaques;

## **Competing Interests**

The authors declare they have no competing interests.

## **Authors' Contributions**

BMW and MJH conceived and coordinated the study. AS, DP, JS, BL and FR and MJH performed the experiments. MJH, BMW and AS analyzed the data. BMW and MJH wrote the paper. PR provided conceptual input. All authors read and approved the manuscript.

## **Acknowledgement**

We thank the DFG (Wo630/7-3), the Bavarian International Graduate School of Science (BIGSS) and the University of Bayreuth for financial support, and Ramona Heissmann for excellent technical assistance.



## References

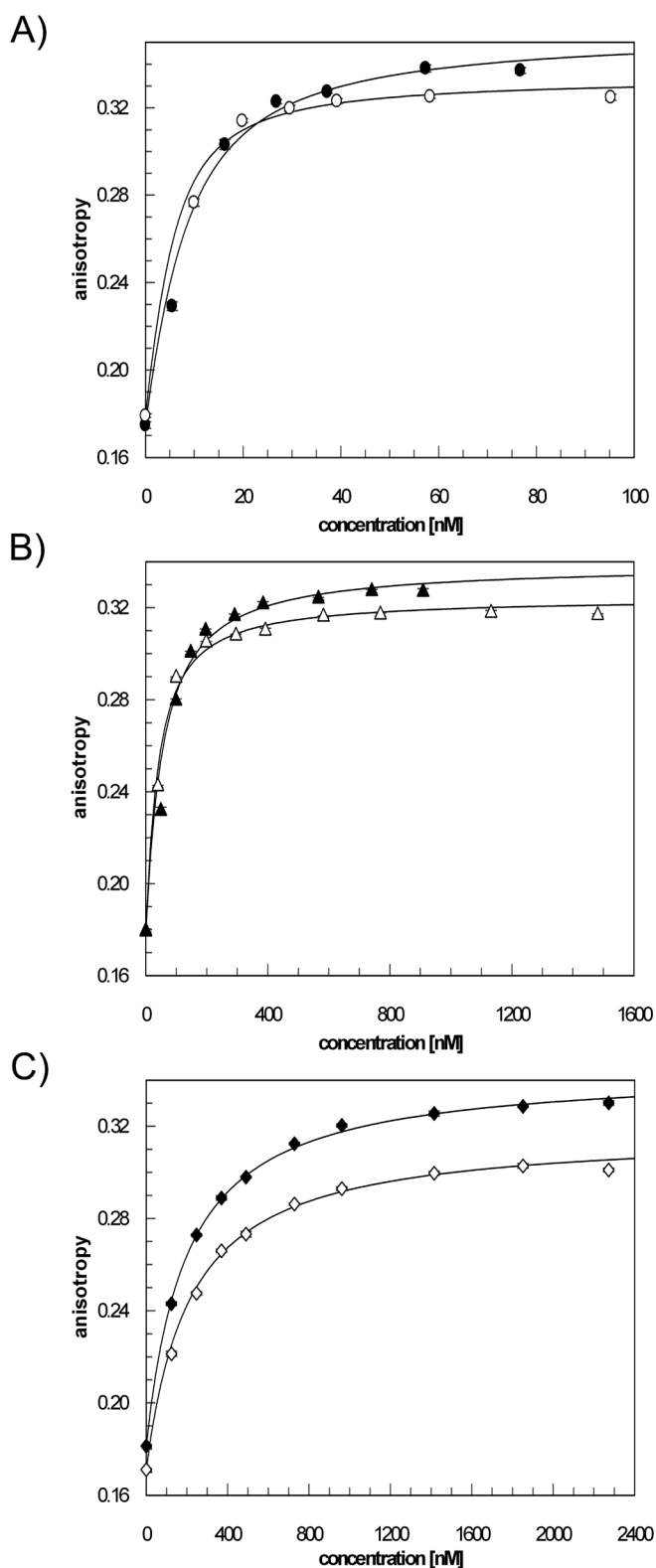
1. Löchelt M, Flügel RM: **The human foamy virus pol gene is expressed as a pro-pol polyprotein and not as a gag-pol fusion protein.** *J Virol* 1996, **70**:1033-1040.
2. Pfrepper KI, Rackwitz HR, Schnölzer M, Heid H, Löchelt M, Flügel RM: **Molecular characterization of proteolytic processing of the pol proteins of human foamy virus reveals novel features of the viral protease.** *J Virol* 1998, **72**:7648-7652.
3. Hartl MJ, Bodem J, Jochheim F, Rethwilm A, Rösch P, Wöhrl BM: **Regulation of foamy virus protease activity by viral RNA - a novel and unique mechanism among retroviruses.** *J Virol* 2011, **85**:4462-4469.
4. Pfrepper KI, Löchelt M, Rackwitz HR, Schnölzer M, Heid H, Flügel RM: **Molecular characterization of proteolytic processing of the gag proteins of human spumavirus.** *J Virol* 1999, **73**:7907-7911.
5. Enssle J, Fischer N, Moebes A, Mauer B, Smola U, Rethwilm A: **Carboxy-terminal cleavage of the human foamy virus gag precursor molecule is an essential step in the viral life cycle.** *J Virol* 1997, **71**:7312-7317.
6. Zemba M, Wilk T, Rutten T, Wagner A, Flügel RM, Löchelt M: **The carboxy-terminal p3Gag domain of the human foamy virus gag precursor is required for efficient virus infectivity.** *Virology* 1998, **247**:7-13.
7. Baldwin DN, Linial ML: **Proteolytic activity, the carboxy terminus of gag, and the primer binding site are not required for pol incorporation into foamy virus particles.** *J Virol* 1999, **73**:6387-6393.
8. Swiersy A, Wiek C, Reh J, Zentgraf H, Lindemann D: **Orthoretroviral-like prototype foamy virus gag-pol expression is compatible with viral replication.** *Retrovirology* 2011, **8**:66.
9. Hartl MJ, Wöhrl BM, Rösch P, Schweimer K: **The solution structure of the simian foamy virus protease reveals a monomeric protein.** *J Mol Biol* 2008, **381**:141-149.
10. Hartl MJ, Schweimer K, Reger MH, Schwarzinger S, Bodem J, Rösch P, Wöhrl BM: **Formation of transient dimers by a retroviral protease.** *Biochem J* 2010, **427**:197-203.

