The human cytidine deaminase APOBEC3C restricts retroviruses independent of editing

–

A biochemical and structural analysis

vom Fachbereich Biologie der Technischen Universität Darmstadt

zur

Erlangung des akademischen Grades

eines Doctor rerum naturalium

genehmigte Dissertation von

Dipl. Biol. Henning Hofmann

aus Erbach/Odenwald

1. Referent: **Prof. Dr. Gerhard Thiel**

2. Referent: **Prof. Dr. Carsten Münk**

3. Referent: **PD Dr. Arnulf Kletzin**

Tag der Einreichung: 18. August 2010

Tag der mündlichen Prüfung: 3. Dezember 2010

Darmstadt 2011

D 17

Eidesstattliche Erklärung

Ich erkläre hiermit an Eides statt, dass ich die vorliegende Dissertation selbstständig und nur mit den angegebenen Hilfsmitteln angefertigt habe. Ich habe bisher noch keinen Promotionsversuch unternommen.

_____

Henning Hofmann

Darmstadt, den 18. August 2010

Im Zuge der Promotion konnten folgende Arbeiten veröffentlicht werden:

Vif of Feline Immunodeficiency Virus from domestic cats protects against APOBEC3 restriction factors from many felids
J. Zielonka, D. Marino, H. Hofmann, N. Yuhki, M. Löchelt, and C. Münk
J. Virol., 2010. 84: 7312-7324

Model structure of APOBEC3C reveals a binding pocket modulating ribonucleic acid interaction required for encapsidation
B. Stauch*, H. Hofmann*, M. Perkovic, M. Weisel, F. Kopietz, K. Cichutek, C. Münk, and G. Schneider;
Proc. Natl. Acad. Sci., Jul 21;106 (29):12079-84. 2009.
*B.S. and H.H. contributed equally to this work

High level expression of the anti-retroviral protein APOBEC3G is induced by influenza A virus but does not confer antiviral activity
E-K. Pauli, M. Schmolke, H. Hofmann, C. Erhardt, E. Flory, C. Münk and S. Ludwig; *Retrovirology, 2009, 6:38*

Species-specific inhibition of APOBEC3C by the prototype foamy virus protein bet
M. Perkovic, S. Schmidt, D. Marino, R. A. Russell, B. Stauch, H. Hofmann, F. Kopietz, B.-P. Kloke, J. Zielonka, H. Ströver, J. Hermle, D. Lindemann, V. K. Pathak, G. Schneider, M. Löchelt, K. Cichutek, and C. Münk; *JBC, 2009. 284(9):5819-26*

Teile dieser Arbeiten wurden in Chapter I der vorliegenden Arbeit verwendet.

Abgesehen von folgenden Abbildungen in Chapter I wurden alle Daten von mir generiert:

Fig. 1, Fig. 2, Fig. 3 (a+b), Fig. 7, Fig. 9 und Fig. S1 bis S5

Contents

Zusammenfassung 1

Summary 2

Introduction 3

Chapter I 15

Chapter II 30

Chapter III 41

Appendix 51

Zusammenfassung

Die Fähigkeit des *Humanen Immundefizienz Virus* (HIV) in Zellen verschiedener Spezies zu replizieren ist häufig abhängig von der An- oder Abwesenheit interagierender oder einschränkender zellulärer Faktoren. Restriktionsfaktoren stellen einen Teil des Abwehrmechanismus des Wirts gegen Pathogene dar. Die Familie der humanen APOBEC3 (A3) Zytidin-Deaminasen ist Teil der intrinsischen Immunität und schützt höhere Säugetiere vor viralen Pathogenen. A3 Proteine tragen eine oder zwei Kopien eines Zytidin-Deaminase Motivs und können die retrovirale Replikation inhibieren, indem sie deren Genom während der reversen Transkription deaminieren. Diese Inhibition wird von HIV durch das Vif-Protein überwunden, welches den Abbau von A3s induziert und damit deren Verpackung in Virionen verhindert.

Obwohl ein Mitglied dieser humanen Proteinfamilie, APOBEC3C (A3C), hoch in vielen menschlichem Geweben und in $CD4^+$ T-Zellen exprimiert wird, kann A3C nur *vif*-defizientes SIV (SIVΔ*vif*), nicht aber HIVΔ*vif* inhibieren. Ziel der vorliegenden Arbeit war es, 1.) die unterschiedlichen Resistenzmechanismen von HIV und SIV gegenüber A3C zu definieren, 2.) den Restriktionsmechanismus von A3C gegenüber SIV zu charakterisieren, und 3.) basierend auf einem Strukturmodel Aminosäurereste zu identifizieren, die relevant für die antivirale Aktivität von A3C sind. Mit der Hilfe eines dreidimensionalen Proteinmodels wurden Aminosäurereste prognostiziert und experimentell identifiziert, die in die Dimerisierung des Proteins oder dessen Verpackung in virale Partikel involviert sind. Die Ergebnisse zeigen, dass die Dimerisierung von A3C notwendig für dessen antivirale Aktivität gegen SIV ist. Neben der Dimerisierung wird die antivirale Aktivität durch Verpackung in virale Partikel reguliert. Die Verpackung von A3C beruht auf einer Substratbindetasche, die distal des Zink-koordinierenden enzymatisch aktiven Zentrums liegt. Diese Bindetasche vermittelt die RNA-abhängige Verpackung des Proteins in das knospende Virus. Es konnte gezeigt werden, dass der Mechanismus der A3C-vermittelten Restriktion von SIVΔ*vif* nicht die Deaminierung virale ssDNA ist. Die experimentelle Untersuchung alternativer Mechanismen zeigte, dass A3C weder virale RNA deaminiert, noch einen signifikanten inhibitorischen Effekt auf die reverse Transkription von SIV hat. Daher ist anzunehmen, dass A3C einen zusätzlichen, nicht bekannten Restriktionsmechanismus gegen SIV besitzt.

Ergebnisse dieser Arbeit zeigen weiterhin, dass die Resistenz von HIV gegen A3C überwunden werden kann. Eine A3C vermittelte Restriktion von HIV wurde durch eine N- oder C-terminale Fusion des viralen Proteins R (Vpr) von HIV-1 an A3C erreicht. Interessanter Weise wurden beide Proteine, A3C und Vpr.3C, ins Virus verpackt und waren im gleichen viralen Kompartiment, dem viralen *Core*, zu finden. Diese Daten weisen darauf hin, dass ein von HIV zusätzlich zu *vif* kodierter Faktor der A3C Aktivität entgegenwirkt. Auf der Suche nach einem derartigen Faktor, konnte die virale Integrase als ein Kandidat identifiziert werden.

Weitere Experimente sind notwendig, um die Interaktion von A3C mit HIV besser zu verstehen. Jedoch implizieren die Ergebnisse dieser Arbeiten, dass eine Steigerung der antiviralen Aktivität des ubiquitär exprimierten A3C die Replikation von HIV in Patienten eindämmen könnte und so einen neuen Behandlungsansatz ermöglicht.

Summary

The capability of the *Human immunodeficiency virus* (HIV) to replicate in cells of different species depends on the presence or absence of interacting or restricting cellular factors. The restriction factors are part of the host's defense mechanism against pathogens. The family of human APOBEC3 (A3) cytidine deaminases forms part of the intrinsic immunity and protects placental mammals from viral pathogens. A3 proteins have one or two copies of a cytidine deaminase motif and can restrict retroviral replication by deamination of their genomes during reverse transcription. HIV counteracts this inhibition by the Vif-protein, which induces the destruction of A3s and thereby prevents its incorporation into virions.

Although one member of this human protein family, APOBEC3C (A3C), is highly expressed in many human tissues and in $CD4^+$ T cells, A3C does not restrict *vif*-deficient HIV (HIV Δ*vif*) but is active against SIV Δ*vif*. This study aimed to characterize i) the different mechanisms of resistance of HIV and SIV against A3C, ii) the restriction mechanism of A3C against SIV, and iii) residues crucial for the antiviral activity of A3C based on a structural model. With the help of a three-dimensional protein model of A3C, amino acids residues were predicted and finally experimentally identified that are involved in protein dimerization and that mediate the incorporation of A3C into virions. The result showed that dimerization of A3C is essential for antiviral activity against SIV. Beside dimerization, the antiviral activity is regulated by encapsidation into viral particles. Encapsidation of A3C depends on the substrate binding-pocket distal from the zinc-coordinating enzymatic centre. This binding pocket mediates the RNA-dependent incorporation of the protein into budding viral particles. It was found that A3C restricts SIV Δ*vif* without cytidine deamination of viral ssDNA. Therefore, alternative mechanisms were investigated. In this study the experiments showed clearly that A3C does neither deaminates viral RNA nor has a significant inhibitory effect on the reverse transcription of SIV. Thus, it is likely that A3C has a so far unknown restriction mechanism against SIV Δ*vif*.

Results of this study also show that the resistance of HIV against A3C can be overcome. An A3C mediated inhibition of HIV was achieved by fusing the Viral Protein R (Vpr) of HIV-1 either to the C- or N- terminus of A3C. Interestingly both, A3C and Vpr.3C were incorporated into viral particles and found to be localized in the same viral compartment, the viral core. These data indicate that HIV encodes a factor additional to *vif* to counteract the A3C activity. In search for such a viral factor, the viral integrase was identified as a candidate inhibitor of A3C.

Additional experiments are required for a better understanding of the interaction of A3C with HIV. However, the results here implicate that enhancing the antiviral activity of the ubiquitously expressed A3C would likely repress HIV-1 replication in patients and thus should be considered as a novel approach for treatment.

Introduction

In the year 1981 a new epidemic in human history, the acquired immunodeficiency syndrome (AIDS), appeared and 3 years later it was found that a novel retrovirus was the aetiologic agent. Infection with the *Human immunodeficiency virus type 1* (HIV-1) causes a loss of CD4 positive T-cells, thereby weakening the human immune system and enhancing the risk of future opportunistic infections. HIV belongs to a large family of retroviruses including simian immunodeficiency viruses (SIVs). HIV-1 developed at least three times in humans by three independent transmissions from an SIV infected chimpanzee to humans forming the HIV-1 clades M, N and O. The transmission of HIV between humans mostly occurs through unsafe sex and from contaminated needles from injection drug users. In 2008 an estimated number of 33.4 million people (UN-AIDS; WHO) in the world were infected with HIV while another 25 million people already died of this disease. Most of the infected people, in numbers 22.4 million, live in Sub-Saharan Africa. Although the expanding understanding of the HIV life cycle has led to the development of antiretroviral drugs like inhibitors of the viral enzymes protease and reverse transcriptase, there is no vaccine or cure for HIV and AIDS available. Currently HIV infected people are treated with the highly active antiretroviral therapy (HAART), a therapy that can prolong the life of infected people to a mostly normal lifespan.
Since the discovery of HIV, much work has been done to understand the replication of HIV in humans and to investigate the interactions of the virus with its host. Understanding the interaction of the virus with host cells restriction factors is a promising new aspect in the HIV research. Host cell restriction factors of non-human species are able to prevent viral infections by blocking different steps of the viral life cycle. Adapted viruses such as HIV to human overcome these restriction factors through accessory gene products that counteract the antiviral restriction. The characterization of the molecular basis of the interaction of viral with cellular determinants can highlight new targets for HIV therapy. This study focuses on the molecular understanding of the interaction between the cellular restriction factor APOBEC3C (A3C) and HIV or SIV.

A| RETROVIRUSES

Retroviruses are enveloped RNA viruses forming the family of *Retroviridae.* They are diploid as each virion contains two copies of its ssRNA genome. During replication the retroviral genome is reverse transcribed into DNA by the viral enzyme reverse transcriptase (RT). The viral integrase (IN) catalyzes the integration of viral DNA into the host cell genome. The integration of the proviral DNA results in a lifelong persistent infection of the host.

The basic genome of retroviruses is composed of three genes: *gag* (*g*roup specific *a*nti*g*en), *pol* (*pol*ymerase) and *env* (*env*elope) (Fig. A-1). The provirus of all replication competent retroviruses is flanked by LTRs (long terminal repeats).

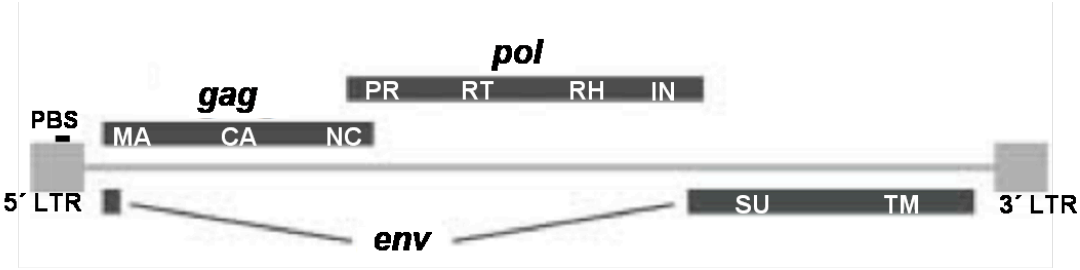


Fig. A-1: Schematic overview of the genome organization of a simple retrovirus. LTRs are depicted in light grey, in dark grey the genes *gag, pol* and *env* are shown with their protein products (white letters). The primer binding site (PBS) is shown as small bar. (Figure modified from Malim, M., 2009)

The *gag* gene encodes for structural proteins of the virus: matrix (MA), capsid (CA) and nucleocapsid (NC). The genetic information for the viral enzymes protease (PR), reverse transcriptase (RT), RNaseH (RH) and integrase (IN) is located on the *pol* gene. The two envelope proteins encoded by the *env* gene are the surface glycoprotein (SU) and the transmembrane glycoprotein (TM).

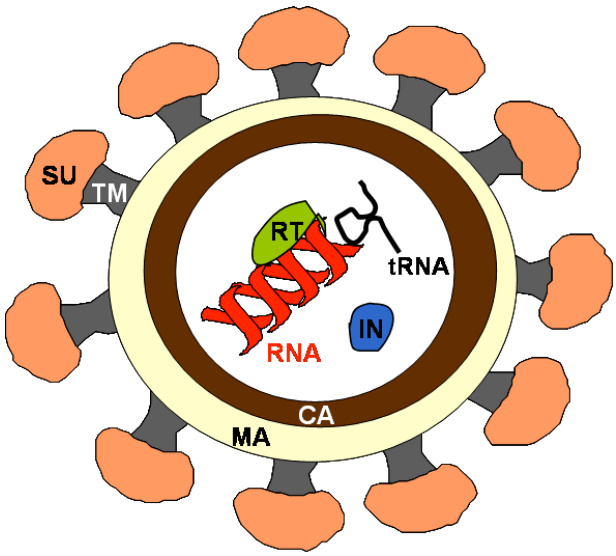


Fig. A-2: Composition of a mature retroviral particle. The dimeric RNA genome together with a cellular tRNA and the viral enzymes IN and RT is surrounded by CA proteins build the viral core. A layer of MA proteins separates the core from the envelope proteins (SU and TM).

The viral particle contains a dimer of the viral ssRNA as well as host cell RNA (Fig.2). This host cell RNA is a transfer RNA (tRNA) which is bound to the PBS of the viral RNA. The tRNA is required for the initiation of reverse transcription. Some viral proteins are associated with the RNA, such as NC that coats the RNA and the viral enzymes IN and RT. A layer of capsid (CA) proteins forms the viral core casing this ribonucleic acid-protein (RNP)-complex. The capsid proteins are surrounded by a layer of MA-proteins building a border between CA and envelope proteins. The envelope of retroviruses is build of two, non-covalently linked proteins, a TM protein and the heavily glycosylated SU protein which is the ligand for the cellular receptors and responsible for the recognition of the target cell. The retroviral life cycle will be illustrated using the *Human immunodeficiency virus* (HIV) by way of example.

B| HUMAN IMMUNODEFICIENCY VIRUS (HIV)

HIV is a complex retrovirus that belongs to the genus of *Lentiviruses*. In addition to the basic retroviral genes *gag*, *pol* and *env* the HIV genome encodes for accessory genes as shown in Fig. B-1. The gene products of these accessory genes play important roles in the HIV life cycle as they either directly support viral replication like (Tat or Rev) (1-3) or counteract host cell restriction factors (like Vif or Vpu) (4, 5). The biological relevant functions of Vpr and Nef are less clear.

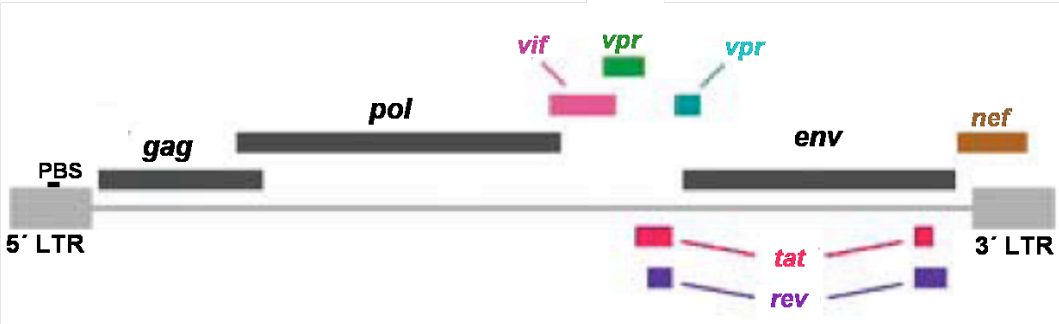


Fig. B-1: Schematic representation of the HIV-1 genome flanked by the 5´ and 3´ LTR regions (light grey). *Gag, pol* and *env* genes are shown in dark grey. HIV encodes the accessory genes *vif* (light pink), *vpr* (green), *tat* (dark pink), *rev* (purple), *vpu* (blue) and *nef* (brown) in addition to the basic retroviral genes. The PBS is shown as small bar. (Figure modified from (6)).

Cells of the human immune system expressing the cellular receptor CD4 ($CD4^+$ cells), like T-helper cells, macrophages or dentritic cells, are the target cells for HIV infection. HIV infection can be separated into three steps: the acute infection, a latency phase and the progression of the Acquired Immunodeficiency Syndrome (AIDS) (Fig B-2). In early HIV infection the number of $CD4^+$ cells dramatically drops down, either as a result of apoptosis triggered by the viral proteins Nef and Vpr or as a result of $CD8^+$ T-cell recognition and killing (reviewed in (7) and (8)). During this phase of acute viral infection the virus spreads, being detectable through a high viral load in the plasma of patients. The phase of acute infection lasts for several weeks before the next stage of HIV infection, the latency phase, begins which in average last for years. In this phase the viral load decreases, due to a strong defence of the host immune system. Viral latency lasts until the number

of $CD4^+$ T-cells drops below a limit where cell-mediated immunity gets lost and the progression of AIDS begins (reviewed in (7)). Loss of $CD4^+$ T-cell mediated immunity enhances the risk of future opportunistic infections.

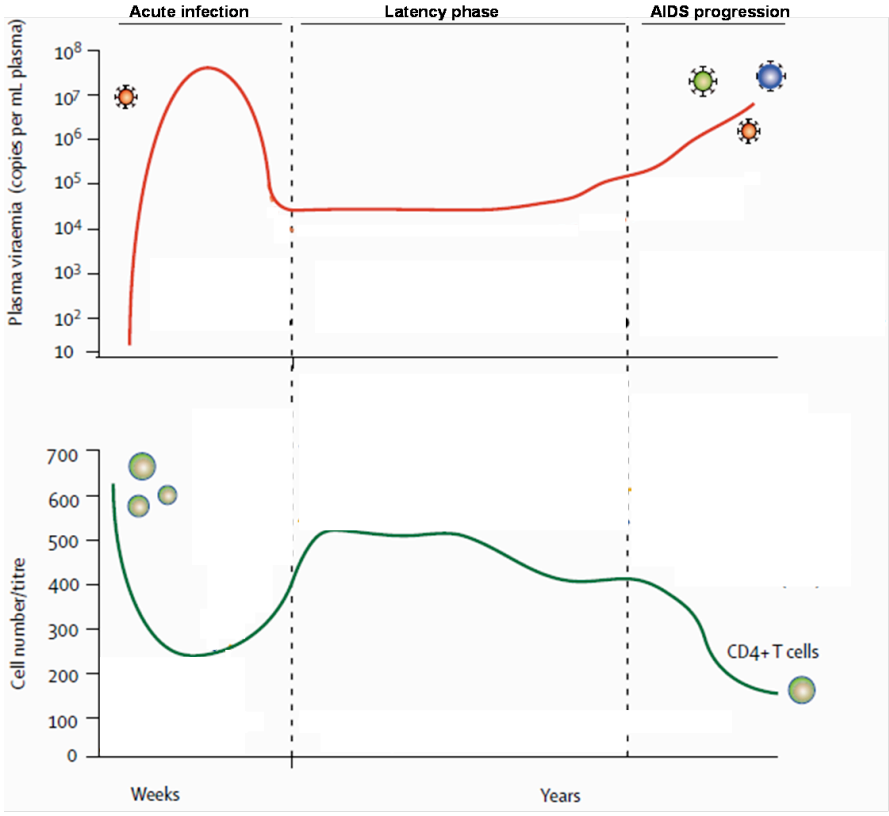


Fig. B-2: The course of HIV infection. The acute infection is characterized by a high plasma viremia (red line, top) and a decrease in $CD4^+$ cells (green line, bottom). After the latency phase displayed through a recovery of the $CD4^+$ cells and a lowered viral load, the amount of $CD4^+$ drops down and the number of viruses in the plasma increases again resulting in AIDS progression. (Figure modified from (8)).