11. Hartl MJ, Mayr F, Rethwilm A, Wöhrl BM: **Biophysical and enzymatic properties of the simian and prototype foamy virus reverse transcriptases.** *Retrovirology* 2010, **7**:5.
12. Spannaus R, Hartl MJ, Wöhrl BM, Rethwilm A, Bodem J: **The prototype foamy virus protease is active independently of the integrase domain.** *Retrovirology* 2012, **9**:41.
13. Jacobo-Molina A, Ding J, Nanni RG, Clark AD, Jr., Lu X, Tantillo C, Williams RL, Kamer G, Ferris AL, Clark P, Hizi A, Hughes SH, Arnold E: **Crystal structure of human immunodeficiency virus type 1 reverse transcriptase complexed with double-stranded DNA at 3.0 Å resolution shows bent DNA.** *Proc Natl Acad Sci U S A* 1993, **90**:6320-6324.
14. Kohlstaedt LA, Wang J, Friedman JM, Rice PA, Steitz TA: **Crystal structure at 3.5 Å resolution of HIV-1 reverse transcriptase complexed with an inhibitor.** *Science* 1992, **256**:1783-1790.
15. Ren J, Bird LE, Chamberlain PP, Stewart-Jones GB, Stuart DI, Stammers DK: **Structure of HIV-2 reverse transcriptase at 2.35-Å resolution and the mechanism of resistance to non-nucleoside inhibitors.** *Proc Natl Acad Sci U S A* 2002, **99**:14410-14415.
16. Das D, Georgiadis MM: **The crystal structure of the monomeric reverse transcriptase from Moloney murine leukemia virus.** *Structure* 2004, **12**:819-829.
17. Müller B, Restle T, Kühnel H, Goody RS: **Expression of the heterodimeric form of human immunodeficiency virus type 2 reverse transcriptase in *Escherichia coli* and characterization of the enzyme.** *J Biol Chem* 1991, **266**:14709-14713.
18. Müller B, Restle T, Weiss S, Gautel M, Sczakiel G, Goody RS: **Co-expression of the subunits of the heterodimer of HIV-1 reverse transcriptase in *Escherichia coli*.** *J Biol Chem* 1989, **264**:13975-13978.
19. Le Grice SFJ, Grüninger-Leitch F: **Rapid purification of homodimer and heterodimer HIV-1 reverse transcriptase by metal chelate affinity chromatography.** *Eur J Biochem* 1990, **187**:307-314.
20. Hansen J, Schulze T, Mellert W, Moelling K: **Identification and characterization of HIV-specific RNase H by monoclonal antibody.** *EMBO J* 1988, **7**:239-243.

- 
21. Leo B, Schweimer K, Rösch P, Hartl MJ, Wöhrl BM: **The solution structure of the prototype foamy virus RNase H domain indicates an important role of the basic loop in substrate binding.** *Retrovirology* 2012, **9**:73.
22. di Marzo Veronese F, Copeland TD, DeVico AL, Rahman R, Oroszlan S, Gallo RC, Sarngadharan MG: **Characterization of highly immunogenic p66/p51 as the reverse transcriptase of HTLV-III/LAV.** *Science* 1986, **231**:1289-1291.
23. Hansen J, Schulze T, Mellert W, Moelling K: **Identification and characterization of HIV-specific RNase H by monoclonal antibody.** *EMBO J* 1988, **7**:239-243.
24. Leo B, Hartl MJ, Schweimer K, Mayr F, Wöhrl BM: **Insights into the structure and activity of prototype foamy virus RNase H.** *Retrovirology* 2012, **9**:14.
25. Kirby KA, Marchand B, Ong YT, Ndongwe TP, Hachiya A, Michailidis E, Leslie MD, Sietsema DV, Fetterly TL, Dorst CA, Singh K, Wang Z, Parniak MA, Sarafianos SG: **Structural and inhibition studies of the RNase H function of xenotropic murine leukemia virus-related virus reverse transcriptase.** *Antimicrob Agents Chemother* 2012, **56**:2048-2061.
26. Switzer WM, Salemi M, Shanmugam V, Gao F, Cong ME, Kuiken C, Bhullar V, Beer BE, Vallet D, Gautier-Hion A, Tooze Z, Villinger F, Holmes EC, Heneine W: **Ancient co-speciation of simian foamy viruses and primates.** *Nature* 2005, **434**:376-380.
27. Linial ML, **Foamy viruses are unconventional retroviruses.** *J Virol* 1999, **73**:1747-1755.
28. Yu SF, Baldwin DN, Gwynn SR, Yendapalli S, Linial ML: **Human foamy virus replication: A pathway distinct from that of retroviruses and hepadnaviruses.** *Science* 1996, **271**:1579-1582.
29. Rethwilm A, **Molecular biology of foamy viruses.** *Med Microbiol Immunol* 2010, **199**:197-207.
30. Kögel D, Aboud M, Flügel RM: **Molecular biological characterization of the human foamy virus reverse transcriptase and ribonuclease H domains.** *Virology* 1995, **213**:97-108.
31. Kretzschmar B, Nowrouzi A, Hartl MJ, Gartner K, Wiktorowicz T, Herchenröder O, Kanzler S, Rudolph W, Mergia A, Wöhrl B, Rethwilm A: **AZT-resistant foamy virus.** *Virology* 2008, **370**:151-157.