To establish preventive or therapeutic measures against HIV infections a closer look on the viral replication is necessary (Fig. B-3). The HIV life cycle starts with the infection of $CD4^+$ target cells, either as a free virus particle or transmitted by dendritic cells within the virologic synapse. The binding of the HIV surface glycoprotein gp120 to a cellular CD4 receptor induces a rearrangement of the envelope complex allowing another domain of gp120 to attach to a chemokine receptor, either CCR5 or CXCR4 (9). This interaction allows gp41 to penetrate the cell membrane resulting in the fusion with the viral membrane (reviewed in (10)). After fusion the viral core is released into the cytoplasm. Prior to reverse transcription uncoating of the viral core occurs. Binding of the tRNA to the PBS is required for reverse transcription and the reaction is catalyzed through the viral reverse transcriptase (11). Viral ssRNA thereby gets transcribed into dsDNA. The nuclear import of the cDNA is triggered by a pre-integration complex that is transported into the nucleus via nuclear pores. In the nucleus the IN catalyzes the integration of the viral DNA genome into the chromosomal genome of the host cell resulting in the provirus formation (12). Transcription of the integrated provirus is initiated by cellular transcription factors that interact with the viral promoter in the 5´ LTR region.

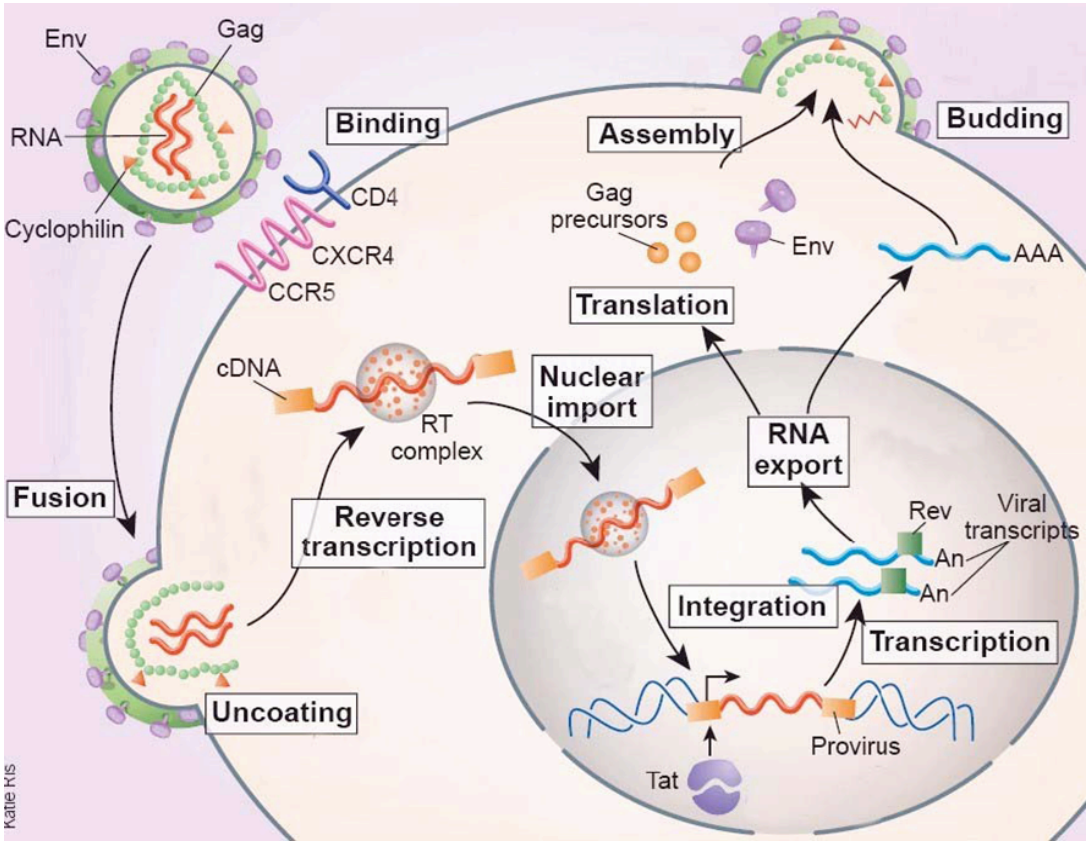


Fig B-3: Replication cycle of HIV. Important steps are highlighted with white boxes. (Figure modified from (7))

The first viral transcripts are multiply spliced and exported from the nucleus to the cytoplasm where the early gene expression of Tat, Rev and Nef occurs (13). Tat and Rev both have a nuclear localization signal that targets them to the nucleus. Tat enhances the proviral transcription (2, 3) and Rev binds to unspliced or single spliced proviral RNA and thereby promotes their transport to the cytoplasm (1). In the cytoplasm the late gene products like Vif, Vpu, Vpr and Env are translated from the single spliced RNAs, whereas the unspliced RNA is translated into the gag-pol polyprotein. In order to facilitate the assembly of newly produced viral particles, two molecules of the unspliced viral RNA interact with the NC of gag-pol polyproteins which become anchored in the plasma membrane (reviewed in (14)). Through the recruitment of the cellular ESCRT (= endosomal sorting complex required for transport) proteins budding of the virion from the cellular membrane takes place (reviewed in (15)). Generation of an infectious viral particle (maturation) occurs after virus release through cleavage of the gag-pol polyproteins by the viral protease (14).
Certain steps in the viral life cycle are targets of known host cell restriction factors. An overview of restriction factors with special focus on APOBEC3 proteins is given in the following.

C| RESTRICTION FACTORS AND APOBEC3 PROTEINS

Some cells are non-permissive for specific viruses. Non-permissivity is based on the absence of essential co-factors or due to the presence of factors restricting viral replication in these cells. These restriction factors can block viruses at different steps of the viral life cycle and adapted viruses have evolved strategies to circumvent these restrictions. HIV-1 is able to counteract

different restriction factors with its own accessory proteins. A lack of such a protein inhibits replication in non-permissive cells, but is not inhibitory in permissive cells.

One example for a restriction factor counteracted by a viral protein is tetherin (also known as BST-2 or CD317). This protein was first described by Neil *et al* 2008 (16). Tetherin-expressing cells were non-permissive for HIV∆*vpu*, whereas *vpu*-deleted viruses could replicate in permissive cells that did not express tetherin. In the absence of the viral protein U (Vpu) HIV particles are tethered at the cellular membrane and viral release is impaired. Although it was observed, that Vpu down-regulates cellular tetherin levels (17-19) and that both proteins show physical interaction (17), it remains unclear by which exact mechanism Vpu counteracts the restriction of tetherin. Additionally, lentiviruses that do not encode Vpu can also counteract the tetherin mediated restriction. In the case of HIV-2, a Vpu-like activity is encoded in its Env glycoprotein (20-22) and the *Simian immunodeficiency virus* from African Green Monkeys (SIVagm) and from Macaques (SIVmac) are able to circumvent the tetherin restriction via their Nef proteins (23, 24). Another example for a complex and diverse virus-host cell interaction is the family of APOBEC3 restriction proteins, which are subject of this study.

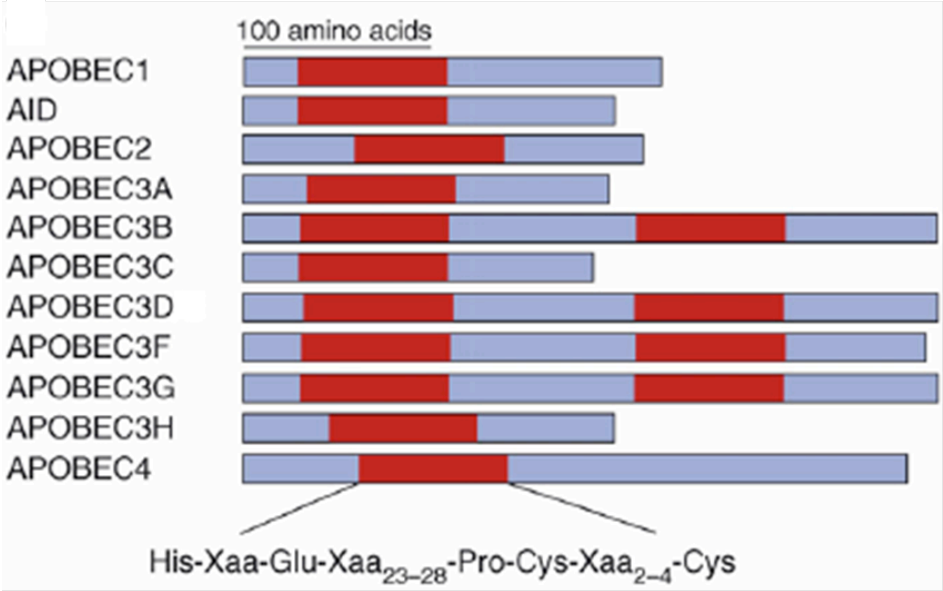


Fig. C-1: Schematic representation of the human APOBEC protein family. The Zn^{2+} -coordinating motifs are shown in red and the crucial, Zn^{2+} - binding amino acids are highlighted. (Figure taken from (25))

The family of APOBEC3 (apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like 3) restriction factors was identified through the observation, that ∆*vif* HIV was able to infect the cell line CEM-S (permissive) but not the CEM-SS cell line (non-permissive) (26). This observation led to the idea of a restriction factor that is expressed in the non-permissive CEM-SS cells and can be inactivated by the Vif protein of HIV, because *w*ild*t*ype (*wt*) HIV was able to replicate in CEM-SS cells. Screening of the expression profiles of these near-isogenic cell lines resulted in the identification of the protein APOBEC3G (A3G). This cytidine deaminase was expressed in CEM-SS cells, but not in the permissive CEM-S cells. A3G is a member of the human APOBEC cytidine deaminase family that consists of 11 members shown in Figure C-1.

Although this family is very homologous sharing one or two Zn^{2+} coordinating domains the proteins exhibit different enzymatic activities. APOBEC1 (A1), the founder of this protein family, is described as a RNA deaminase expressed in gastrointestinal tissue (27). A1 deaminates the cytosine 6666 in apolipoprotein mRNA resulting in an in-frame stop codon (28, 29). The truncated and the full-length protein have different functions in the lipid metabolism. The DNA editing activity of the activation induced deaminase (AID) in B-cells leads to somatic hypermutation and class-switch recombination of antibodies, and thus plays a key role in the maturation of the natural antibody response (30-32). Whereas APOBEC2 was shown to be necessary for muscle development (33), the function of APOBEC4 (A4) is still unknown. The human APOBEC3 (A3) family has been described as antiviral proteins restricting retroviruses and retrotransposons. The antiviral activity of A3 proteins is ascribed to the deamination of cytidine to uridine of single-stranded viral DNA formed during

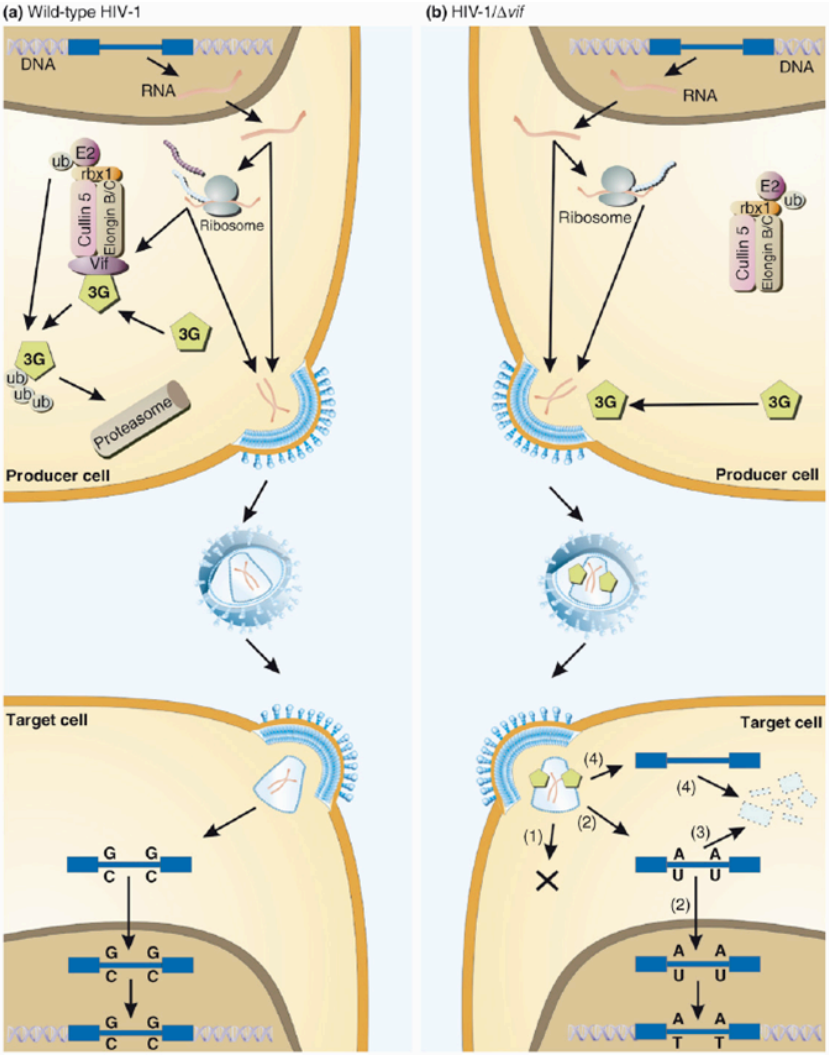


Fig. C-2: APOBEC-dependent restriction of HIV infection. (a) In cells infected with *wt* HIV A3 proteins are excluded from budding viral particles due to Vif-induced proteasomal degradation. (b) A3 proteins present in producer cells get incorporated into Vif-deficient viral particles and deaminate the viral genome after infection of a target cell. (Figure taken from (25))

reverse transcription (34-37). This DNA editing either leads to non-functional viral genes as a result of extensive G to A hypermutation on the coding strand of viral DNA or to degradation of viral DNA

by the DNA base repair enzymes uracil DNA glycosylase and apurinic-apyrimidinc endonuclease (38). A prerequisite for this antiviral activity is the incorporation of the A3 proteins from the producer cell into budding virions. In most cases, A3s being present in viral particles deaminate the viral genome after infection of a target cell (Fig. C-2b). Besides editing other mechanisms of A3 mediated viral restriction, like inhibition of integration or reverse transcription have been discussed (39-43). The Vif protein defeats the activity of A3 proteins by binding them and recruiting an E3 ubiquitin ligase complex (44). A3 proteins get poly-ubiquitylated and subsequent degraded by the proteasome (5, 45)(Fig C-2a). Thus A3 proteins are no longer incorporated into viral particles and the virus remains fully infectious.

Interestingly, the Vif – APOBEC3 interaction is often species-specific. Human APOBEG3G (hA3G) for example is recognized by the Vif_{HIV-1}, but the Vif protein of the *Simian immunodeficiency virus*

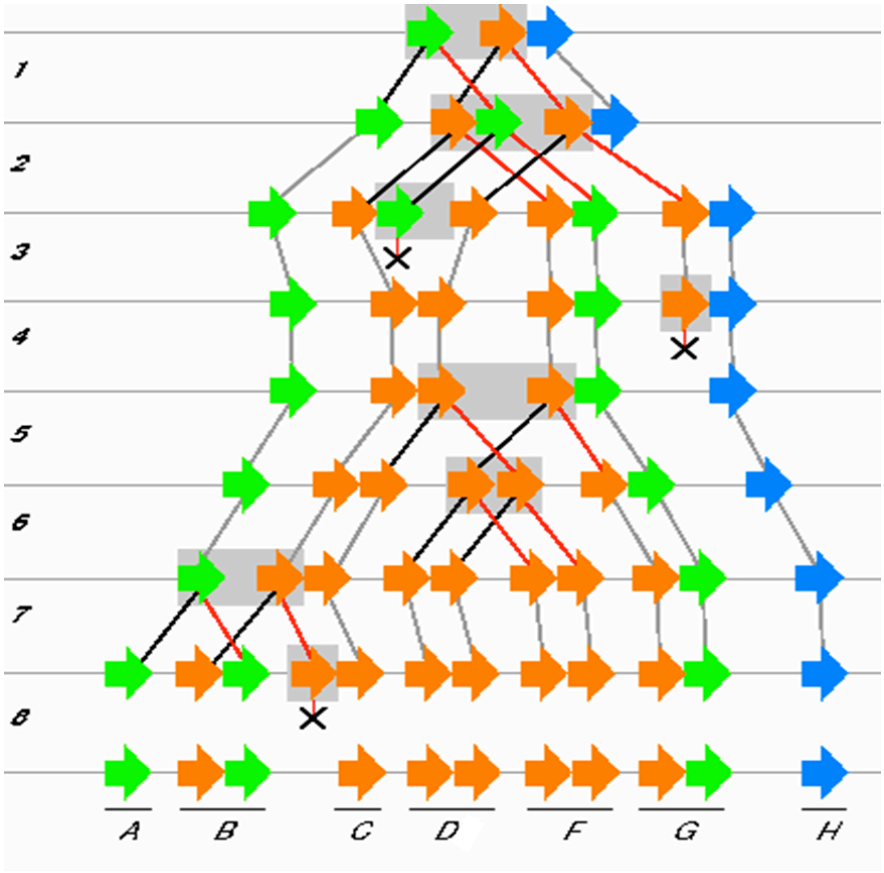


Fig. C-3: A model for the evolutionary origin of human APOBEC3 genes located on chromosome 22. The ancestor domains of A3A, A3C and A3H are colored in green, orange and blue, respectively. Black lines indicate duplications within the original segment, red lines represent duplication events resulting in a duplicated segment and grey lines indicate no change. Black crosses symbolize gene deletions. (Figure taken from (46))

from African Green Monkeys (SIVagm) is not capable to induce the degradation of hA3G (47). *Vice versa*, agmA3G is only recognized by Vif_{SIVagm}, but not by Vif_{HIV-1}. The restriction phenotype of the individual human A3 proteins highly differs from each other. The most prominent member of the human APOBEC3 protein family is A3G that exhibits antiviral activity against Δ*vif* HIV, many other retroviruses and retroelements (for review (25); see table 1 and references within). A3G has two

Zn^{2+} coordinating motifs, a C-terminal domain (CTD) and a N-terminal domain (NTD), with different activities (48-50). While the NTD is required for its incorporation into viral particles, the CTD catalyzes the deamination of the viral genome. Another member of the A3 protein family, A3C, contains only one Zn^{2+} coordinating domain and is not active against Δ*vif* HIV, but shows a robust restriction of Δ*vif* SIV (51) and Long Interspersed Element-1 (LINE-1) retrotransposons (52). Whether A3 proteins have one or two domains is a result of gene-duplications and –deletions from 3 ancestor A3 proteins A3Z1, A3Z2 and A3Z3 (46)(Fig. C-3).

Interestingly, the ancestor domain of A3C, A3Z2, is present in each of the two-domain APOBEC3 proteins, showing that A3C is an evolutionary highly conserved protein. Additionally, it is broadly expressed in many human tissues (53), a variety of cancer cell lines (54) and is up regulated in $CD4^+$ cells, the target cells of HIV (55). As HIV is able circumvent the A3C mediated restriction also in absence of its Vif protein (51), there is likely an additional factor encoded by HIV that can counteract the antiviral activity of A3C. Through these observations A3C has become an interesting candidate to explore virus-host interactions.

D| SCOPE OF THIS WORK

This study aimed to understand the interaction of the restriction factor APOBEC3C with HIV and SIV. Therefore residues of A3C crucial for its antiviral activity were identified based on a structural model, the different restriction phenotypes of A3C against HIV and SIV and putative mechanisms of antiviral restriction were analyzed. In the different chapters the following questions were answered or aimed to answer:

CHAPTER I: Which amino acids and protein structures are crucial for the antiviral activity of A3C against SIV Δ*vif*?

CHAPTER II: Why is A3C antiviral against SIV Δ*vif*, but not against HIV Δ*vif*?

CHAPTER III: What is the restriction mechanism of A3C against SIV Δ*vif* and is the Zn^{2+}-coordinating domain the enzymatic active site of the protein?

References

1. Pollard VW & Malim MH (1998) The HIV-1 Rev protein. *Annu. Rev. Microbiol.* 52, 491-532.
2. Dayton AI *et al.* (1986) The trans-activator gene of the human T cell lymphotropic virus type III is required for replication. *Cell* 44, 941-947.
3. Fisher AG *et al.* (1986) The trans-activator gene of HTLV-III is essential for virus replication. *Nature* 320, 367-371.
4. Neil SJ, Zang T & Bieniasz PD (2008) Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. *Nature* 451, 425-430.
5. Yu X *et al.* (2003) Induction of APOBEC3G ubiquitination and degradation by an HIV-1 Vif-Cul5-SCF complex. *Science* 302, 1056-1060.
6. Malim MH (2009) APOBEC proteins and intrinsic resistance to HIV-1 infection. *Philos. Trans. R. Soc. Lond B Biol. Sci.* 364, 675-687.
7. Stevenson M (2003) HIV-1 pathogenesis. *Nat. Med.* 9, 853-860.

8. Simon V, Ho DD & Abdool KQ (2006) HIV/AIDS epidemiology, pathogenesis, prevention, and treatment. *Lancet* 368, 489-504.

9. Clapham PR & McKnight A (2002) Cell surface receptors, virus entry and tropism of primate lentiviruses. *J. Gen. Virol.* 83, 1809-1829.

10. Sierra S, Kupfer B & Kaiser R (2005) Basics of the virology of HIV-1 and its replication. *J. Clin. Virol.* 34, 233-244.

11. Harrich D & Hooker B (2002) Mechanistic aspects of HIV-1 reverse transcription initiation. *Rev. Med. Virol.* 12, 31-45.

12. Van MB & Debyser Z (2005) HIV-1 integration: an interplay between HIV-1 integrase, cellular and viral proteins. *AIDS Rev.* 7, 26-43.