32. Hartl MJ, Kretzschmar B, Frohn A, Nowrouzi A, Rethwilm A, Wöhrl BM: **AZT resistance of simian foamy virus reverse transcriptase is based on the excision of AZTMP in the presence of ATP.** *Nucleic Acids Res* 2008, **36**:1009-1016.
33. Najmudin S, Cote ML, Sun D, Yohannan S, Montano SP, Gu J, Georgiadis MM: **Crystal structures of an N-terminal fragment from moloney murine leukemia virus reverse transcriptase complexed with nucleic acid: Functional implications for template-primer binding to the fingers domain.** *J Mol Biol* 2000, **296**:613-632.
34. Coté ML, Roth MJ: **Murine leukemia virus reverse transcriptase: Structural comparison with HIV-1 reverse transcriptase.** *Virus Res* 2008, **134**:186-202.
35. Roy J, Linial ML: **Role of the foamy virus pol cleavage site in viral replication.** *J Virol* 2007, **81**:4956-4962.
36. Konvalinka J, Löchelt M, Zentgraf H, Flügel RM, Kräusslich H-: **Active foamy virus proteinase is essential for virus infectivity but not for formation of a pol polyprotein.** *J Virol* 1995, **69**:7264-7268.
37. Sambrook, J., Fritsch, E.,F and Maniatis, T. (1994) *Molecular cloning - A laboratory manual.* , Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
38. Meyer O, Schlegel HG: **Biology of aerobic carbon monoxide-oxidizing bacteria.** *Annu Rev Microbiol* 1983, **37**:277-310.

## Supplementary Information



**Figure S1: Determination of  $K_d$ -values by fluorescence anisotropy measurements.** 15 nM of a fluorescently labeled DNA/DNA substrate was titrated with different SFVmac PR-RT constructs at 25 °C. The curves represent the best fit based upon a two-component binding model.  $K_d$ -values with standard deviations are given in Table 1. (A) SFVmac PR-RT (●) and RT(107-751) (○), (B) SFVmac PR-RT(1-590) (▲) and RT(107-590) (△), (C) SFVmac PR-RT(1-454) (◆) and RT(107-454) (◇).

```

1 MDPLQ LLQPL EAEIK GTKLK AHWDS GATIT CVPEA FLEDE RPIQT MLIKT
51 IHGEK QQDVY YLTFK VQGRK VEAEV LASPY DYILL NPSDV PWSMK KPLQL
101 TVLVP LHEYQ ERLLO QTALP KEQKE LLQKL FLKYD ALWQH WENQV GHRRI
151 KPHNI ATGTL APRPQ KQYPI NPKAK PSIQI VIDDL LKQGV LIQQN STMNT
201 PVYVP PKPDG KWRMV LDYRE VNKII PLIAA QNQHS AGILS SIYRG KYKTT
251 LDLTN GFWAH PITPE SYWLT AFTWQ GKQYC WTRLP QGFLN SPALF TADV
301 DLLKE IPNVQ AYVDD IYISH DDPQE HLEQL EKIFS ILLNA GYVVS LKKSE
351 IAQRE VEFLG FNITK EGRGL TDTEK QKLLN ITPPK DLKQL QSILG LLNFA
401 RNFIP NYSEL VKPLY TIVAN ANGKF ISWTE DNSNQ LQHII SVLNQ ADNLE
451 ERNPE TRLLI KVNSS PSAGY IRYYN EGSKR PIMYV NYIFS KAEAK FTQTE
501 KLLTT MHKGL IKAMD LAMGQ EILVY SPIVS MTKIQ RTPLP ERKAL PVRWI
551 TWMTY LEDPR IQFHY DKSLP ELQQI PNVTE DVIK TKHPS EFAMV FYTDG
601 SAIKH PDVVK SHSAG MGIAQ VQFIP EYKIV HQWSI PLGDH TAQLA EIAAV
651 EFACK KALKI SGPVL IVTDS FYVAE SANKE LPYWK SNGFL NNKKK PLRHV
701 SKWKS IAACL QKPD IIMH EKGHQ QPMTT LHTEG NNLAD KLATQ GSYVV
751 HLEHH HHHH

```

**Figure S2: Domain organization of SFVmac PR-RT.** The (sub-)domains are color-coded: blue: PR; dark green: polymerase domain, the region 107-143 is shown in italics; bright green: connection subdomain; red: RNase H; Numbers on the left indicate amino acid positions. Amino acids labeled in black are either not assigned (aa 102 – 106) or belong to the C-terminal tag (aa 752 – 759)

## 10 Danksagung

An dieser Stelle möchte ich Prof. Dr. Birgitta M. Wöhrl für dieses interessante Projekt und die damit verbundenen zahlreichen Diskussionen und Aufmunterungen danken. Ich danke ihr auch für die sehr gute Betreuung, die Geduld, den Einsatz und die Unterstützung während der ganzen Arbeit, menschlich sowie wissenschaftlich.

Prof. Dr. Paul Rösch danke ich für die Möglichkeit, dass ich die Arbeit am Lehrstuhl Biopolymere unter hervorragenden Arbeitsbedingungen durchführen konnte. Der BIGSS-Graduiertenschule und der Stabsabteilung Chancengleichheit danke ich für alle Förderungsmaßnahmen, auch in finanzieller Hinsicht. Der DFG danke ich für die finanzielle Unterstützung.

Dr. Kristian Schweimer danke ich für die Unterstützung und die Erklärungen in allen Fragenstellungen bezüglich NMR.

Dr. Maximilian Hartl, Anna Schneider, Dr. Hanna Berkner und Dr. Stefan Knauer danke ich vielmals für ihre Diskussionsbereitschaft und ihre kritischen Betrachtungsweisen sowie für die zahlreiche Hilfe und die Beantwortung meiner Fragen. Vor allem aber danke ich für die aufmunternden Worte, die immer dann kamen, wenn sie nötig waren.

Ich danke dem TA-Team - Andrea Hager, Ramona Heißmann, Britta Zimmermann und Ulrike Persau - für alle Unterstützung und für unerschöpfliche Geduld im und außerhalb des Labors. Außerdem bin ich Rainer Hofmann für die schnelle Hilfe bei Computerfragen und -problemen sehr dankbar. Den aktuellen und ehemaligen Mitarbeiterinnen des Sekretariats LS Biopolymere möchte ich für die Hilfe in Verwaltungsangelegenheiten danken.

Ich bedanke mich auch bei Christian Seutter von Loetzen, Johanna Drögemüller, Martin Strauß, Philipp Weiglmeier, Sabine Wenzel, Sophie Pleißner und Dr. Olivia Hartl-Spiegelhauer für die Unterstützung in wissenschaftlichen sowie nicht-wissenschaftlichen Fragestellungen und für die angenehme Arbeitsatmosphäre. Johanna Koch danke ich für ein stets offenes Ohr und ihre fachkompetente Unterstützung während des Studiums und der Promotion.

Mein Dank gilt auch allen Sportkameraden, vor allem aber Dorothea Wischnat, Monika Luthner, Margitta Grötsch und Marina Behr, die mich jederzeit angespornt haben. Ich danke meiner Familie für ihre Unterstützung.

Holger danke ich für alles.





## **11 Erklärung**

Hiermit erkläre ich, dass ich die Arbeit selbständig verfasst und keine anderen als die von mir angegebenen Quellen und Hilfsmittel benutzt habe.

Ferner erkläre ich, dass ich anderweitig mit oder ohne Erfolg nicht versucht habe, diese Dissertation einzureichen. Ich habe keine gleichartige Doktorprüfung an einer anderen Hochschule endgültig nicht bestanden.

Bayreuth,