13. Purcell DF & Martin MA (1993) Alternative splicing of human immunodeficiency virus type 1 mRNA modulates viral protein expression, replication, and infectivity. *J. Virol.* 67, 6365-6378.

14. Freed EO (2001) HIV-1 replication. *Somat. Cell Mol. Genet.* 26, 13-33.

15. Bieniasz PD (2009) The cell biology of HIV-1 virion genesis. *Cell Host. Microbe* 5, 550-558.

16. Neil SJ, Zang T & Bieniasz PD (2008) Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. *Nature* 451, 425-430.

17. Douglas JL *et al.* (2009) Vpu directs the degradation of the human immunodeficiency virus restriction factor BST-2/Tetherin via a {beta}TrCP-dependent mechanism. *J. Virol.* 83, 7931-7947.

18. Miyagi E, Andrew AJ, Kao S & Strebel K (2009) Vpu enhances HIV-1 virus release in the absence of Bst-2 cell surface down-modulation and intracellular depletion. *Proc. Natl. Acad. Sci. U. S. A* 106, 2868-2873.

19. Van DN *et al.* (2008) The interferon-induced protein BST-2 restricts HIV-1 release and is downregulated from the cell surface by the viral Vpu protein. *Cell Host. Microbe* 3, 245-252.

20. Abada P, Noble B & Cannon PM (2005) Functional domains within the human immunodeficiency virus type 2 envelope protein required to enhance virus production. *J. Virol.* 79, 3627-3638.

21. Bour S & Strebel K (1996) The human immunodeficiency virus (HIV) type 2 envelope protein is a functional complement to HIV type 1 Vpu that enhances particle release of heterologous retroviruses. *J. Virol.* 70, 8285-8300.

22. Bour S, Schubert U, Peden K & Strebel K (1996) The envelope glycoprotein of human immunodeficiency virus type 2 enhances viral particle release: a Vpu-like factor? *J. Virol.* 70, 820-829.

23. Jia B *et al.* (2009) Species-specific activity of SIV Nef and HIV-1 Vpu in overcoming restriction by tetherin/BST2. *PLoS. Pathog.* 5, e1000429.

24. Zhang F *et al.* (2009) Nef proteins from simian immunodeficiency viruses are tetherin antagonists. *Cell Host. Microbe* 6, 54-67.

25. Holmes RK, Malim MH & Bishop KN (2007) APOBEC-mediated viral restriction: not simply editing? *Trends Biochem. Sci.* 32, 118-128.

26. Sheehy AM, Gaddis NC, Choi JD & Malim MH (2002) Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature* 418, 646-650.

27. Nakamuta M *et al.* (1995) Alternative mRNA splicing and differential promoter utilization determine tissue-specific expression of the apolipoprotein B mRNA-editing protein (Apobec1) gene in mice. Structure and evolution of Apobec1 and related nucleoside/nucleotide deaminases. *J. Biol. Chem.* 270, 13042-13056.

28. Knott TJ *et al.* (1986) Complete protein sequence and identification of structural domains of human apolipoprotein B. *Nature* 323, 734-738.

29. Yang CY *et al.* (1986) Sequence, structure, receptor-binding domains and internal repeats of human apolipoprotein B-100. *Nature* 323, 738-742.

30. Martin A *et al.* (2002) Activation-induced cytidine deaminase turns on somatic hypermutation in hybridomas. *Nature* 415, 802-806.

31. Muramatsu M *et al.* (2000) Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell* 102, 553-563.

32. Okazaki IM *et al.* (2002) The AID enzyme induces class switch recombination in fibroblasts. *Nature* 416, 340-345.

33. Sato Y *et al.* (2010) Deficiency in APOBEC2 leads to a shift in muscle fiber type, diminished body mass, and myopathy. *J. Biol. Chem.* 285, 7111-7118.

34. Bishop KN *et al.* (2004) Cytidine deamination of retroviral DNA by diverse APOBEC proteins. *Curr. Biol.* 14, 1392-1396.

35. Lecossier D, Bouchonnet F, Clavel F & Hance AJ (2003) Hypermutation of HIV-1 DNA in the absence of the Vif protein. *Science* 300, 1112.

36. Mangeat B *et al.* (2003) Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts. *Nature* 424, 99-103.

37. Zhang H *et al.* (2003) The cytidine deaminase CEM15 induces hypermutation in newly synthesized HIV-1 DNA. *Nature* 424, 94-98.

38. Yang B *et al.* (2007) Virion-associated uracil DNA glycosylase-2 and apurinic/apyrimidinic endonuclease are involved in the degradation of APOBEC3G-edited nascent HIV-1 DNA. *J. Biol. Chem.* 282, 11667-11675.

39. Bishop KN, Holmes RK & Malim MH (2006) Antiviral potency of APOBEC proteins does not correlate with cytidine deamination. *J. Virol.* 80, 8450-8458.

40. Guo F *et al.* (2006) Inhibition of formula-primed reverse transcription by human APOBEC3G during human immunodeficiency virus type 1 replication. *J. Virol.* 80, 11710-11722.

41. Holmes RK, Koning FA, Bishop KN & Malim MH (2007) APOBEC3F can inhibit the accumulation of HIV-1 reverse transcription products in the absence of hypermutation. Comparisons with APOBEC3G. *J. Biol. Chem.* 282, 2587-2595.

42. Mbisa JL *et al.* (2007) Human immunodeficiency virus type 1 cDNAs produced in the presence of APOBEC3G exhibit defects in plus-strand DNA transfer and integration. *J. Virol.* 81, 7099-7110.

43. Mbisa JL, Bu W & Pathak VK (2010) APOBEC3F and APOBEC3G inhibit HIV-1 DNA integration by different mechanisms. *J. Virol.* 84, 5250-5259.

44. Yu Y *et al.* (2004) Selective assembly of HIV-1 Vif-Cul5-ElonginB-ElonginC E3 ubiquitin ligase complex through a novel SOCS box and upstream cysteines. *Genes Dev.* 18, 2867-2872.

45. Marin M, Rose KM, Kozak SL & Kabat D (2003) HIV-1 Vif protein binds the editing enzyme APOBEC3G and induces its degradation. *Nat. Med.* 9, 1398-1403.

46. LaRue RS *et al.* (2008) The artiodactyl APOBEC3 innate immune repertoire shows evidence for a multi-functional domain organization that existed in the ancestor of placental mammals. *BMC. Mol. Biol.* 9, 104.

47. Mariani R *et al.* (2003) Species-specific exclusion of APOBEC3G from HIV-1 virions by Vif. *Cell* 114, 21-31.

48. Hache G, Liddament MT & Harris RS (2005) The retroviral hypermutation specificity of APOBEC3F and APOBEC3G is governed by the C-terminal DNA cytosine deaminase domain. *J. Biol. Chem.* 280, 10920-10924.

49. Navarro F *et al.* (2005) Complementary function of the two catalytic domains of APOBEC3G. *Virology* 333, 374-386.

50. Gooch BD & Cullen BR (2008) Functional domain organization of human APOBEC3G. *Virology* 379, 118-124.

51. Yu Q *et al.* (2004) APOBEC3B and APOBEC3C are potent inhibitors of simian immunodeficiency virus replication. *J. Biol. Chem.* 279, 53379-53386.

52. Muckenfuss H *et al.* (2006) APOBEC3 proteins inhibit human LINE-1 retrotransposition. *J. Biol. Chem.* 281, 22161-22172.

53. Koning FA *et al.* (2009) Defining APOBEC3 expression patterns in human tissues and hematopoietic cell subsets. *J. Virol.* 83, 9474-9485.

54. Chiu YL & Greene WC (2008) The APOBEC3 cytidine deaminases: an innate defensive network opposing exogenous retroviruses and endogenous retroelements. *Annu. Rev. Immunol.* 26, 317-353.

55. Refsland EW *et al.* (2010) Quantitative profiling of the full APOBEC3 mRNA repertoire in lymphocytes and tissues: implications for HIV-1 restriction. *Nucleic Acids Res.* 38, 4274-4284.

CHAPTER I

Model Structure of APOBEC3C Reveals a Binding Pocket Modulating RNA Interaction Required for Encapsidation

Benjamin Stauch[1,2]*, Henning Hofmann[3,4]*, Mario Perković[3,4], Martin Weisel[1], Ferdinand Kopietz[3], Klaus Cichutek[3], Carsten Münk[3,4†], Gisbert Schneider[1†]

[1]Johann Wolfgang Goethe-University Frankfurt, Chair for Chem- and Bioinformatics, Siesmayerstr. 70, 60323 Frankfurt am Main, Germany. [2]present address: EMBL Heidelberg, Structural and Computational Biology Unit, Meyerhofstr. 1, 69117 Heidelberg, Germany. [3]Paul-Ehrlich-Institute, Division of Medical Biotechnology, Paul-Ehrlich-Str. 51-59 63225, Langen, Germany. [4]present address: Heinrich-Heine-University Düsseldorf, Clinic for Gastroenterology, Hepatology and Infectiology, Building 23.12_U1 Room 81, Moorenstr. 5, 40225 Düsseldorf, Germany.

* BS And HH contributed equally to this work

† Corresponding authors: CM and GS

ABSTRACT

Human APOBEC3 (A3) proteins form part of the intrinsic immunity to retroviruses. Carrying one or two copies of a cytidine deaminase motif, A3s act by deamination of retroviral genomes during reverse transcription. HIV-1 overcomes this inhibition by the Vif-protein, which prevents incorporation of A3 into virions. In this study, we modeled and probed the structure of APOBEC3C (A3C), a single-domain A3 with strong anti-lentiviral activity. The three-dimensional protein model was used to predict the effect of mutations on antiviral activity, which was tested in a Δ*vif* SIV-reporter-virus assay. We found that A3C-activity requires protein dimerization for antiviral activity against SIV. Furthermore, by using a structure-based algorithm for automated pocket extraction we detected a putative substrate binding-pocket of A3C distal from the zinc-coordinating deaminase motif. Mutations in this region diminished antiviral activity by excluding A3C from virions. We found evidence that the small 5.8S RNA specifically binds to this locus and mediates incorporation of A3C into virus particles.

KEY WORDS

Bioinformatics, immunodeficiency, protein-protein interaction, protein structure, retrovirus

INTRODUCTION

One of the best-characterized cellular proteins efficiently restricting *Human immunodeficiency virus type-1* (HIV-1) is APOBEC3G (A3G) (1). Encapsidation of A3G in HIV-1 virus particles leads to deamination of cytosine residues to uracil in growing single-stranded DNA during reverse transcription (2-6). A3G has additional, still ill-defined antiviral activities (7). HIV-1 uses the viral infectivity factor (Vif) to prevent or reduce incorporation of A3G into progeny virions (4, 8, 9).

The human genome contains seven APOBEC3 (A3) genes which can be classified according to the presence of the Z1, Z2 and Z3 zinc-coordinating motifs (10, 11). Z2, the A3C family, consists of A3C, the C- and N-terminal domains of A3DE and A3F, and the N-terminal domains of A3B and A3G. The Z1 group, the A3A family, contains A3A and the C-terminal domains of A3B and A3G. A3H represents the Z3 zinc-finger domain. Accordingly, A3B, A3G, A3DE and A3F have two domains, while A3A, A3C and A3H possess only one domain (12). In the human A3 locus there is evidence for gene expansion, and it was speculated that duplications of single-domain genes led to the evolution of the two-domain A3s (13). Phylogentic analysis of primate and non-primate antiviral cytidine deaminases showed that in the early evolution of mammals genes for A3C (Z2), A3A (Z1) and A3H (Z3) were already present (11). Among these antetype A3s, human A3A and most variants of A3H are not antiviral against HIV (14-16). While A3C is packaged into Δ*vif* HIV with a weak antiviral effect (17), A3C is a strong inhibitor of Δ*vif* SIV (18). The study of A3C gains further importance from the fact that for A3s until now only structures of Z1-derived domains, e.g. A3G-CD, have been solved experimentally. Notably, both A3C and the still ill-defined N-terminal domain of A3G are of type Z2. A study by Bourara *et al.* (19) shows that in target cells A3C can induce limited G-to-A mutations in HIV. These mutations do not block viral replication, but rather contribute to viral diversity.

Fundamental biochemical aspects of the A3 protein structure and their relevance for antiviral activity are still a matter of discussion. Here, we performed comparative protein modeling of A3C, and assessed the model using A3C-mutants in the SIV_{agm} system. This study provides a first structural basis for rational antiviral intervention targeting A3C. We found evidence that A3C dimerization is critical for antiviral activity. Furthermore, we found a previously undescribed cavity in A3C that is similar to nucleic acid binding pockets of known enzymes. A point mutation near the pocket diminishes encapsidation of A3C and reduces 5.8S-RNA-binding. We hypothesize that the natural substrate of this pocket of A3C is a nucleic acid, possibly mediating its incorporation into the virion by interaction with nucleocapsid protein.

RESULTS

Comparative modeling of A3C

Structures of APOBEC2 (A2) and A3G, C-terminal domain (A3G-CD), have been solved experimentally (20-23). A2 crystallizes as a homo-tetramer (PDB identifier: 2NYT, 2.5 Å resolution), composed of two outer and two inner monomers, forming a dimer of dimers, each of whose β-strands form an extended β-sheet. Each monomer possesses one copy of the conserved deaminase motif H-X-E-X_{23-28}-C-X_{2-4}-C coordinating one catalytic Zn^{2+}-ion. While the overall conformations of the inner and outer monomers − chains A and C, B and D − differ only slightly (0.2-1.1 Å pairwise root mean square deviation (RMSD)), the orientation of E60 with respect to the Zn^{2+} differs remarkably, possibly representing a molecular switch between the active (outer monomers) and inactive (inner monomers) conformation (22). We thus choose chains B and D as possible templates for comparative modeling. Since several residues are not resolved in chain D, chain B was chosen as the template for the A3C model.

Recently, two solution structures (PDB identifiers 2JYW, 2KBO) and a crystal structure (PDB identifier 3E1U, 2.3 Å resolution) of A3G-CD have been obtained (20, 21, 23). Superposition of A2 and the crystal structure of A3G-CD (21) reveals the common fold of the two polypeptides despite their relatively low sequence identity (< 30%) (Fig. 1). Notably, in the solution structure of A3G-CD, one β-strand adopts a loop conformation (20). As this segment is anticipated to be involved in domain dimerization in native full-length A3G, it has a limited suitability as a template structure for comparative modeling of A3C. Stereochemical quality was higher in the A2 structure (Fig. S1). As we anticipate A3C to form oligomers and A3G-CD was crystallized as a monomer, while A2 has been crystallized as a tetramer, oligomerization properties might be better represented by native A2 as a template than by A3G-CD. Modeling of A3C on both templates at the same time led to poor stereochemistry of the resulting models, which could not be resolved by geometry optimization (data not shown).

Alignments of A2 and A3G-CD to A3C were carried out using MODELLER (24), explicitly considering structural information of the templates (Fig. S2). We yielded favorable BLAST (25) e-values (A3C to A2: 6×10^{-23}, A3C to A3G-CD: 10^{-30}; BLOSUM62). All residues involved in the Zn^{2+} coordination are conserved in both alignments. Insertions and deletions were placed in loop regions. Alignments are supported by matching secondary structure predictions (PSIPRED (26)) to those from the crystal structures (data not shown). For both models, there is no correspondence in the templates for the N-terminal amino acids of A3C, so this region had to be constructed without template. Residue conservation was mapped back to the templates (Fig. 1a, c) and is substantially higher in the protein core and around the Zn^{2+}-coordinating center, showing a striking pattern of alternating conserved / non-conserved residues in buried / exposed parts of both α-helices (conserved: *i*, *i*+3, ...) and β-strands (conserved: *i*, *i*+2, ...).

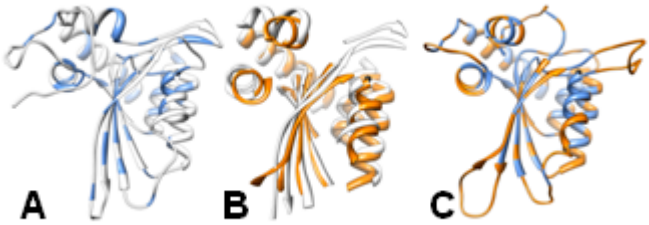


Fig. 1. (A). Structure of A2, chain B. Conserved residues in A3C are shown in blue. Black sphere: catalytic Zn^{2+}-ion. **(B)** Superposition of A2 (white) and A3G, catalytic domain (A3G-CD, orange). Loop regions are not shown. **(C)** Structure of A3G-CD. Residues conserved in A3C are shown in blue.

Ten initial models were built for each template, energy-minimized and evaluated for robustness (27). For each of the templates, the model with the fewest violations of the stereochemistry was comparable to the quality of the template structures (Fig. S1, Fig. S3) and subjected to three independent runs of 20 ns molecular dynamics (MD) simulations (Fig. S4). Convergence of RMSD and energy parameters suggested the fold to be stable (Fig. S5). Identical protocols were applied for the template structures, also showing convergence. Simulated B-factors were calculated from the trajectories for the template structures as described (28), averaged for each system and compared to the experimental B-factors reported for the X-Ray structures (r = 0.68 for A2, r = 0.76 for A3G-CD, Fig. S5), suggesting the dynamics of the structure to be well captured by our MD simulation (29). Taken together, these results suggest that both models of A3C are valid from a structural point of view and thus useful to deduce further hypotheses.

While the two minimized models show a moderate pairwise RMSD of 2.7 Å (without the N-terminal amino acids in the loop region preceding the β-sheet), all residues considered in this study are located at overlapping positions (RMSD 0.7-1.4 Å) in the two models within precision expected from given levels of sequence identity and thus are practically equivalent (Fig. 2). Here, only the model structure of A3C based on the structure of A2 is shown. All experiments conducted in this study have been replicated with the A3G-CD-based model without significant change of predictions (data not shown).

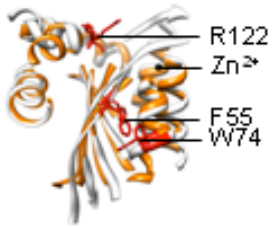


Fig. 2. Superposition of model of A3C, derived from A2 (white) and A3G-CD (orange). Residues shown to be of functional importance in this study are shown in red.

A3C functions as a dimer.

For A3C, which possesses only one domain, it is reasonable to assume a mode of dimerization analogous to that of A2, where the β-strands of two monomers build a single extended sheet. We would assume a similar mode of domain interaction in full-length A3G. We posed the questions whether i) A3C oligomerizes, and ii) there is differential

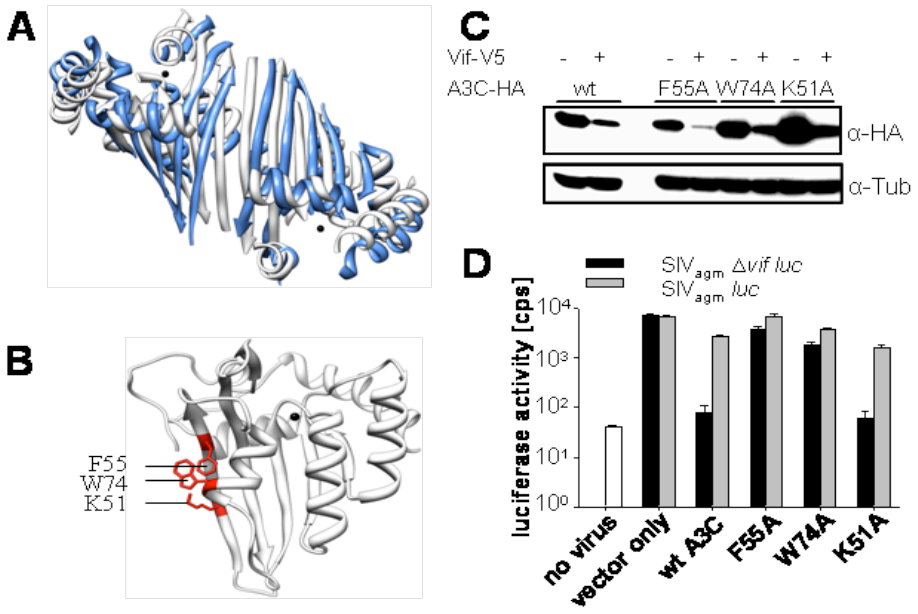


Fig. 3. (A) Dimerization pose of A2 (experimental, white) and A3C (predicted by ClusPro, blue). Loop regions are not shown. **(B)** Residues in the predicted dimerization interface of A3C (red) were mutated to alanine. **(C)** Immunoblot analysis of the expression and Vif-dependent degradation of wild-type(wt) A3C and different dimerization mutants (K51A, F55A, W74A). A3C constructs were detected by an anti(α)-HA antibody. Tubulin (Tub) served as loading control. **(D)** Antiviral activity of dimerization mutants F55A, W74A, K51A and wt A3C against SIV_{agm}, compared to non-transduced cells (no virus) and vector only control without A3C. wt or Δ*vif* SIV_{agm}*luc* (VSV-G) virions were generated by co-transfection with the respective A3C mutant. Virions were normalized by RT activity. Luciferase activity was measured 3 days post infection.

activity between the monomeric and oligomeric forms.

First, protein-protein interaction interfaces were predicted using ProMate (30). Amino acids corresponding to the A2 dimerization interface were highlighted as potentially involved in the interaction. The protein docking and clustering technique ClusPro (31) accurately reproduced the A2 dimer (RMSD = 1.7 Å), and was applied to generate an A3C dimer model *in silico* (Fig. 3a). The overall topology of both the A2 and A3C dimer models is similar.

Selected amino acids in the predicted dimerization interface of A3C were mutated. As exposed aromatic amino acids often participate in protein-protein-interaction (32), two prominent aromatic amino acids (F55, W74) and K51, possibly contributing to electrostatic interactions, were mutated to alanine (Fig. 3b). All constructs (K51A, F55A, W74A) were expressed, showing wt-like degradation in presence of Vif (Fig. 3c). To determine whether these A3C-mutants display antiretroviral activity, virions were generated by co-transfection with A3C expression plasmids. In transduced cells, A3C and K51A reduced the infectivity of the Δ*vif* SIV ~90-120-fold, F55A and W74A showed a clearly diminished inhibitory activity (~2-4-fold inhibition of Δ*vif* viruses) (Fig. 3d). In contrast, wt SIV was not inhibited by any of the A3C-mutants. Inhibition of Δ*vif* SIV by wt A3C and K51A was shown to be dose-dependent, whereas F55A and W74A were still inactive although using highest amounts of expression plasmid (Fig. 4).

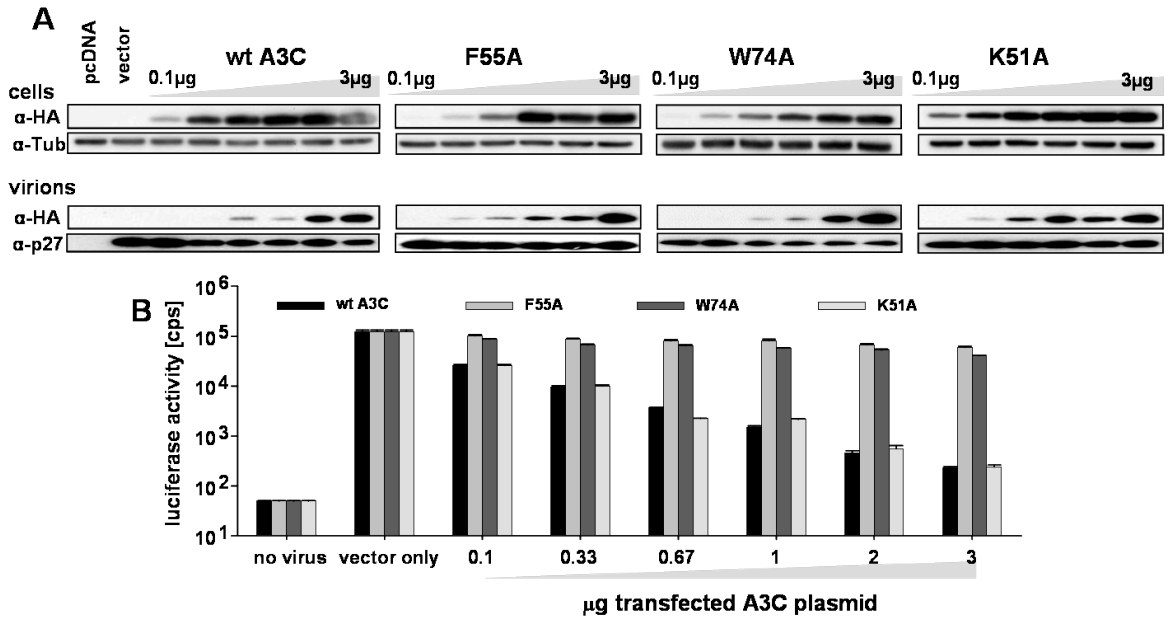


Fig. 4. (A) Immunoblot analysis of the dose-dependent expression in 293T cells and encapsidation in viral particles of *wt* A3C and different dimerization mutants (F55A, W74A and K51A). pcDNA, transfection control; vector, wt SIVagm without A3C. The respective amounts of the A3C constructs were detected by an anti(α)-HA antibody. Tubulin (Tub) and capsid (p27) served as loading control. **(B)** Dose-dependent antiviral activity of dimerization mutants F55A, W74A and K51A against SIVagm Δ*vif*, compared to wild-type(wt)-huAPOBEC3C (A3C) and background of non-transduced cells (no virus) and vector only control (vector) without A3C. Δ*vif* SIVagm *luc* (VSV-G) virions were generated by co-transfection with the indicated amount (µg) of the respective A3C mutant. Virions were normalized by RT activity and used for transduction. Luciferase activity (in counts per second, cps) was measured 3 days post infection.

Packaging of A3 into the virion is crucial for its antiretroviral activity (4, 8, 9). To determine whether missing encapsidation of the inactive mutants, is responsible for the absence of inhibition, virions were generated and analyzed for A3C content by immunoblot analysis. Fig. 5a shows that all mutants were efficiently packaged, comparable to wt A3C.

To determine whether oligomerization of the mutant proteins correlates with antiretroviral activity, expression vectors for V5-tagged wt A3C were co-transfected with expression plasmids for HA-tagged A3Cs (wt or mutants). F55A and W74A barely precipitated V5-tagged wt A3C. In contrast, K51A and HA-tagged wt A3C were able to precipitate V5-tagged wt A3C (Fig. 5b). Quantification of the precipitated V5-tagged wt A3C in presence of wt A3C compared to W74A resulted in a significant higher binding efficiency (~10 fold, Fig. S6). W74A-HA co-precipitated the V5-tagged wt A3C only very weakly. Crosslinking of wt A3C in total cell lysate of transfected cells showed monomeric, dimeric and tetrameric forms, whereas W74A only formed monomers and dimers (Fig. S6). We postulate that the W74A mutation in the dimerization interface prevents the formation of the inner dimer, but does not influence the formation of the outer dimer. This correlates with the observation of a weaker binding activity seen in immunoprecipitation assuming that the inner dimer is more stable, where as the outer

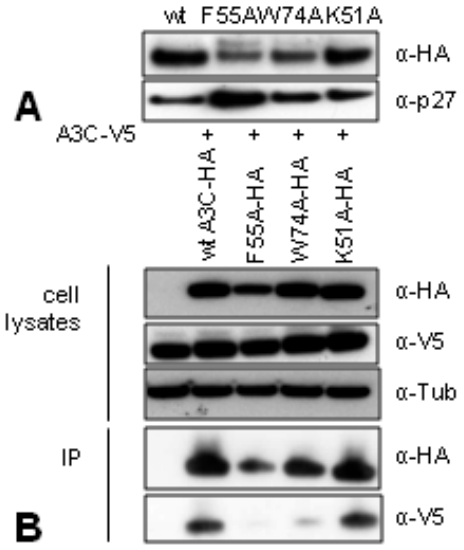


Fig. 5. (A) Immunoblot analysis of A3C packaging. 293T cells were co-transfected with SIVagm *Δvif luc* (VSV-G), the respective HA-tagged A3C construct (mutants F55A, W74A, and wild-type (wt)). Virions were harvested and normalized by RT. Physically equal amounts of virions were lysed and subjected to immunoblot analysis. Presence of A3C in the virons was detected using anti(α)-HA-antibodies. p27 (capsid) served as loading control. **(B)** Immunoblot analysis of A3C dimerization. 293T cells were co-transfected with the respective HA-tagged A3C construct (mutants K51A, F55A, W74A, and wt) and V5-tagged wt A3C. Immune precipitates (IP) were subjected to immunoblot analysis. Proteins were probed using α-HA- / -V5-antibodies, respectively. Tubulin (Tub) served as loading control in cell lysates.

dimer gets disrupted during the washing steps. These results indicate that F55 and W74 participate in dimerization of A3C, and that oligomerization is crucial for antiviral activity of the enzyme. It is noteworthy that the dimer mutants did not exhibit DNA editing activity in an *E. coli* mutation assay (compared to wt A3C, Fig. 6), which is in perfect agreement with the requirement for dimeric protein for enzymatic activity.

A cavity of A3C mediates its encapsidation.

Enzyme-substrate interactions typically occur over well-defined protein binding pockets, where the active site typically is one of the largest cavities of the protein (33). Our software PocketPicker (33) was employed to identify potential ligand binding pockets in the A3C model. The presumed active site was found to be the 5th biggest cavity in the A2-based structure (A3C/A2) (~80 Å^3). The largest cavity (~200 Å^3) is found 15 Å apart from the Zn^{2+} ion and has not yet been described in literature (Fig. 7a). The A3G-CD-based model structure of A3C contains a similar binding pocket (Fig. 7b).

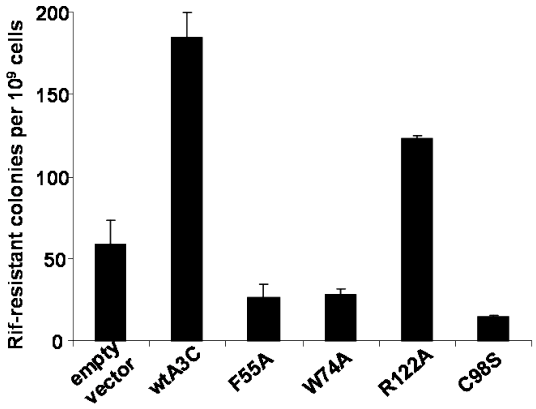


Fig. 6. Editing activity of wt and mutant A3C proteins measured as Rif^R colonies per 10^9 viable cells in *E. coli* mutation assay.

Four residues near this pocket were mutated to alanine: K22, T92, R122, and N177 (Fig. 7c). Co-expression of these mutants with Vif demonstrated protein degradation similar to wt A3C (Fig. 8a). To test whether mutant A3C-proteins inhibit virus replication, wt and Δ*vif* SIV were generated by co-transfection with A3C expression plasmids. Wild-type A3C as well as its mutants K22A, T92A, and N177A inhibited Δ*vif* SIV (~60-84-fold) but not wt SIV (Fig. 8b). In contrast, the mutant

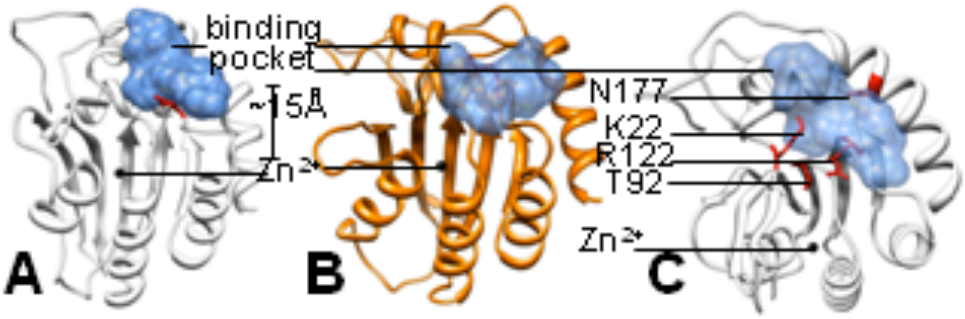


Fig. 7. (A) Model structure of A3C, derived from A2 (white). The binding pocket distal to the Zn^{2+}-ion (black sphere) is indicated in blue, R122 in red. **(B)** Model structure of A3C, derived from A3G-CD (orange). The binding pocket is indicated in blue, the Zn^{2+}-ion in black, R122 in red. **(C)** To test the function of this protein cavity, residues in red were mutated to alanine.

R122A lost the inhibitory activity. Immunoblot analysis of A3C content in viral particles showed a greatly reduced packaging of R122A compared to A3C wt and K22A, T92A, N177A (Fig. 8c), although it was detectable in cell lysates (Fig. 6a). By fusing VPR to R122A the mutant could be re-targeted into viral particles and showed antiviral activity (Fig. 8d, e). In addition, R122A showed DNA editing activity in bacteria (Fig. 6). We conclude that R122 is critically relevant for particle packaging but not for antiviral activity of A3C.

In search for potential ligands of this presumed binding pocket, it was compared to pockets extracted from the PDBBind (34) collection with known protein function and ligands. The four hits that were identified as most similar stem from human papillomavirus (HPV) type 11 E2 transactivation domain (TAD) complex (PDB identifier: 1R6N), a DNA binding protein of HPV; and three *holo*-structures of bovine RNase A (PDB identifiers: 1QHC, 1JN4, 1O0M), co-crystallized with different nucleotides. For the A3C/A3G-CD pocket, four holo-structures of bovine Rnase A (PDB identifiers: 1U1B, 1W4P, 1O0M, 1QHC) were among the eight top scoring pockets. The natural substrates of these pockets are nucleic acids: ds DNA for HPV 11 E2 TAD, and ss and dsRNA for bovine RNase A. Steric and electrostatic properties of the hypothesized pocket in A3C would allow for nucleic acids as a substrate, possibly by R122 interacting with the negatively charged sugar-phosphate backbone.

To test for RNAs interacting with this binding pocket, wt A3C and R122A were precipitated from transfected cells, interacting RNA was isolated and detected by ^{32}P-labelling (Fig. 8f). Mutation of R122 resulted in strongly decreased amounts of RNA bound to the protein compared to wt A3C or a C98S active site mutation. The isolated RNA was further subjected to RT-PCR to amplify 7SL or 5.8S RNA (Fig. 8g). A3C wt protein showed an interaction with 7SL and 5.8S RNA, while the mutant R122A lost the binding to 5.8S RNA. Furthermore, RNA binding was shown to be crucial for interaction of A3C to SIV-nucleocapsid (NC). A3C wt interacts in a RNA-dependent manner with SIV-NC, while R122A exhibits no binding activity (Fig. 9).

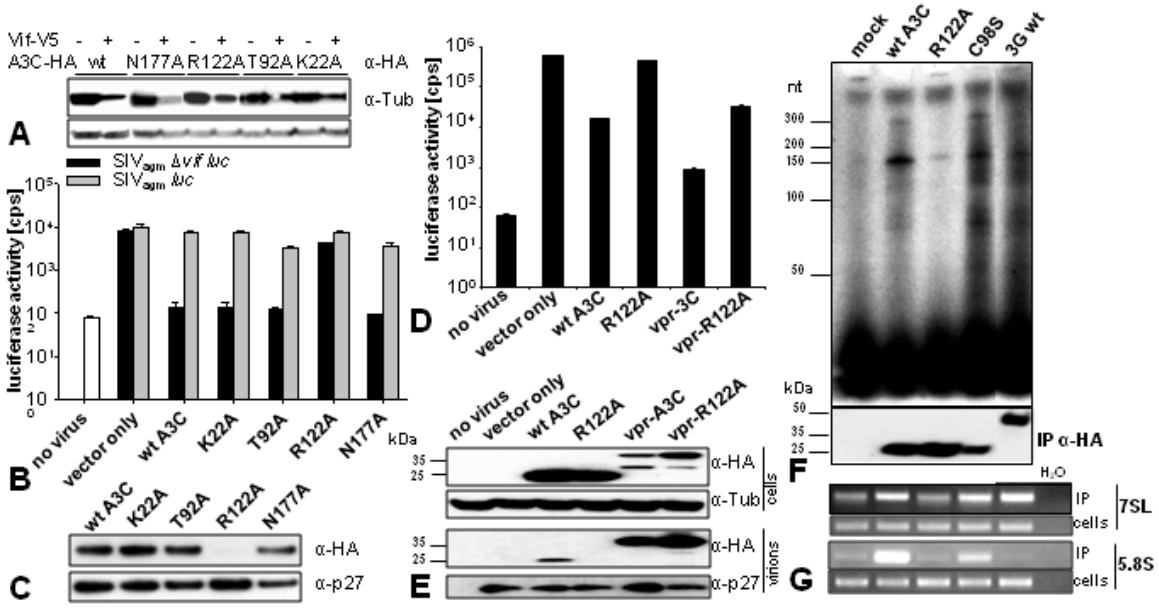


Fig. 8. (A) Immunoblot analysis of the expression and Vif-dependent degradation of wild-type(wt) A3C and the mutants K22A, T92A, R122A, N177A. The respective A3C constructs were detected by an anti(α)-HA antibody. Tubulin (Tub) served as loading control. **(B)** Antiviral activity of the mutants K22A, T92A, R122A and N177A against SIV_{agm}, compared to wt A3C and background of non-transduced cells (no virus) and vector only control (vector) without A3C. 293T cells were co-transfected with wt or Δ*vif* SIV_{agm}*luc* (VSV-G), respectively, and the respective A3C mutant. Virions were normalized by RT and HOS cells transduced. Luciferase activity was determined at 3 days post infection. **(C)** Immunoblot analysis of A3C packaging. 293T cells were co-transfected with Δ*vif* SIV_{agm}*luc*(VSV-G), the respective HA-tagged A3C construct (mutant R122A and wt). Virions were harvested and normalized by RT. Physically equal amounts of virions were lysed and subjected to immunoblot analysis. Presence of A3C in the virons was detected using α-HA-antibodies. p27 (capsid) served as loading control. **(D)** Antiviral activity of Vpr-A3C and Vpr-R122A fusion proteins against SIV_{agm}, compared to wt A3C and R122A and background of non-transduced cells (no virus) and vector only control without A3C. Δ*vif* SIV_{agm}*luc* (VSV-G) virions were generated by co-transfection with the respective A3C mutant. Virions were normalized by RT activity and used for transduction. Luciferase activity was measured 3 days post infection. **(E)** Immunoblot analysis of the expression and encapsidation of wt A3C and R122A compared to the respective Vpr-fusion-proteins. A3C constructs (+ or – Vpr) were detected by an anti(α)-HA antibody. Tubulin (Tub) served as loading control for cell lysates and p27 (capsid) for viral lysates. **(F)** RNA interacting with A3 proteins. A3C wt or mutant proteins and wt A3G were expressed in 293T and cell lysates were subjected to immuno precipitation. RNA bound to immunoprecipitated proteins was radioactively labeled with ^{32}P through RT-PCR and separated on a 12% PAA Gel and exposed on X-ray film. Background was set to signal of untransfected cells (mock). Equal amount of precipitated A3 protein was proven by immunoblot analysis of the elution fraction with an anti(α)-HA antibody. **(G)** RT-PCR on RNA interacting with A3 proteins. Isolated and A3 bound RNA (IP) was reverse transcribed and amplified using specific primers for 7 SL and 5.8 S RNA. Background signal was determined with RNA from untransfected cells (mock). Availability of the tested RNAs was confirmed for each sample through RT-PCR on RNA from cells prior to IP (cells).

We conclude that the large pocket detected on the surface of A3C plays a key role in incorporation of A3C into viral particles, mediated by RNA-dependent interaction with SIV-NC.

DISCUSSION

We have presented two three-dimensional models of A3C derived by comparative protein structure modeling taking the crystal structures of A2 and the catalytic domain of A3G as templates. These models were used to deduce hypotheses regarding dimerization and to characterize a presumed substrate binding pocket of A3C. Although sequence identity between A2 and A3C falls into the "twilight zone" (35, 36), homology between A2 and A3C can be assumed due to the conservation of the Zn^{2+}-coordinating domain, the comparable class of enzyme function, and predicted similar secondary structure. The stereochemical quality of energy-minimized models of A3C was comparable to that of the templates, and folding stability was demonstrated by MD simulations. The level of sequence identity of the templates to the targets *a priori* indicates an expected medium accuracy of the model (~85% of residues within 3.5 Å of the actual conformation), rendering them suitable to support site-directed mutagenesis experiments (37), although predictions requiring exact side chain orientations cannot be made.

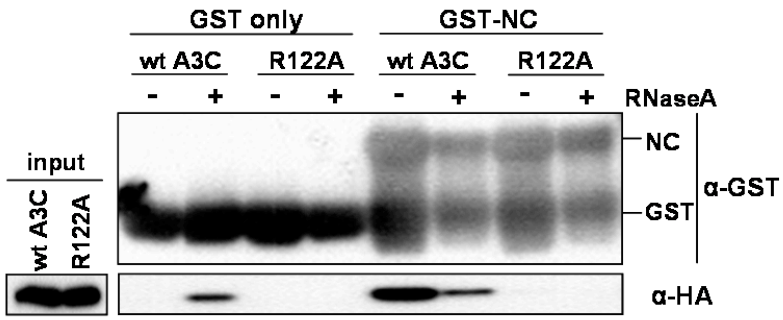


Fig. 9. Immunoblot analysis of A3C interaction with nucleocapsid (NC) from SIVagm. 293T cells were transfected with A3C wt or R122A expression plasmids and incubated with GST or GST-NC expressed in *E. coli.* After RNaseA treatment and immunoprecipitation, bound proteins were subjected to immunoblot analysis. Proteins were detected using anti-HA mAb or anti-GST mAb, respectively. Equal expression of wt A3C and R122A in 293T cells was confirmed (input).

Using an automated approach we suggest a potential substrate binding pocket in A3C, which is distal from the Zn^{2+}-coordinating site. Mutation of R122 at the pocket entrance resulted in the loss of antiviral activity due to diminished incorporation of the mutated protein into the virion. This arginine is conserved in A2 A3G-CD and A3G-ND (Fig. S2). Packaging of A3G has been demonstrated to be dependent on interaction with 7SL RNA (38). As a mutation of R122 in A3C impedes the RNA-dependent NC-interaction and the incorporation into virions one can speculate on R122 being important for RNA dependent packaging of the protein. With regard to the accuracy level of our protein model, it is possible that this binding pocket partially overlaps with the active site, resulting in a large binding pocket of bipartite function, similar to the substrate binding channel proposed for A3G-CD (21). R122 could then be thought of interacting with DNA as a substrate anchor for deaminase activity. Both activities might be mutually independent, as suggested by comparing the

characteristics of active-site mutant C98S with R122A.

Our A3C model is consistent with dimer formation of A3C analogous to that of A2. Mutations in the predicted interaction surface revealed that the antiviral function of A3C requires dimerization. In contrast to previous data (39), it was later shown that monomeric A3G is an active inhibitor of Δ*vif* HIV (40). Because of the inherent dimeric character of A3G, which possesses two copies of the Zn^{2+}-coordinating motif, additional dimerization of A3G might not be required for antiviral activity.

Summarizing, our results demonstrate that a predicted binding pocket of A3C interacts with RNA, *e.g.* 5.8 S RNA, and that RNA interaction is required for encapsidation mediated by binding to NC, but not for antiviral activity. Why dimerization of A3C is critical for antiviral activity remains an open question and an important subject for future studies.

During revision of this manuscript, Huthoff *et al.* (41) presented a homology model of A3G and demonstrated its RNA-dependent packaging that is determined by a residue inside its N-terminal domain being equivalent to R122 (for sequence alignment see Fig. S2), thereby additionally supporting our hypothesis.

EXPERIMENTAL PROCEDURES

Model building. Target A3C sequence, NCBI accession: NP_055323. Chain B of the tetrameric crystal structure of A2 (PDB identifier 2NYT, 2.5 Å resolution (22, 42)) and catalytic domain of A3G (PDB 3E1U, 2.3 Å resolution (21)) served as template. The initial sequence-alignment of A3C to monomeric A2, chain B, and the catalytic domain of A3G, was performed by the align2d function of MODELLER 9v4 (24, 43). Both target-template alignment and structural coordinates of A2, chain B, and A3G-CD, were used to build ten initial models by satisfaction of spatial restraints, subjected to energy minimization and MD simulation (Supplementary Material).

Characterization of binding pockets. PocketPicker (33) was used to automatically identify potential binding pockets and encode them as correlation vectors as described (33, 44). Each vector was then compared to 1,300 binding pockets with annotated function from the “refined set” of the PDBBind data set (34) using the Euclidean distance metric.

Mapping of functional residues. MOE 2006.08 (Chemical Computing Group, Montreal, Canada) was used to calculate the solvent accessible surface of the protein model and map electrostatic properties by a Poisson-Boltzmann potential. Putative protein-protein interaction interfaces were selected manually by looking for solvent-exposed hydrophobic patches, and fully automated using ProMate (30). Protein-protein docking and clustering of docking poses was carried out with ClusPro (31).

Plasmids. C-terminally hemagglutinin (HA)-tagged A3C expression plasmid has been described (45). Viral vectors were produced by co-transfecting $pSIV_{agm}$ Δ*vif luc* (4), pMD.G, a VSV.G expression plasmid, and supplemented by pcVif-SIV_{agm}-V5 (46). For co-immunoprecipitation studies pcDNA3.1-APOBEC3C-V5-6xHis (47) was used. HA-tagged mutated A3C constructs were derived by fusion PCR and cloned into pcDNA 3.1(+) (Invitrogen), using BamHI and NotI restriction sites. The pcVPR-A3C-HA expression plasmid was generated by fusion of SIVagm-VPR cDNA to the N-Terminus of HA-tagged wt or mutant A3C with a Gly_4-Ser-linker and cloned into pcDNA 3.1(+) using the same restriction sites. The premature stop-codon in VPR of SIVagm_TAN-1 was rectified by site directed mutagenesis. Sequences of primers in Supplementary Material. PCRs were performed with Phusion™ DNA Polymerase (Finnzymes) at: 1 cycle at 98°C, 3

min; 35 cycles at 98°C, 15 s; 65-71°C, 30 s; and 72°C, 20 s; and 1 cycle at 72°C, 10 min.

Immunoblot analysis. For analysis of expression of A3C-constructs, 293T cells were transfected with 1 µg A3C-expression plasmid and 2 µg of Vif-expression plasmid, using LipofectamineLTX (Invitrogen). Two days post transfection, cells were harvested and lysed using Western Lysis Buffer (100 mM NaCl, 20 mM Tris, pH 7.5, 10 mM EDTA, 1% sodium deoxycholate, 1% Triton X-100, and complete protease inhibitor (Roche)). Lysates were cleared by centrifugation and subjected to SDS-PAGE followed by transfer to a PVDF membrane. A3C-HA-constructs were detected using an anti (α)-HA antibody (1:10^4 dilution, Covance) and α-mouse horseradish peroxidase (1:7,500 dilution, Amersham Biosciences). For detection of Vif-V5 or A3C-V5, an α-V5 antibody (1:4,000 dilution, Serotec) was applied. α -tubulin was detected using an α-tubulin antibody (1: 10^4 dilution, Sigma). Signals were visualized by ECL plus (Amersham Biosciences).

Packaging of A3C. To detect A3C in virions, A3C expression constructs were co-transfected with $pSIV_{agm}$ Δ*vif luc* and pMD.G in 293T cells, as described above. Particles were precipitated by ultracentrifugation over a 20% sucrose cushion, normalized for activity of reverse transcriptase and lysed using Western Lysis Buffer. Lysates were directly subjected to SDS-PAGE and transferred to a PVDF-membrane. p27 was detected using a p24/p27 monoclonal antibody AG3.0 (1:250) (48).

Co-immunoprecipitation. To detect protein interaction, expression plasmids of the respective proteins were co-transfected into 293T-cells. Cells were lysed in ice cold lysis buffer (25mM Tris (pH 8.0), 137mM NaCl, 1% Glycerol, 0.1% SDS, 0.5% Na-deoxycholat, 1% NP-40, 2mM EDTA and complete protease inhibitor cocktail (Roche)). The cleared lysates were incubated with 30 µL α-HA Affinity Matrix Beads (Roche), 60 min, 4°C. The samples were washed 5 times with ice cold lysis buffer. Bound proteins were eluted by boiling the beads for 5 min at 95°C in SDS loading buffer. Immunoblot analysis and detection was done as described. Light units of the elution fractions were directly quantified from membranes incubated with ECL plus using the Lumianalyst 3.0 software (Roche).

Reporter virus assay. To measure antiretroviral activity of A3C-constructs, expression plasmids were co-transfected to 293T cells with $pSIV_{agm}$Δ*vif luc*, pMD.G, in presence and absence of a Vif SIV_{agm} expression plasmid. Two days after transfection, supernatants were harvested, normalized by RT activity and transduced 2 × 10^3 HOS cells in a 96-well dish. Three days after transduction, intracellular luciferase activity was quantified using Steady Lite HTS (Perkin Elmer). Data are presented as the average counts/second of the triplicates ± standard deviation. RT activity was quantified using the Lenti-RT Activity Assay (Cavidi Tech). The cell lines 293T and HOS were cultured in Dulbecco's high glucose modified Eagle's medium (Invitrogen), 10% fetal bovine serum, 0.29 mg/ml L-glutamine, and 100 units/m penicillin/streptomycin at 37°C with 5% CO_2.

A3C-RNA-interaction. To detect protein-RNA interaction, expression plasmids of the respective proteins were transfected to 293T-cells as described above. Cells were lysed in ice cold lysis buffer (PBS with 1% Triton X-100, 16U/mL RiboLock RNase Inhibitor (Fermentas) and complete protease inhibitor cocktail (Roche)). The cleared lysates were incubated with 50 µL α-HA Affinity Matrix Beads (Roche), 60 min, 4°C. The samples were washed 5 times with ice cold lysis buffer. RNA bound to immobilized proteins was extracted from HA-beads using TRIzol reagent (Invitrogen) (according to the companies instructions). RNAs co-precipitated with A3C proteins were dissolved in DEPC-treated H_2O and equal amounts were used for reverse transcription using random hexamer primers (Fermentas) in the presence of α-^{32}P-dATP (Hartmann Analytic). Radioactive labeled cDNA was separated on a 12% TBE 8M Urea PAA Gel and visualized with Amersham Hyperfilm (GE healthcare). Specific RT-PCR on A3C bound RNAs was performed using Reverd Aid First Strand cDNA synthesis kit (Fermentas) and random hexamer primers.

Protein-Protein-Crosslinking. For crosslinking experiments, expression plasmids of the respective proteins were transfected to 293T-cells, two days later cells were lysed and whole cell lysate was incubated with 50mM N-Ethylmaleimid (NEM/Calbiochem) for 2hrs at RT. Immunoblot analysis was performed as described above but no DTT or β-mercapthoethanol was added to the SDS-loading buffer during electrophoresis.

***E. coli* mutation assay**

Uracil DNA glycosylase-deficient *E. coli* strain BW310 was transformed with isopropyl 1-thio-β-D-galactopyranoside (IPTG) - inducible APOBEC3 expression constructs (APOBEC3 encoding cDNA was inserted in pTrc99A using XhoI and SalI restriction sites) or pTrc99A empty vector. Individual colonies were picked and grown to saturation in a rich medium containing 100µg/ml ampicillin and 1 mM IPTG. Appropriate dilutions were plated onto a rich medium containing 100µg/ml rifampicin to select Rif^R colonies after an overnight incubation. Mutation frequencies were reported as the number of Rif^R colonies per 10^9 viable cells.

A3C – NC interaction

Plasmids: A bacterial plasmid (pGEX-p2NCp1-SIVtan) expressing p3:NC:p1 fused to GST on its N-terminus was generated by PCR amplification of relevant part of SIVagmTAN and subsequent insertion into pGEX-6P via XmaI and XhoI restriction sites.

Escherichia coli (strain: BL21 DE3 Rosetta) transformed with pGEX-p2NCp1-SIVtan and parental pGEX-6p-1 were used for protein expression. Cells were harvested 5 h post induction with 100 µM IPTG (Fermentas), lysed by sonification in PBS supplemented with 1 % NP 40 and protease inhibitor coctail (Calbiochem). Glutathione sepharose beads (GE Healthcare) were incubated with clarified lysates for 30 min at 25 °C, washed 3 times with lysis buffer (10fold beads volume), resuspended in 10 mM Tris pH 7,5; 150 mM NaCl; 0,5 % NP40 with or without 34 µg/ml RNaseA (Fermentas) and incubated for 1h at 37 °C. Subsequently samples were washed and incubated with lysates of 293T cells expressing wt A3C or R122A mutant over night at 4 °C in PBS, 1% NP 40 and 300 mM NaCl. After the beads were washed extensively with PBS/1 % NP 40, proteins were eluted with 10 mM glutathione, 50 mM TRIS (pH 8) and used for immunoblot analysis using anti-HA mAb or anti-GST mAb (1:1.000 dilution; cell signaling).

ACKNOWLEDGMENTS

We thank Marion Battenberg, Elea Conrad and Norbert Dichter for expert technical assistance, and Nathaniel R. Landau and Bryan Cullen for the gift of reagents. The following reagents were obtained through the NIH AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH: pcDNA3.1-APOBEC3C-V5-6XHis from B. Matija Peterlin and Yong-Hui Zheng, monoclonal antibody to HIV-1 p24 (AG3.0) from Jonathan Allan. We are grateful to the Chemical Computing Group for providing an MOE license. This study was supported by the Beilstein-Institut zur Förderung der Chemischen Wissenschaften, and the Deutsche Forschungsgemeinschaft (SFB 579, project A11.2). Part of the study was funded by DFG grant 1608/3-1 to CM. CM is supported by the Ansmann foundation for AIDS research. We thank Dieter Häussinger for support. MW and GS are grateful to Boehringer-Ingelheim Pharma for funding.

References

1. Sheehy AM, Gaddis NC, Choi JD & Malim MH (2002) Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature* 418, 646-650.

2. Lecossier D, Bouchonnet F, Clavel F & Hance AJ (2003) Hypermutation of HIV-1 DNA in the absence of the Vif protein. *Science* 300, 1112.

3. Mangeat B *et al.* (2003) Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts. *Nature* 424, 99-103.

4. Mariani R *et al.* (2003) Species-specific exclusion of APOBEC3G from HIV-1 virions by Vif. *Cell* 114, 21-31.

5. Zhang H *et al.* (2003) The cytidine deaminase CEM15 induces hypermutation in newly synthesized HIV-1 DNA. *Nature* 424, 94-98.

6. Bishop KN *et al.* (2004) Cytidine deamination of retroviral DNA by diverse APOBEC proteins. *Curr. Biol.* 14, 1392-1396.

7. Holmes RK, Malim MH & Bishop KN (2007) APOBEC-mediated viral restriction: not simply editing? *Trends Biochem. Sci.* 32, 118-128.

8. Marin M, Rose KM, Kozak SL & Kabat D (2003) HIV-1 Vif protein binds the editing enzyme APOBEC3G and induces its degradation. *Nat. Med.* 9, 1398-1403.

9. Sheehy AM, Gaddis NC & Malim MH (2003) The antiretroviral enzyme APOBEC3G is degraded by the proteasome in response to HIV-1 Vif. *Nat. Med.* 9, 1404-1407.

10. LaRue RS *et al.* (2009) Guidelines for naming nonprimate APOBEC3 genes and proteins. *J. Virol.* 83, 494-497.

11. Münk C *et al.* (2008) Functions, structure, and read-through alternative splicing of feline APOBEC3 genes. *Genome Biol.* 9, R48.

12. Jonsson SR *et al.* (2006) Evolutionarily conserved and non-conserved retrovirus restriction activities of artiodactyl APOBEC3F proteins. *Nucleic Acids Res.* 34, 5683-5694.

13. Jarmuz A *et al.* (2002) An anthropoid-specific locus of orphan C to U RNA-editing enzymes on chromosome 22. *Genomics* 79, 285-296.

14. Chen H *et al.* (2006) APOBEC3A is a potent inhibitor of adeno-associated virus and retrotransposons. *Curr. Biol.* 16, 480-485.

15. Harari A, Ooms M, Mulder LC & Simon V (2009) Polymorphisms and splice variants influence the antiretroviral activity of human APOBEC3H. *J. Virol.* 83, 295-303.

16. OhAinle M, Kerns JA, Malik HS & Emerman M (2006) Adaptive evolution and antiviral activity of the conserved mammalian cytidine deaminase APOBEC3H. *J. Virol.* 80, 3853-3862.

17. Langlois MA, Beale RC, Conticello SG & Neuberger MS (2005) Mutational comparison of the single-domained APOBEC3C and double-domained APOBEC3F/G anti-retroviral cytidine deaminases provides insight into their DNA target site specificities. *Nucleic Acids Res.* 33, 1913-1923.

18. Yu Q *et al.* (2004) APOBEC3B and APOBEC3C are potent inhibitors of simian immunodeficiency virus replication. *J. Biol. Chem.* 279, 53379-53386.

19. Bourara K, Liegler TJ & Grant RM (2007) Target cell APOBEC3C can induce limited G-to-A mutation in HIV-1. *PLoS. Pathog.* 3, 1477-1485.

20. Chen KM *et al.* (2008) Structure of the DNA deaminase domain of the HIV-1 restriction factor APOBEC3G. *Nature* 452, 116-119.

21. Holden LG *et al.* (2008) Crystal structure of the anti-viral APOBEC3G catalytic domain and functional implications. *Nature* 456, 121-124.

22. Prochnow C *et al.* (2007) The APOBEC-2 crystal structure and functional implications for the deaminase AID. *Nature* 445, 447-451.

23. Furukawa A *et al.* (2009) Structure, interaction and real-time monitoring of the enzymatic reaction of wild-type APOBEC3G. *EMBO J.*

24. Eswar N *et al.* (2007) Comparative protein structure modeling using MODELLER. *Curr. Protoc. Protein Sci.* Chapter 2, Unit.

25. Altschul SF *et al.* (1990) Basic local alignment search tool. *J. Mol. Biol.* 215, 403-410.

26. Jones DT (1999) Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol.* 292, 195-202.

27. Melo F & Feytmans E (1998) Assessing protein structures with a non-local atomic interaction energy. *J. Mol. Biol.* 277, 1141-1152.

28. Hunenberger PH, Mark AE & van Gunsteren WF (1995) Fluctuation and cross-correlation analysis of protein motions observed in nanosecond molecular dynamics simulations. *J. Mol. Biol.* 252, 492-503.

29. Bond PJ, Faraldo-Gomez JD, Deol SS & Sansom MS (2006) Membrane protein dynamics and detergent interactions within a crystal: a simulation study of OmpA. *Proc. Natl. Acad. Sci. U. S. A* 103, 9518-9523.

30. Neuvirth H, Raz R & Schreiber G (2004) ProMate: a structure based prediction program to identify the location of protein-protein binding sites. *J. Mol. Biol.* 338, 181-199.

31. Comeau SR, Gatchell DW, Vajda S & Camacho CJ (2004) ClusPro: a fully automated algorithm for protein-protein docking. *Nucleic Acids Res.* 32, W96-W99.

32. Mitchell JBO *et al.* (1993) Amino/aromatic interactions. *Nature* 366, 413.

33. Weisel M, Proschak E & Schneider G (2007) PocketPicker: analysis of ligand binding-sites with shape descriptors. *Chem. Cent. J.* 1, 7.

34. Wang R, Fang X, Lu Y & Wang S (2004) The PDBbind database: collection of binding affinities for protein-ligand complexes with known three-dimensional structures. *J. Med. Chem.* 47, 2977-2980.

35. Abagyan RA & Batalov S (1997) Do aligned sequences share the same fold? *J. Mol. Biol.* 273, 355-368.

36. Rost B (1999) Twilight zone of protein sequence alignments. *Protein Eng* 12, 85-94.

37. Marti-Renom MA *et al.* (2000) Comparative protein structure modeling of genes and genomes. *Annu. Rev. Biophys. Biomol. Struct.* 29, 291-325.

38. Wang T *et al.* (2007) 7SL RNA mediates virion packaging of the antiviral cytidine deaminase APOBEC3G. *J. Virol.* 81, 13112-13124.

39. Shindo K *et al.* (2003) The enzymatic activity of CEM15/Apobec-3G is essential for the regulation of the infectivity of HIV-1 virion but not a sole determinant of its antiviral activity. *J. Biol. Chem.* 278, 44412-44416.

40. Opi S *et al.* (2006) Monomeric APOBEC3G is catalytically active and has antiviral activity. *J. Virol.* 80, 4673-4682.

41. Huthoff H *et al.* (2009) RNA-dependent oligomerization of APOBEC3G is required for restriction of HIV-1. *PLoS. Pathog.* 5, e1000330.

42. Berman HM *et al.* (2000) The Protein Data Bank. *Nucleic Acids Res.* 28, 235-242.

43. Sali A & Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* 234, 779-815.

44. Stahl M, Bur D & Schneider G (1999) Mapping of proteinase active sites by projection of surface-derived correlation vectors. *J. Comput. Chem.* 20, 336-347.

45. Muckenfuss H *et al.* (2006) APOBEC3 proteins inhibit human LINE-1 retrotransposition. *J. Biol. Chem.* 281, 22161-22172.

46. Perkovic M *et al.* (2008) Species-specific inhibition of APOBEC3C by the prototype foamy virus protein Bet. *J. Biol. Chem.*

47. Zheng YH *et al.* (2004) Human APOBEC3F is another host factor that blocks human immunodeficiency virus type 1 replication. *J. Virol.* 78, 6073-6076.

48. Simm M *et al.* (1995) Aberrant Gag protein composition of a human immunodeficiency virus type 1 vif mutant produced in primary lymphocytes. *J. Virol.* 69, 4582-4586.

CHAPTER II

$VPR_{(HIV-1)}$ fused to APOBEC3C alters its restriction properties but not its sub-viral localization

Henning Hofmann and Carsten Münk

Heinrich-Heine-University Düsseldorf, Clinic for Gastroenterology, Hepatology and Infectiology, Building 23.12_U1 Room 81, Moorenstr. 5, 40225 Düsseldorf, Germany

Data not published.

ABSTRACT

The APOBEC3 family of cytidine deaminases protects vertebrates from viral pathogens. Human APOBEC3C (A3C) restricts the *Simian immunodeficiency virus* (SIV) in absence of its Vif protein, but has no antiviral effect against *vif*-deficient *Human immunodeficiency virus* (HIV Δ*vif*).

Our study confirmed that A3C lacks the ability to restrict HIV Δ*vif* although it is incorporated into viral particles. This phenotype could be altered through fusing the Viral Protein R (Vpr) of HIV-1 either to the C- or N- terminus of A3C. Interestingly, also *wt* HIV was inhibited by the Vpr-A3C fusion protein indicating a block of A3C interaction with Vif. Exploring the sub-viral localization of the fusion protein compared to *wt* A3C did not show significant differences. Thus, HIV likely encodes for a factor which counteracts A3C activity additional to *vif*. Preliminary experiments indicate that the viral integrase (IN) acts as an A3C inhibitor.

INTRODUCTION

Although the APOBEC3 protein family is very homologues sharing a characteristic motif for Zn^{2+}-coordination, the specificity of each APOBEC3 differs from the others (1). Several APOBEC3 proteins (*e.g.* A3F and A3G) restrict HIV replication in the absence of the Viral infectivity factor (Vif) protein (2, 3). Another member, A3C is only active against SIV and Long Interspersed Element-1 (LINE-1) retrotransposons, but not against HIV (4, 5). The inhibitory effect of A3C against SIV Δ*vif* was shown to be 10-100 fold. However, one report described an influence of A3C on HIV in target cells (6). Here single G-A mutations could be observed in spreading HIV infection of cells expressing A3C. The authors speculate on a potential benefit of these mutations to the mutation rate of the HIV genome. The viral counterpart that mediates the HIV Δ*vif* resistance against A3C restriction was not described yet.

Another protein of the APOBEC3 family, A3A, also has no antiviral activity against HIV Δ*vif*. This Vif-independent counteraction of HIV could be circumvented by fusing the Viral Protein R (Vpr) of HIV-1 to A3A (7). Several functions in HIV replication have been suggested for Vpr such as modest transactivation of viral transcription (8),

contribution to the nuclear import of the pre-integration complex (9, 10) or induction of a cell cycle arrest in the G2/M phase in mammalian cells (11, 12). Another characteristic of Vpr is the interaction with the p6 domain of the Gag precursor polyprotein of HIV (13, 14). Via this interaction Vpr gets efficiently incorporated into HIV particles and enables the protein to deliver heterologous proteins (as Vpr fusion proteins) into the virus (15, 16). The fusion of Vpr to A3A resulted in alterations of the sub-viral localization of the fusion protein compared to *wt* A3A. The inactive A3A was not localized within the viral core whereas the antiviral active Vpr-A3A fusion protein was found to be part of the viral core thus being in close proximity to its potential target, the viral genome. Further, the fusion of Vpr of SIV of the African Green Monkey (SIVagm) to a packaging deficient R122A mutant of A3C (Vpr_{agm}.R122A) could rescue its incorporation (see Chapter I).

The aim of this study was to identify whether the fusion of HIV-1_{Vpr} to A3C modulates its restriction properties against HIV and / or alters the localization of A3C within the viral particle.

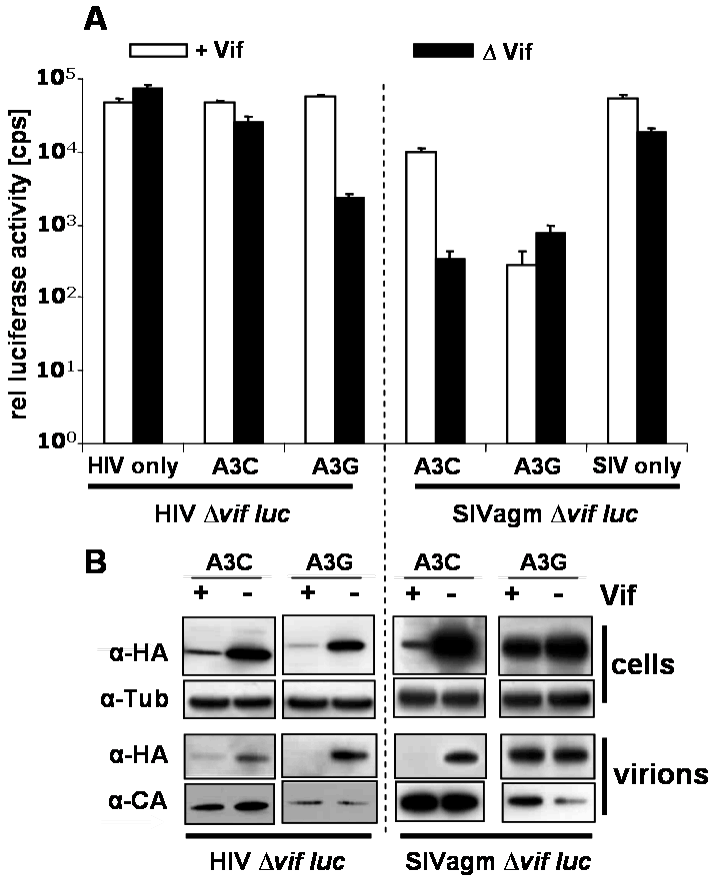


Fig. 1: A3C is active against SIVΔ*vif*, but not against HIVΔ*vif*. **(A)** Antiviral activity of A3C and A3G proteins against SIV_{agm} Δ*vif luc* and HIV Δ*vif luc* in presence (white bars) or absence (black bars) of the respective co-expressed Vif protein compared to controls without A3 expression (HIV / SIV only). **(B)** Immunoblot analysis of the expression, Vif-dependent degradation in producer cells and incorporation into virions of HA-tagged A3C and A3G using an α-HA antibody. Co-expression of the respective Vif proteins of HIV and SIV is indicated with a plus sign (+), the absence of Vif indicates a minus sign (-). Tubulin (α-Tub) and capsid (α-CA) served as loading control.

RESULTS

A3C acts antiviral only against SIV∆*vif*

APOBEC3G (A3G) is well described as a potent inhibitor of HIV ∆*vif* and SIV as well as other viruses. The antiretroviral activity of A3C is restricted to SIV ∆*vif* (5). A *vif*-deficient HIV is not inhibited by A3C. As incorporation into viral particles during budding is crucial for the antiviral activity of APOBEC3 cytidine deaminases, we first examined whether the lack of restriction against HIV ∆*vif* is due to the exclusion of A3C from HIV particles (Fig.1). Therefore virions were produced in 293T cells in presence and absence of A3C and A3G as a positive control. As expected A3G restricted SIV in a *vif*-independent manner and HIV ∆*vif* was also inhibited up to 30-fold (Fig.1A). Immunoblot analysis of lysates from transfected cells showed comparable expression levels of A3C and A3G and a Vif-dependent degradation of both proteins (Fig. 1B). As expected, A3G was not degraded by

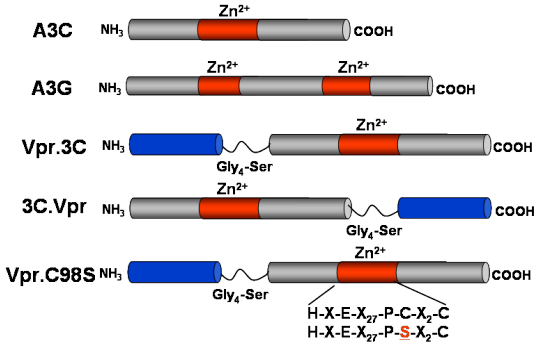


Fig. 2: Schematic representation of the used APOBEC3 and A3C-Vpr fusion proteins. APOBEC proteins are shown in grey having one (A3C) or two (A3G) Zn^{2+} - coordinating motifs shown in red. HIV-1Vpr (blue) fused to the N- or C- terminus of A3C with a Gly_4-Ser- linker in between.

the Vif protein of SIVagm (17). In both cases of viral restriction A3G was incorporated into the viral particles (Fig.1B). HIV ∆*vif* was resistant against A3C, but unexpectedly it was also incorporated into viral particles. The A3C-mediated restriction of SIV ∆*vif* demonstrated that the deaminase was enzymatically active (Fig.1A).

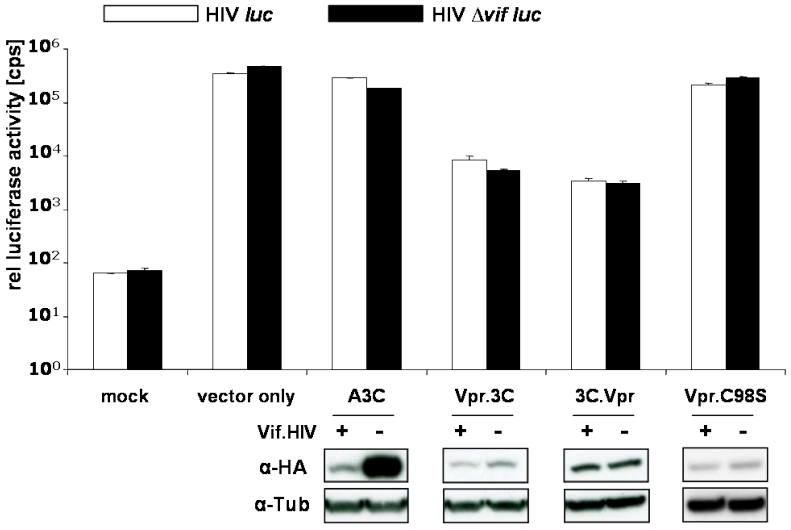


Fig. 3: Antiviral activity of *wt* A3C and indicated Vpr-A3C fusion proteins against HIV in presence (white bars) or absence (black bars) of Vif fromHIV compared to non-transduced cells (mock) and vector only control without A3C. Expression of the different HA-tagged proteins in producer cells was determined via immunoblot analysis with an α-HA antibody. Tubulin (α-Tub) served as loading control.

A $VPR_{(HIV-1)}$ - A3C fusion protein exhibits anti-HIV activity

Vpr of HIV was fused to the N- and the C-terminus of A3C (Fig. 2) via fusion PCR. In both cases a Gly_4-Ser-linker was introduced to separate the different protein domains. The same *vpr* construct was also fused to the N-terminus of the A3C C98S mutant. C98 is part of the Zn^{2+}-coordinating motif. Mutating only one of the 4 amino acids that coordinates the Zn^{2+}-ion in the active site of the protein results in loss of function proteins (see Chapter III). Here, the Vpr.C98S construct served as a negative control. To investigate the antiviral potential of the Vpr fusion proteins, HIV Δ*vif* particles were produced in 293T cells in presence or absence of the different expression plasmids. Using luciferase reporter viruses, A3C showed no antiviral activity against HIV Δ*vif*, but both A3C Vpr fusion proteins reduced the infectivity of HIV Δ*vif* ~90-fold. (Fig. 3A). Interestingly, this effect was Vif-independent because co-expression of HIV Vif could not counteract the activity of Vpr.3C. This effect was shown to be specific, as titration experiments resulted in a dose-dependent inhibition of Δ*vif* and *wt* HIV, whereas A3C was inactive against HIV Δ*vif* at any concentration of used expression plasmid (Fig. 4). Fusion of Vpr to the C-terminus of A3C resulted in an even stronger inhibition of HIV (up to ~150-fold) in presence or absence of Vif (Fig. 3A). The increased antiviral activity of 3C.Vpr correlated with higher expression compared to Vpr.3C (Fig. 3B). The antiviral activity of 3C.Vpr was likely caused by an enzymatic activity of the A3C part of the fusion protein because fusion of the enzymatic inactive C98S mutant to Vpr did not restrict HIV (Fig. 3A).

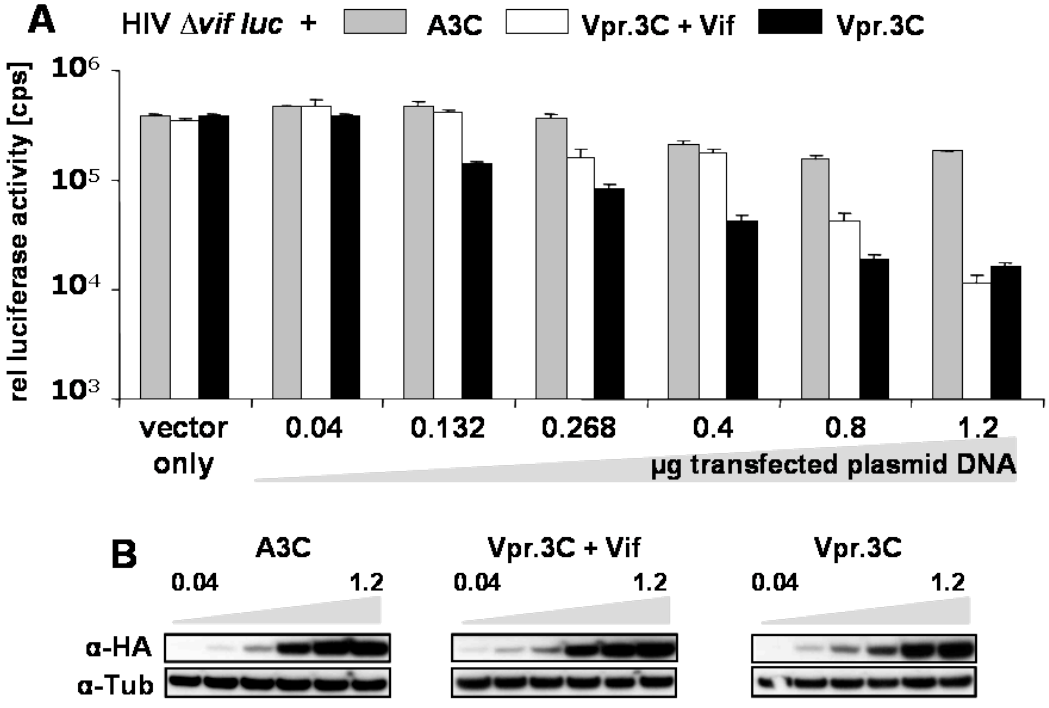


Fig. 4: (A) Dose-dependent antiviral activity of Vpr.3C against HIVΔ*vif* in presence (white bars) and absence (black bars) of co-expressed HIV Vif, compared to *wt* A3C (grey bars) and vector only control without A3C. **(B)** Immunoblot analysis of the dose-dependent expression of *wt* A3C and Vpr.3C in producer cells. The respective amounts of A3C were detected by an α-HA antibody, tubulin (Tub) served as loading control.

VPR-fusion does not change sub-viral localization

To explore the relevance of sub-viral localization of A3C and Vpr.3C, VSV-G pseudotyped *vif*-deficient HIV particles were produced in the presence of A3C or Vpr.3C. After concentration through a 20% sucrose cushion the virions were applied to a 20-60% OptiPrep™ gradient (Fig. 5).

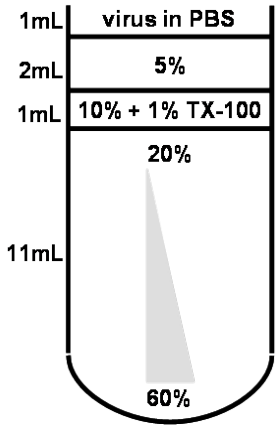


Fig. 5: Schematic representation of the constitution of the 20 - 60% OptiPrep™ gradient. Numbers (in %) indicate the OptiPrep™ concentration (diluted with H_2O) in the different compartments. TX-100 = Triton X–100.

A 1% TritonX-100 containing layer on top of the gradient was used to destroy the envelope structures resulting in a separation of the viral core pelleting at fractions with a higher density and the free proteins that can be found in fractions with lower density (18-20). After ultracentrifugation 1mL fractions were collected from bottom to top and analyzed by immunoblotting (Fig. 6B). Both, A3C and Vpr.3C localized in the fractions 5 to 8 with a density of 1.5 to 2 g/mL (Fig 6A and B). The reverse transcriptase (RT) is part of the viral core and therefore can be used to identify the core containing fractions of the gradient. Measuring the RT-activity for all fractions resulted in two peaks (Fig. 6A). The first peak of RT in both gradients was found in fractions co-localizing with A3C and Vpr.3C. RT activity was also detectable in the fractions on top of the gradient likely indicating free RT, possibly as result of a too harsh treatment with 1% TritonX-100 resulting in disruption of some viral cores. The envelope protein (here VSV-glycoprotein) was only seen in the top fractions of the gradient supporting a proper separation of free viral proteins from core structures (Fig. 6B).

As A3C and Vpr.3C show identical sub-viral localizations, modes of encapsidation cannot explain the different restriction phenotypes of both proteins. It is more likely that HIV is able to circumvent the restriction of A3C by a virus encoded factor.

HIV integrase as promising antagonist of A3C mediated restriction

To investigate whether other HIV encoded factors besides the Vif protein might be able to counteract A3C, various viral proteins were analyzed on their potential to interact with A3C (data not shown). The HIV integrase (HIV-IN) turned out to be the most promising candidate. Integrase proteins of HIV-1 and SIVmac with a C-terminal 6x Histidin tag (His-tag) were expressed in *E. coli* and bound to Ni-NTA agarose (Fig. 7, upper panel). A3C was expressed in 293T cells and supplied to the Ni-NTA bound integrases. The HIV-IN showed interaction with A3C whereas the integrase of SIV was not able to bind A3C with the same efficiency (Fig. 7, lower panel). To test the binding of IN to Vpr.3C, both proteins, A3C and Vpr.3C, were expressed in 293T cells and added to Ni-NTA bound HIV-IN expressed in *E. coli* (Fig. 8A). In support of the model, by a direct

measurement of the relative light units from the immunoblot A3C was found to interact 10-fold stronger to the HIV-IN compared to Vpr.3C (Fig. 8B).

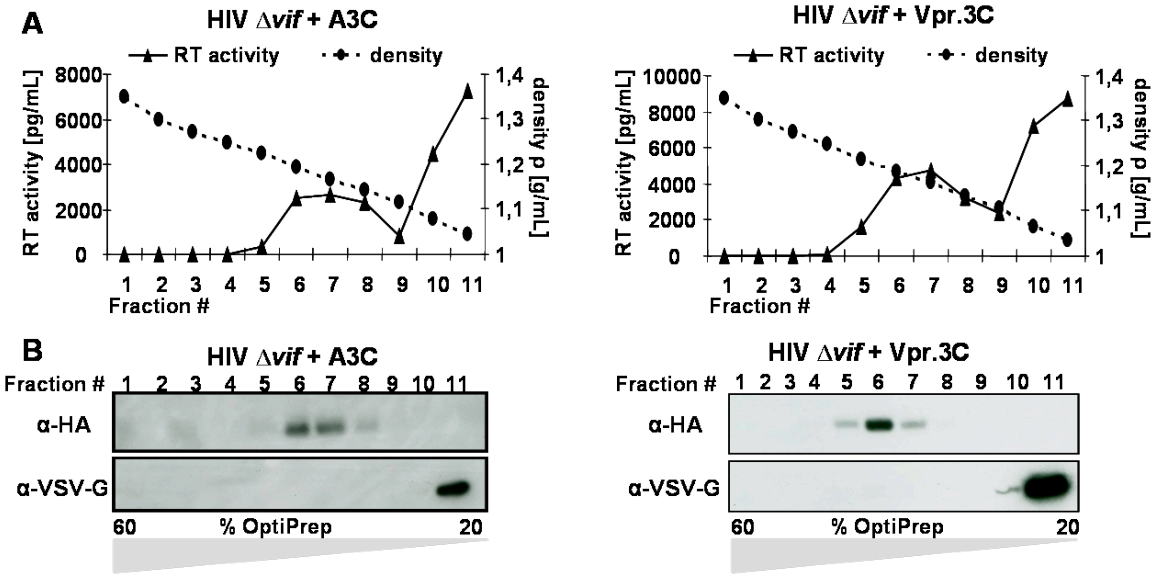


Fig. 6: OptiPrep™ gradients of HIV particles show no difference in sub-viral localization of A3C and Vpr.3C. **(A)** Determination of the RT activity in pg/mL (black lines) to display the localization the viral core containing fractions. The density (in g/mL) of each gradient fractions is displayed as dotted lines. **(B)** Immunoblot analysis to detect A3C and Vpr.3C (α-HA) in the different gradient fractions. An antibody against the VSV-glycoprotein (α-VSV-G) was used to determine the fractions that contain free viral proteins.

DISCUSSION

The antiviral activity of the ubiquitously expressed A3C (21-23)(18-20) is restricted to SIV Δ*vif* and retroelements. The Vif-independent resistance of HIV-1 to A3C implicated an unknown viral mechanism to counteract A3 proteins. Unfortunately this study could not entirely elucidate the presumed novel viral mechanism, but the results indicate a new role for IN as an antagonist against A3C.

A prerequisite for A3-mediated viral restriction is the incorporation of A3 proteins in newly budding virions. Although A3C is incorporated into HIV Δ*vif* particles no inhibition of viral infectivity could be observed. Fusion of the protein of HIV-1_{Vpr} to A3C resulted in a restriction of HIV-1. The mechanism by which Vpr.3C and 3C.Vpr inhibit the viral infectivity is so far not entirely understood. However, this restriction of HIV-1 by Vpr-A3C fusion proteins could not be counteracted through co-expression of Vif. It is thus possible that a sterical hindrance through the fused Vpr protein masked the Vif recognition site in A3C. But the Vif resistance to Vpr.3C and 3C.Vpr is difficult to explain. An inhibition of Vif binding by sterical hindrance by both, the N-terminal and the C-terminal Vpr domain is mechanistically not obvious and a valuable topic for further experiments. Vpr itself behaved inert in this system and an effect of Vpr on the cell cycle (11, 12) can be excluded, because the fusion of Vpr to the enzymatic inactive A3C mutant C98S had no influence on viral infectivity. The Vpr fusion did also not change the sub-viral localization of A3C in contrast to the results observed for A3A (7).

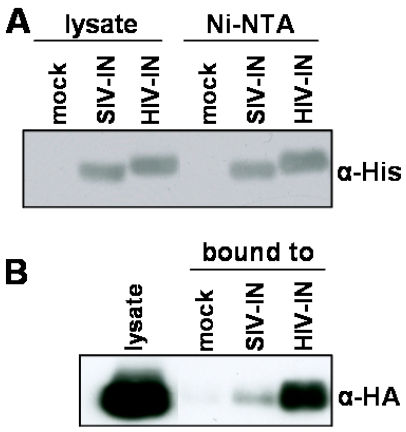


Fig. 7: A3C interacts with integrase of HIV. **(A)** Immunblot analysis of HIV and SIV integrase expression in *E. coli* and their binding to Ni-NTA agarose via a C-terminal His-tag using an α-His antibody. **(B)** Binding of HA-tagged A3C to the integrases was detected using an α-HA antibody.

To investigate the sub-viral localization of A3C and Vpr.3C a fractionation via a 20-60% OptiPrep™ gradient was performed. Unexpectedly, both A3C and the Vpr.3C fusion protein were localized in fractions with similar density. Further a peak in RT activity could be observed in these fractions, strongly suggesting presence of viral core structures. These results exclude a random packaging of A3C into newly budding virions as a result of intensive overexpression in producer cells, because A3C and Vpr.3C are not found in fractions of free viral proteins after detergent treatment such as the envelope proteins. Therefore the incorporation of A3C and Vpr.3C into HIV virions must be a specific and directed process. As differences in the sub-viral localization cannot answer the question why Vpr.3C is antiviral active and A3C is not, one might speculate on Vpr.3C dimers which have a higher affinity to each other. This would be a reasonable explanation, as the Vpr protein itself forms dimers (24) and dimerization was shown to be crucial for antiviral activity of A3C (see Chapter I). Obviously HIV must have another factor besides Vif that is counteracting the antiviral activity of A3C. This so far unknown factor could destroy A3C dimers but is not potent enough to disrupt dimers of the Vpr.3C fusion protein. This factor must be

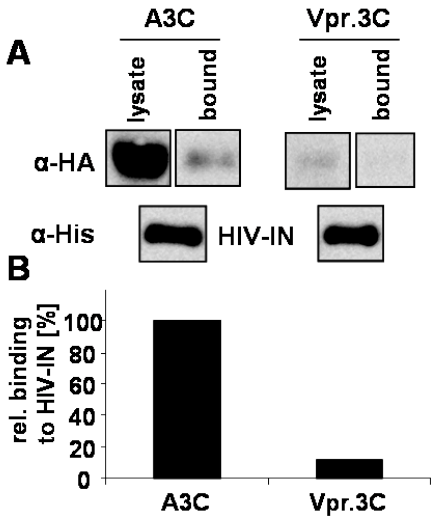


Fig. 8: HIV-IN binds A3C more efficiently than Vpr.3C. **(A)** Immunoblot analysis of A3C and Vpr.3C expression in 293T cells (lysate) and the binding of both proteins to HIV-IN (bound) using an α-HA antibody. A3C and Vpr.3C were bound to comparable amounts of HIV-IN as determined via immunoblot analysis of Ni-NTA bound integrase using an α-His antibody. **(B)** Relative binding of A3C and Vpr.3C to the HIV-IN was quantified by direct measurement of the relative light units from the immunoblots in (A). Binding of A3C to HIV-IN was set to 100%.

encoded by HIV, but not by SIV and this factor must show differences in the interaction with A3C and Vpr.3C. One of the first candidates was the Viral protein U (Vpu), as this protein is encoded by HIV, but not by SIV. Unfortunately A3C was not able to restrict the infectivity of *vpu*-deficient HIV (data not shown). After screening different other viral proteins for their physical interaction with A3C, the most promising candidate for an HIV encoded antagonist against A3C was thought to be the HIV-1

integrase. Interaction studies of A3C with the integrases of HIV and SIV showed that A3C only interacts with the HIV integrase. This emphasizes the idea of an antagonist, because HIV Δ*vif* is not restricted by A3C, in contrast to SIV Δ*vif*. Additionally, the amino acid seqeuence of the HIV and SIV integrases are 57.8% identical, which facilitates different binding properties. Furthermore the fact, that the HIV restricting fusion protein Vpr.3C shows a 10-fold lower interaction to the HIV-IN supports this hypothesis. As interaction studies can only provide an indication on a possible function, an obvious experiment was to create an HIV Δ*vif luc* reporter construct that encodes for the SIV-IN instead of its own integrase. This chimeric virus should therefore be restricted by A3C. Unfortunately, such a chimeric virus is unable to integrate into the host cell genome.

Future experiments are necessary to emphazise the idea of the HIV integrase as antagonist of A3C mediated viral restriction. A fusion of Vpr to HIV-IN could deliver the protein together with A3C into SIV Δ*vif* particles, for instance. The presence of the HIV-IN in those SIV particles should then counteract the restriction of A3C against SIV Δ*vif*. As inhibition of dimerization could be an A3C neutralizing effect of the HIV-IN, experiments targeting the dimerization of A3C and Vpr.3C in presence and absence of the integrases of HIV and SIV could be performed.

Altogether, these indications and early findings clearly lead to the integrase as potential A3C antagonist of HIV-1.

Materials and Methods

Plasmids. HA-tagged A3C and A3G constructs have been described previously (4, 25). To express the fusion proteins of Vpr_{HIV-1} and A3C the cDNA of Vpr of HIV-1 was fused either to the 5´ start (Vpr.3C) or to the 3´ end (3C.Vpr) of A3C with a Gly_4-Ser-linker in between. The fusion PCR was performed using KOD XL Polymerase (Novagen) according to the company's instructions. The final fusion product was cloned into pcDNA 3.1 (+) from Invitrogen using *BamHI* and *NotI* restriction sites. Primer sequences are listed in Suppl. table S2. Viral vectors were produced by cotransfecting $pSIV_{agm}$ Δ*vif luc* (17) or pNL 4-3 Δ*vif luc* (26), pMD.G (a VSV.G expression plasmid), and supplemented (or not) with pcVif-SIV_{agm}-V5 (27) or pc.Vif.HIV1-V5 (28).

Protein expression and virus production

293T cells were co-transfected with A3C, A3G or Vpr-fusion protein expression plasmids and $pSIV_{agm}$ Δ*vif luc* or pNL 4-3 Δ*vif luc* using Lipofectamin LTX (Invitrogen). To pseudotype the viral particles with the VSV-glycoprotein pMD.G was added. Vif expression plasmids were co-transfected to put back Vif to counteract A3 activity. 2days post-transfection virions and cells were harvested. Cells were immediately lysed in the appropiate volume of RIPA lysis buffer (25mM Tris, pH 8, 137mM NaCl, 1% glycerol, 0.1% SDS, 0.5% sodiumdeoxycholate, 1% NP-40) for 10 min on ice and lysates were clarified by centrifugation in a table top centrifuge for 10 min at 4°C with 14.000rpm. Cell lysates were subsequently used for immunoblot analysis to detect protein expression using an α-HA antibody (1:10.000 dilution, Covance) and α-mouse horseradish peroxidase (1:7.500, Amersham Biosciences). Tubulin detection served as loading control (1:10.000, Sigma). Binding was visualized by ECL plus (Amersham Biosciences). Harvested viral particles were filtered using 0.45µm MiniSart Filters (Sartorius) and downstream used for luciferase reportervirus assays or to analyze the incorporation of A3C, A3G and the fusion proteins into those particles.

Luciferase reporter virus assay. To measure the infectivity of viral particles produced in the presence and absence of co-transfected proteins, the 2 days post transfection harvested virions were normalized by RT-concentration and used to transduce 2×10^3 HOS cells in a 96-well dish. Three days after transduction, intracellular luciferase activity was quantified using Steady Lite HTS (Perkin Elmer). Data are presented as the average counts per second of the triplicates ± standard deviation. RT concentration was quantified using the Lenti-RT Activity Assay (Cavidi Tech).

Incorporation of A3 proteins into virions. To detect co-transfected proteins in viral particles, the 2days post transfection harvested, filtered and RT -normalized virions were precipitated by ultracentrifugation over a 20% sucrose cushion and lysed using RIPA lysis buffer. These lysates of whole viral particles were then analyzed by immunoblot analysis for incorporation of A3 proteins with an α-HA antibody as described above. Equal loading amounts were analyzed using the capsid monoclonal antibody AG 3.0 (29) in a 1:250 dilution. This antibody detects HIV p24 and cross-reacts with SIV p27 and was therefore used in both cases.

Determination of sub-viral localization.

Harvested and filtered virion containing supernatants were pelleted over a 20% sucrose cushion by ultracentrifugation. The particles were then resuspended in 1 mL PBS. An 11 mL density gradient consisting of 20-60% OptiPrep™ (Sigma) was produced. On top of the gradient was a 1 mL layer of 10% OptiPrep™ and 1% TritonX-100 followed by 2mL layer with 5% OptiPrep™ to build a boarder to the top 1 mL layer of resuspended viral particles. Ultracentrifugation of these gradients was performed for 20 hrs at 4°C with 80.000 rcf. Subsequently eleven 1 mL fractions from bottom to top of the gradient were collected and either used for determination of RT concentration or for immunoblot analysis. The density of each fraction was calculated through measuring the refraction index at 20°C. Specific RT activity was quantified using the Lenti-RT Activity Assay (Cavidi Tech). Localization of A3 proteins in the different fractions of the gradient was determined by immunoblot using the described α-HA antibody. To detect the envelope proteins outside the viral core an anti-VSV-G antibody (Sigma; 1:20.000) was used.

A3C-Integrase interaction. The integrases were expressed from the plasmids pKB-IN6H (HIV-1 IN) (30) and pCP-SIVIN6H (SIVmac-IN; a gift from P. Cherepanov) in in the *E. coli* strain Rosetta BL21 DE3. Empty expression vector pET-20b(+) (31) was used as negative control. Bacteria were transformed with the respective plasmid DNA and grown upon an OD of 0.5 at 600nm and heterologous protein expression was induced with 0.1mM isopropyl 1-thio-β-D-galactopyranoside (IPTG). After over night incubation at room temperature cells were harvested, resuspended in binding buffer (50mM $Na_xH_yPO_4$, 300mM NaCl, 10mM imidazole) and lysed by sonification. Ni-NTA Agarose (Invitrogen) was incubated with clarified lysates for 2 hours at 4 °C and washed 5 times with washing buffer (50mM $Na_xH_yPO_4$, 300mM NaCl, 20mM imidazole). 293T cells were transfected with A3C or Vpr.3C expression plasmids and cells were harvested 2 days post transfection, resuspended in binding buffer and lysed by sonification. The Ni-NTA agarose bound integrases were incubated with clarified lysates from 293T cells expressing A3C or Vpr.3C for 3 hours at 4°C and washed 5 times with washing buffer. After washing Ni-NTA beads were incubated with elution buffer (50mM $Na_xH_yPO_4$, 300mM NaCl, 250mM imidazole) for 10min at room temperature to elute bound proteins. Proteins in the lysates and the elution fraction were analyzed by immunoblotting using an α-His-C-terminal antibody (Invitrogen; 1:4000) (integrases) or the described α-HA antibody (A3C and Vpr.3C). Light units of the elution fractions were directly quantified from membranes incubated with ECL plus using the Lumianalyst 3.0 software (Roche).

References

1. Holmes RK, Malim MH & Bishop KN (2007) APOBEC-mediated viral restriction: not simply editing? *Trends Biochem. Sci.* 32, 118-128.

2. Sheehy AM, Gaddis NC, Choi JD & Malim MH (2002) Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature* 418, 646-650.

3. Zheng YH *et al.* (2004) Human APOBEC3F is another host factor that blocks human immunodeficiency virus type 1 replication. *J. Virol.* 78, 6073-6076.

4. Muckenfuss H *et al.* (2006) APOBEC3 proteins inhibit human LINE-1 retrotransposition. *J. Biol. Chem.* 281, 22161-22172.

5. Yu Q *et al.* (2004) APOBEC3B and APOBEC3C are potent inhibitors of simian immunodeficiency virus replication. *J. Biol. Chem.* 279, 53379-53386.

6. Bourara K, Liegler TJ & Grant RM (2007) Target cell APOBEC3C can induce limited G-to-A mutation in HIV-1. *PLoS. Pathog.* 3, 1477-1485.

7. Aguiar RS, Lovsin N, Tanuri A & Peterlin BM (2008) Vpr.A3A chimera inhibits HIV replication. *J. Biol. Chem.* 283, 2518-2525.

8. Cohen EA *et al.* (1990) Identification of HIV-1 vpr product and function. *J. Acquir. Immune. Defic. Syndr.* 3, 11-18.

9. Heinzinger NK *et al.* (1994) The Vpr protein of human immunodeficiency virus type 1 influences nuclear localization of viral nucleic acids in nondividing host cells. *Proc. Natl. Acad. Sci. U. S. A* 91, 7311-7315.

10. Subbramanian RA *et al.* (1998) Human immunodeficiency virus type 1 Vpr localization: nuclear transport of a viral protein modulated by a putative amphipathic helical structure and its relevance to biological activity. *J. Mol. Biol.* 278, 13-30.

11. Jowett JB *et al.* (1995) The human immunodeficiency virus type 1 vpr gene arrests infected T cells in the G2 + M phase of the cell cycle. *J. Virol.* 69, 6304-6313.

12. Planelles V *et al.* (1996) Vpr-induced cell cycle arrest is conserved among primate lentiviruses. *J. Virol.* 70, 2516-2524.

13. Kondo E, Mammano F, Cohen EA & Gottlinger HG (1995) The p6gag domain of human immunodeficiency virus type 1 is sufficient for the incorporation of Vpr into heterologous viral particles. *J. Virol.* 69, 2759-2764.

14. Paxton W, Connor RI & Landau NR (1993) Incorporation of Vpr into human immunodeficiency virus type 1 virions: requirement for the p6 region of gag and mutational analysis. *J. Virol.* 67, 7229-7237.

15. Fletcher TM, III *et al.* (1997) Complementation of integrase function in HIV-1 virions. *EMBO J.* 16, 5123-5138.

16. Wu X *et al.* (1997) Functional RT and IN incorporated into HIV-1 particles independently of the Gag/Pol precursor protein. *EMBO J.* 16, 5113-5122.

17. Mariani R *et al.* (2003) Species-specific exclusion of APOBEC3G from HIV-1 virions by Vif. *Cell* 114, 21-31.

18. Dettenhofer M & Yu XF (1999) Highly purified human immunodeficiency virus type 1 reveals a virtual absence of Vif in virions. *J. Virol.* 73, 1460-1467.

19. Mouland AJ *et al.* (2000) The double-stranded RNA-binding protein Staufen is incorporated in human immunodeficiency virus type 1: evidence for a role in genomic RNA encapsidation. *J. Virol.* 74, 5441-5451.

20. Wang X, Dolan PT, Dang Y & Zheng YH (2007) Biochemical differentiation of APOBEC3F and APOBEC3G proteins associated with HIV-1 life cycle. *J. Biol. Chem.* 282, 1585-1594.

21. Chiu YL & Greene WC (2008) The APOBEC3 cytidine deaminases: an innate defensive network opposing exogenous retroviruses and endogenous retroelements. *Annu. Rev. Immunol.* 26, 317-353.

22. Koning FA *et al.* (2009) Defining APOBEC3 expression patterns in human tissues and hematopoietic cell subsets. *J. Virol.* 83, 9474-9485.

23. Refsland EW *et al.* (2010) Quantitative profiling of the full APOBEC3 mRNA repertoire in lymphocytes and tissues: implications for HIV-1 restriction. *Nucleic Acids Res.* 38, 4274-4284.

24. Zhao LJ, Wang L, Mukherjee S & Narayan O (1994) Biochemical mechanism of HIV-1 Vpr function. Oligomerization mediated by the N-terminal domain. *J. Biol. Chem.* 269, 32131-32137.

25. Lochelt M *et al.* (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.

26. Loewen N *et al.* (2003) FIV Vectors. *Methods Mol. Biol.* 229, 251-271.

27. Perkovic M *et al.* (2008) Species-specific inhibition of APOBEC3C by the prototype foamy virus protein Bet. *J. Biol. Chem.*

28. Zielonka J *et al.* (2010) Vif of feline immunodeficiency virus from domestic cats protects against APOBEC3 restriction factors from many felids. *J. Virol.* 84, 7312-7324.

29. Simm M *et al.* (1995) Aberrant Gag protein composition of a human immunodeficiency virus type 1 vif mutant produced in primary lymphocytes. *J. Virol.* 69, 4582-4586.

30. Maertens G *et al.* (2003) LEDGF/p75 is essential for nuclear and chromosomal targeting of HIV-1 integrase in human cells. *J. Biol. Chem.* 278, 33528-33539.

31. Cherepanov P (2007) LEDGF/p75 interacts with divergent lentiviral integrases and modulates their enzymatic activity in vitro. *Nucleic Acids Res.* 35, 113-124.

CHAPTER III

APOBEC3C mediated restriction of SIV is editing-independent

Henning Hofmann and Carsten Münk

Heinrich-Heine-University Düsseldorf, Clinic for Gastroenterology, Hepatology and Infectiology, Building 23.12_U1 Room 81, Moorenstr. 5, 40225 Düsseldorf, Germany

Data not published.

ABSTRACT

Different retroviruses and retroelements are restricted by cellular cytidine deaminases of the APOBEC3 protein family. These proteins inhibit viral replication by deamination of cytidines to uridines on ssDNA during retroviral reverse transcription, a well characterized mechanism called editing. However, APOBEC3C's (A3C) antiviral activity against *Simian immunodeficiency virus* (SIV) lacking its *vif* gene (SIV Δ*vif*) was not caused by editing.

This study aimed to find an alternative restriction mechanism of A3C against SIV Δ*vif*. Different steps in viral replication that are potential targets of APOBEC3 mediated restriction were examined. The results demonstrate, that A3C neither induces RNA-editing nor inhibits reverse transcription. Thus, A3C likely uses an additional, so far unidentified mechanism to restrict SIV Δ*vif*.

INTRODUCTION

APOBEC3 proteins differ from each other in having one or two Zn^{2+}-coordinating (Z) domains, that belong either to the Z1, Z2, or Z3 type (1). The best described member of the APOBEC3 protein family, APOBEC3G (A3G), comprises an N-terminal Z2- and a C-terminal Z1-domain. Only the C-terminal domain mediates the enzymatic activity whereby retroviruses and transposable elements are inhibited (2-4). The restriction pattern of A3G is much broader than that of A3C (reviewed in (5)) which owns a single Z2-domain. A3G inhibits HIV Δ*vif* mainly by deamination of cytidine to uridine of ssDNA formed during reverse transcription (6-9). These mutations either lead to degradation of viral DNA by DNA base repair enzymes (uracil DNA glycosylase and apurinic-apyrimidinic endonuclease) (10), or to extensive G to A hypermutations on the coding strand of viral DNA. These editing events could result in alterations of the amino acid code or in introduction of inappropriate translation termination codons (7-9, 11). However, other mechanisms including inhibition of reverse transcription or

integration have been discussed recently. (12-16).

This study aimed to answer two questions: a) is the Zn^{2+}-coordinating domain of A3C responsible for its antiviral activity against SIV Δ*vif* and b) what is the restriction mechanism of A3C.

RESULTS

The Zn^{2+}-coordinating motif is the active site of A3C

The two Zn^{2+}-coordinating domains of APOBEC3G (A3G) display different functional properties. The N-terminal Z2-domain (NTD) is responsible for virion incorporation of the protein and the C-terminal Z1-domain (CTD) was shown to be the enzymatic active site of A3G (2, 4). To determine the function of the single Z2-domain of A3C the zinc ion-binding amino acids were mutated as shown in Fig. 1.

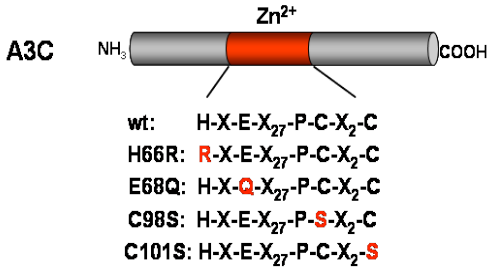


Fig. 1: Schematic overview of the point mutations in the Zn^{2+}-coordinating domain (red) of A3C (grey). The single amino acids changes are shown as red letters.

These A3C mutants were tested against SIV Δ*vif* using luciferase reporter virus. VSV-G pseudotyped SIVagm *luc* Δ*vif* virions were produced in 293T cells in presence of *wt* or mutant A3C and in presence or absence of the Vif protein of SIV.

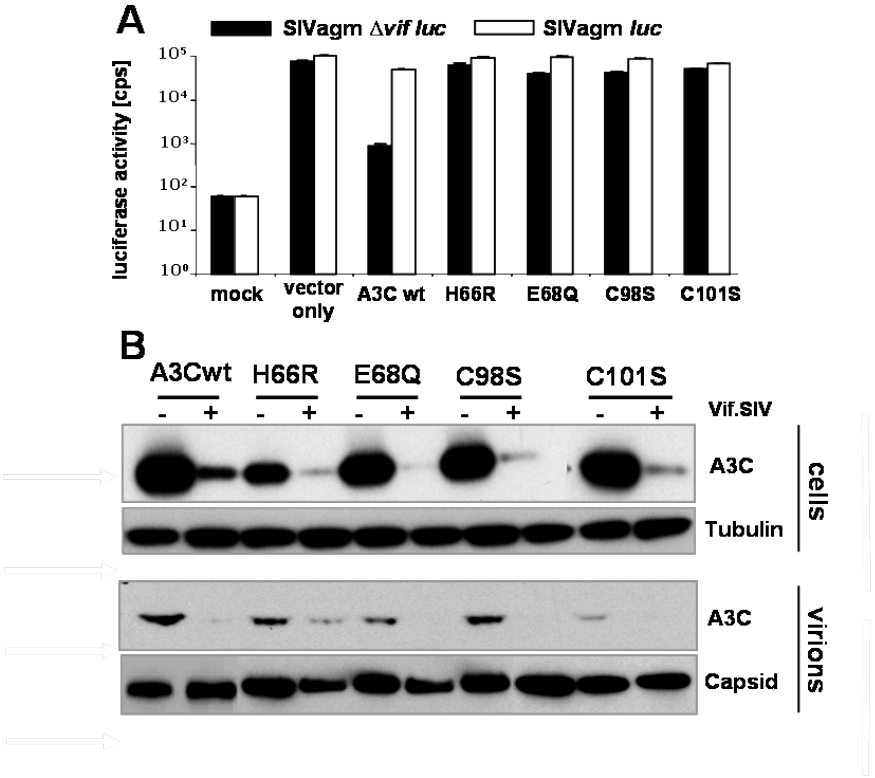


Fig. 2: Mutations in the Zn^{2+} - coordinating domain results in loss of function proteins. **(A)** Antiviral activity of A3C *wt* and mutant proteins against SIV_{agm} Δ*vif luc* and SIV_{agm} *luc*, compared to non-transduced cells (mock) and vector only (control without A3C). **(B)** Immunoblot analysis of the expression, Vif-dependent degradation in producer cells and incorporation into virions of A3C *wt* and mutant proteins. Tubulin and p27 (capsid) served as loading control.

Normalized viral particles were used to transduce human osteosarcoma (HOS) cells. Intracellular luciferase activity in those cells determined the infectivity of used virions. Only *wt* A3C could reduce viral infectivity displayed as a decrease in luciferase activity (Fig. 2A). Mutation of any Zn^{2+}-coordinating amino acid resulted in loss-of-function proteins, likely by destroying the active site. As these mutants were recognized by the Vif protein of SIV inducing their proteasomal degradation (Fig. 2B) they should not be subject to severe misfolding. Additionally, virion incorporation of A3C *wt* and mutant proteins was examined as this is a prerequisite for antiviral activity (Fig. 2B). The mutants H66R, E68Q, C98S were incorporated into viral particles similar to *wt* A3C, indicating that the non-restricting phenotype is not based on the exclusion of the mutants from the virions. Although the mutated proteins were packaged, mutations impaired their enzymatic active domain. Mutant C101S was packaged to a significantly lower extent and for this protein a low encapsidation into virions and/or the damaged active site can explain its inactivity against SIV Δ*vif*.

Viral DNA is not deaminated by A3C

As deamination of viral cDNA is the most prominent mechanism of A3-mediated retroviral restriction, we examined whether DNA editing is also the reason for SIV Δ*vif* inhibition by A3C. Therefore pseudotyped SIVΔ *vif luc* virions, generated in presence or absence of A3C, were used to transduce HOS cells. At indicated time points post infection total DNA was isolated from those cells and a 600bp long fragment of the SIV *gag* gene was amplified via PCR. The resulting products were then sequenced and analyzed for G to A mutations (Fig. 3). The results display that A3C does not deaminate the viral cDNA of SIV Δ*vif*. DNA isolated at 4, 9, 12 or 24 hours post infection in the presence of A3C showed G to A mutation ratios similar to that of the negative control (SIV Δ*vif* only) (Fig. 3). As a positive control, HOS cells were transduced with virions produced in the presence of A3G. PCR products of this infection showed the expected extensive hypermutation of the viral genome caused by A3G.

A3C does not edit viral RNA

As deamination of viral cDNA was likely not involved in the antiviral activity of A3C, another possible mechanism was editing of viral RNA prior to reverse transcription. Total RNA was isolated from cells transduced with VSV-G pseudotyped SIVΔ*vif luc* virions produced in the presence or absence of A3C. After cDNA synthesis from total RNA a region including the primer binding site (PBS) was amplified and analyzed for mutations. Primers amplifying the PBS were chosen, as the initial step of reverse transcription starts with binding of the tRNA to the PBS (Fig. 4A) (17). The sequences (n=7) of SIV made in the presence of A3C showed that more non-specific (non C to U) mutations were detectable than mutations that could have resulted from A3C's enzymatic activity (Fig. 4B). Analyzing sequences after infection with SIV made in the absence of A3C ended up with a similar result. Taken together, A3C does neither deaminate viral cDNA nor viral RNA.

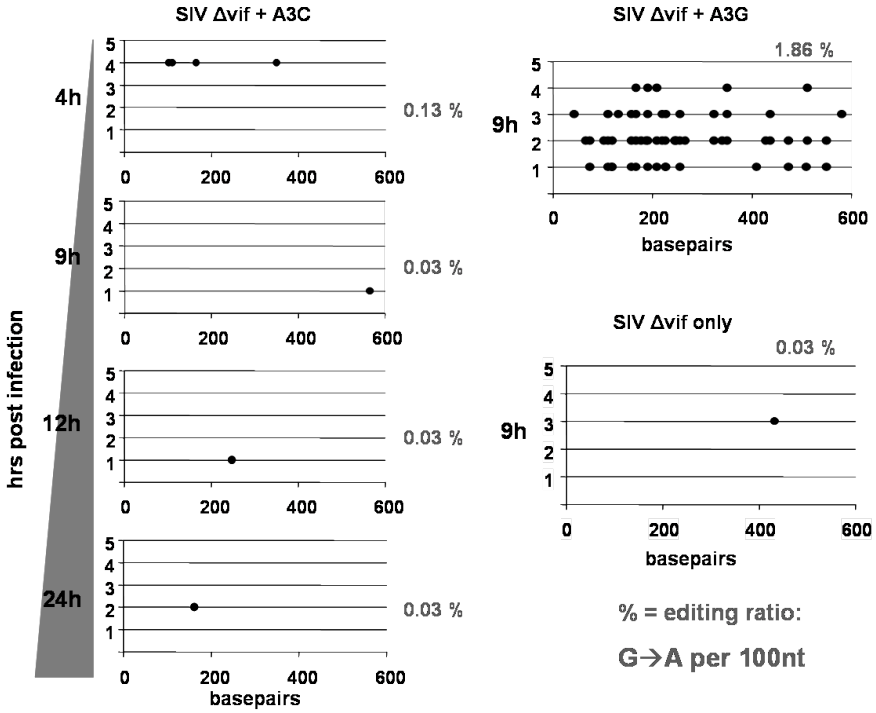
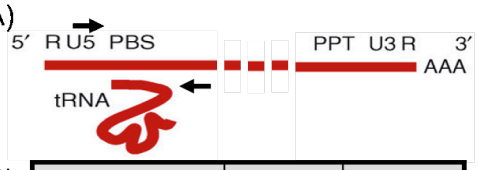


Fig. 3: A3C does not deaminate the genome of SIV Δ*vif*. Each G to A mutation is depicted as a dot. Five different clones were analyzed after different time points post infection. Editing ratios were calculated as the percentage of G to A mutations per 100 nt (grey). A3G served as positive control, SIV Δ*vif* only displays the background of A3-independend mutations.

A3C does not block reverse transcription

The products of retroviral reverse transcription (RT) can be divided in early and late RT products according to their appearance during the retroviral life cycle (18).

A)



B)

Mutation	+ A3C n = 7	w/o A3C n = 4
C→U / 100bp	0.09	0.038
non C→U / 100bp	0.195	0.16

Fig. 4: No deamination of SIV RNA detectable in the presence of A3C. **(A)** Primers are shown as black arrows flanking the PBS in the SIV genome. **(B)** Mutation ratios (in %) in presence (+) and absence (w/o) of A3C were calculated as C to U mutation on viral RNA per 100 nt.

To investigate, if the RT reaction is influenced by A3C at all, we first analyzed late RT products. HOS cells were transduced with normalized amounts of virions (SIVΔ*vif luc*) produced with or without A3C and 2, 4, 24, 48 and 96 hours post infection total DNA was isolated from these cells. Via quantitative real time PCR the amount of viral, reverse transcribed cDNA was quantified as fg per 100ng of total DNA using primers specific for late RT products. The results shown in Fig. 5 indicate a delay in the appearance of late RT products of SIV*Δvif* made in the presence of A3C, whereas the overall amount of RT products remained the same. To proof that the observed A3C dependent decrease of luciferase activity in SIVΔ*vif luc* infected cells after 3 days (Fig. 2) does not result from delayed viral replication, intracellular luciferase activity was also measured at day

4 post infection (Fig. 6). However, luciferase activity of virions produced in presence of A3C analyzed on day 3 and day 4 post infection were comparable. Thus, the 10 – 100-fold antiviral effect of A3C likely cannot be explained by inhibition of reverse transcription.

DISCUSSION

The mechanisms responsible for the retroviral restriction by APOBEC3 proteins are still not completely solved. Besides deamination of the viral genome, other possibilities of A3-mediated restriction have been discussed recently (12-16), such as inhibition of reverse transcription or integration. In two domain APOBEC3 proteins only one domain exhibits enzymatic activity, whereas the other domain is necessary for virion incorporation (2-4). Since A3C comprises only a single Zn^{2+}-coordinating domain, it was unclear how the protein manages both, encapsidation and antiviral activity.

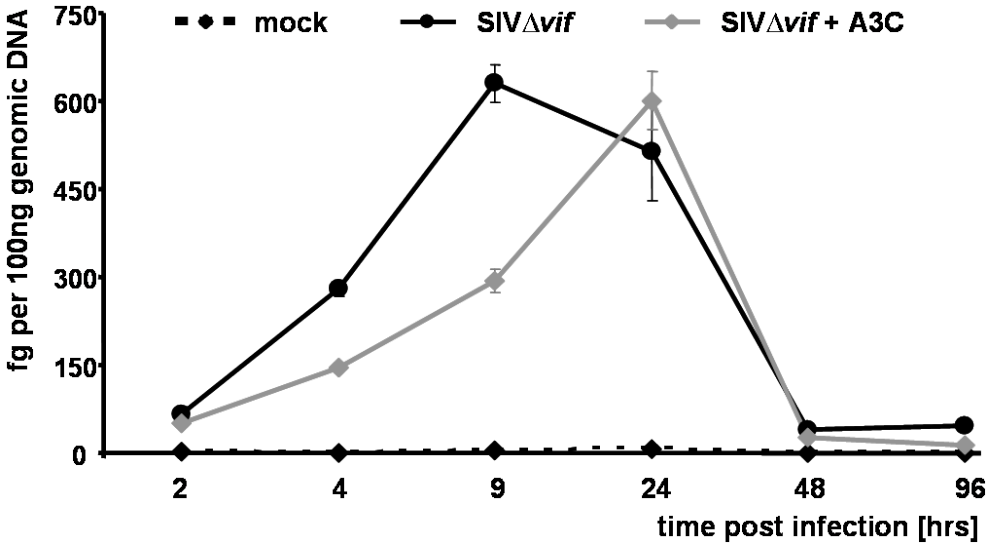


Fig. 5: Late reverse transcription products appear slightly later in the presence of A3C. Quantitative real-time PCR analysis of SIV_{agm} Δ*vif luc* late RT products after indicated time points post infection in target cells. Amount of reverse transcribed products was quantified as fg per 100ng of total isolated DNA.

To answer whether the Zn^{2+}-coordinating domain is the enzymatic active site of A3C, the crucial amino acids within this domain were mutated. These mutations resulted in a complete loss of antiviral activity although most of the mutants were efficiently incorporated, proofing that the single Zn^{2+}-coordinating domain is responsible for enzymatic activity. Later it was shown that a binding pocket proximal to this active site is crucial for RNA-dependent incorporation of A3C into viral particles (see Chapter I). Although in a bacterial deamination assay it was shown that A3C has the intrinsic capacity to deaminate cytidines (see Chapter I), there was no editing detectable in the context of SIV Δ*vif* infections in human cells. This result is in contrast to published data (19), where a very low editing ratio of A3C has been described. One could speculate about A3C inducing mutations resulting in viral restriction that might have occurred elsewhere in the viral genome instead of the analyzed part of the *gag* gene. This

possibility is very unlikely, as A3 proteins have certain recognition sequences (20) that consist of 2-4 nt being found within the whole viral genome. Also A3G clearly showed a strong deamination pattern within this part of the SIV genome. As the experimental results argue against cytidine deamination being the main mechanism of A3C-mediated restriction of SIV Δ*vif*, other possible molecular consequences of A3C activity were analyzed.

To investigate inhibition of reverse transcription, two strategies were pursued. The first was RNA editing at the PBS. The RT reaction is initiated through binding of a tRNA at the PBS. The tRNA binding could be inhibited by destroying the recognition sequence of the PBS and reverse transcription would be blocked at the very first step. In agreement that A3 proteins act on ssDNA (20), RNA editing of A3C could not

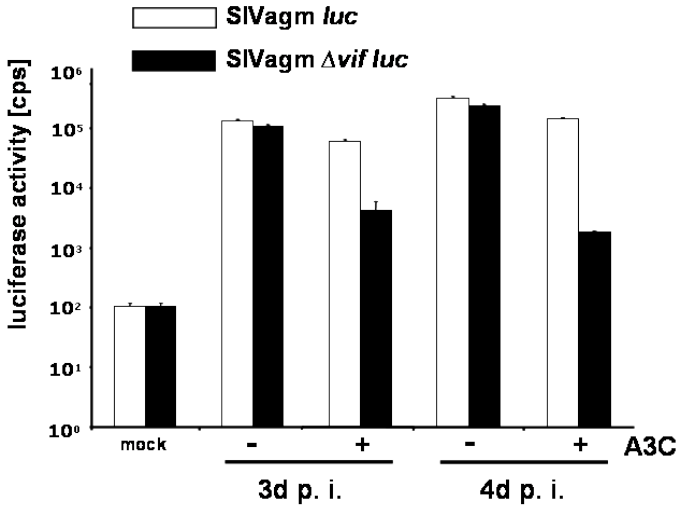


Fig. 6: Measurement of the antiviral activity of A3C against Δ*vi* SIV_{agm} *luc* at 3 or 4 days post infection (p. i.) in presence (white bars) or absence (black bars) of the Vif protein SIV_{agm} showed no significant differences.

be proven. The second approach was to examine whether the RT reaction was influenced through the presence of A3C. For A3G a decrease in HIV-1 reverse transcription products was determined when the protein was present (14, 21). It remains unclear whether this inhibition is caused by direct physical interaction with the viral enzyme reverse transcriptase. A3G could also bind to viral cDNA after first strand synthesis and thereby sterically block the synthesis of the second cDNA strand after strand transfer (15). To investigate whether A3C inhibits the RT reaction during infection, a quantitative PCR approach for late RT products was chosen amplifying a 600nt long part of the *gag* gene that is reverse transcribed at late stages of the dsDNA synthesis. Quantifying these late occurring RT products would show whether reverse transcription is inhibited at all. The results displayed a delay in the appearance of those late RT products caused by A3C. But as the overall amount of viral cDNA produced in presence or absence of A3C was comparable, it can be concluded that A3C does not inhibit reverse transcription. In the presence of A3C the RT products are moderately delayed. But this delay in the RT reaction did not correlate with a “restored” infectivity at later time points post infections. More experiments are required to test,

whether the next steps in replication, nuclear import and integration are affected by A3C.

Materials and Methods

Plasmids. HA-tagged A3C and A3G constructs have been described previously (22, 23). The point mutations of A3C were inserted using site directed mutagenesis. Therefore A3C expression plasmid was used as template for a PCR reaction amplifying the whole plasmid with primers carrying those nucleotides to be changed. Primer sequences: H66R-5': 5'-tctgcacgacaatgggtctca-3', H66R-3': 5'-gagacccgttgtcatgcaga-3', E68Q-5': 5'-caccttttgtgcatgacaatg-3', E68Q-3': 5'-atgcacaaag gtgcttcctc-3', C98S-5': 5'-tctggagaagggctccaagat-3', C98S-3': 5'-gcccttctccagactgtgca-3', C101S-5': 5'-ctgcagagtctgggcaaggg-3' and C101S-3': 5'-cagactctgcaggggaggtg-3'. After PCR the products were digested with DpnI to disrupt methylated, parental DNA and transformed into *E. coli* strain Top 10 (Invitrogen). Mutations were verified by sequencing. Viral vectors were produced by co-transfection of $pSIV_{agm}$ Δ*vif luc* (24) or pNL 4-3 Δ*vif luc* (25), pMD.G, a VSV-G expression plasmid, and supplemented (or not) with pcVif-SIV_{agm}-V5 (26) or pc.vif.HIV1-V5 (27).

Protein expression and virus production. 293T cells were co-transfected with A3C expression plasmids and $pSIV_{agm}$Δ*vif luc* or pNL 4-3Δ*vif luc* using Lipofectamine LTX (Invitrogen). To pseudotype the viral particles with the VSV-glycoprotein pMD.G was co-transfected, too. Vif expression plasmids were co-transfected to counteract A3 activity. 2 days post transfection virions and cells were harvested. Cells were immediately lysed in an appropriate volume of RIPA lysis buffer (25mM Tris, pH 8, 137mM NaCl, 1% glycerol, 0.1% SDS, 0.5% sodiumdeoxycholate, 1% NP-40) for 10min on ice and lysates were clarified by centrifugation in a table top centrifuge for 10min at 4°C with 14.000rpm. Cell lysates were then used for immunoblot analysis to detect protein expression using an anti (α)-HA antibody (1:10.000, Covance) and α-mouse horseradish peroxidase (1:7.500, Amersham Biosciences). α -Tubulin was used as loading control and detected using α-tubulin antibody (1:10.000, Sigma). Signals were visualized by ECL plus (Amersham Biosciences). Harvested viral particles were filtered using 0.45µm MiniSart Filters (Sartorius) and downstream used for luciferase reporter virus infections or to analyze the incorporation of A3C *wt* or mutant proteins into those particles.

Luciferase reporter virus assay. To measure the infectivity of viral particles produced in the presence and absence of co-transfected proteins, 2 days post transfection virions were harvested, normalized by RT-concentration and used to transduce 2×10^3 HOS cells in a 96-well dish. Three days after infection, intracellular luciferase activity was quantified using Steady Lite HTS (Perkin Elmer). Data are presented as the average counts per second of the triplicates ± standard deviation. RT concentration was quantified using the Lenti-RT Activity Assay (Cavidi Tech).

Incorporation of A3 proteins into virions. To detect co-transfected proteins in viral particles, the 2 days post transfection harvested, filtered and RT -normalized virions were precipitated by ultracentrifugation over a 20% sucrose cushion and lysed using RIPA lysis buffer. These lysates of whole viral particles were then analyzed by immunoblot analysis for incorporation of A3 proteins with an α-HA antibody as described above. Equal loading amounts were analyzed using the capsid monoclonal antibody AG 3.0 (28) (1:250). This antibody detects HIV p24 and cross-reacts with SIV p27 and was therefore used in both cases.

DNA Deamination assay. To detect deamination of viral cDNA, 2 days post transfection virions were harvested, filtered and RT normalized, as well as incubated with DNaseI (Roche) to remove residual plasmid DNA. HOS cells were transduced with these virions and 4, 9, 12 and 24 hours post infection total DNA was isolated using DNeasy Blood & Tissue Kit (QIAGEN). 300ng of total DNA was used as template to amplify a part of the *gag* gene with the primers CM101 (5'-caggctgagaaatctccagcag-3') and CM102 (5'-ccatgtctgccactaggtgtcgc-3') using *Taq* polymerase

(Fermentas). The PCR products were cloned into pJET1.2 cloning vector (Fermentas) and sequenced. G to A mutations were detected by aligning these sequences with the SIV_{agm}TAN-1 *gag* sequence. Editing ratios were calculated as G to A mutations per 100nt analyzed.

RNA Deamination assay. Infection of HOS cells was performed similar to the DNA deamination assay. 6 hours post infection total RNA was isolated using RNeasy Kit (QIAGEN) and reverse transcribed into total cDNA with Super Script III Reverse Transcriptase (Invitrogen). 500ng of cDNA were used for the PCR reaction with primers PBS5´ (5'-cttaagagtctatctgagcaag-3') and PBS3´ (5'-gtaattaccatgtctgccact-3') flanking the primer binding site. PCR products were analyzed similar to the DNA deamination assay.

Quantitative PCR to detect late RT products. HOS cells were transduced with SIV *Δvif* virions produced in presence and absence of A3C. At 2, 4, 9, 24, 48 and 96 hours post infection total DNA was isolated using DNeasy Blood & Tissue Kit (QIAGEN). 100ng of each total DNA were used for a quantitative real-time PCR to detect late RT products with the primers: CS: 5'-cactcgg cactgtcaggga-3' and CR: 5'-ggttctagcgggctcaata cttctat-3'. Specificity of the products was determined by use of Fluorescein (Fl) and Light Cycler Red 640 (LC) labeled hybridization probes: C-Fl: 5'-tgatacttttttctttccgttcgggcg—Fl-3' and 5'-LC—aagcgtattttctcaaatgtgtccaaattcc-3'. The following cycling conditions were used: 50 cycles with 95°C for 10sec, 57°C for 20sec and 72°C for 10sec in Light Cycler 3 instrument from Roche. The amount of late RT products in fg per 100ng was determined by standardization of each run with a serial dilution of $pSIV_{agm}$ Δ*vif* plasmid. Therefore 10fg, 100fg, 1pg and 10pg of plasmid DNA were used.

References

1. LaRue RS *et al.* (2008) The artiodactyl APOBEC3 innate immune repertoire shows evidence for a multi-functional domain organization that existed in the ancestor of placental mammals. *BMC. Mol. Biol.* 9, 104.

2. Gooch BD & Cullen BR (2008) Functional domain organization of human APOBEC3G. *Virology* 379, 118-124.

3. Hache G, Liddament MT & Harris RS (2005) The retroviral hypermutation specificity of APOBEC3F and APOBEC3G is governed by the C-terminal DNA cytosine deaminase domain. *J. Biol. Chem.* 280, 10920-10924.

4. Navarro F *et al.* (2005) Complementary function of the two catalytic domains of APOBEC3G. *Virology* 333, 374-386.

5. Holmes RK, Malim MH & Bishop KN (2007) APOBEC-mediated viral restriction: not simply editing? *Trends Biochem. Sci.* 32, 118-128.

6. Bishop KN *et al.* (2004) Cytidine deamination of retroviral DNA by diverse APOBEC proteins. *Curr. Biol.* 14, 1392-1396.

7. Lecossier D, Bouchonnet F, Clavel F & Hance AJ (2003) Hypermutation of HIV-1 DNA in the absence of the Vif protein. *Science* 300, 1112.

8. Mangeat B *et al.* (2003) Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts. *Nature* 424, 99-103.

9. Zhang H *et al.* (2003) The cytidine deaminase CEM15 induces hypermutation in newly synthesized HIV-1 DNA. *Nature* 424, 94-98.

10. Yang B *et al.* (2007) Virion-associated uracil DNA glycosylase-2 and apurinic/apyrimidinic endonuclease are involved in the degradation of APOBEC3G-edited nascent HIV-1 DNA. *J. Biol. Chem.* 282, 11667-11675.

11. Harris RS *et al.* (2003) DNA deamination mediates innate immunity to retroviral infection. *Cell* 113, 803-809.

12. Bishop KN, Holmes RK & Malim MH (2006) Antiviral potency of APOBEC proteins does not correlate with cytidine deamination. *J. Virol.* 80, 8450-8458.

13. Guo F *et al.* (2006) Inhibition of formula-primed reverse transcription by human APOBEC3G during human immunodeficiency virus type 1 replication. *J. Virol.* 80, 11710-11722.

14. Holmes RK, Koning FA, Bishop KN & Malim MH (2007) APOBEC3F can inhibit the accumulation of HIV-1 reverse transcription products in the absence of hypermutation. Comparisons with APOBEC3G. *J. Biol. Chem.* 282, 2587-2595.

15. Mbisa JL *et al.* (2007) Human immunodeficiency virus type 1 cDNAs produced in the presence of APOBEC3G exhibit defects in plus-strand DNA transfer and integration. *J. Virol.* 81, 7099-7110.

16. Mbisa JL, Bu W & Pathak VK (2010) APOBEC3F and APOBEC3G inhibit HIV-1 DNA integration by different mechanisms. *J. Virol.* 84, 5250-5259.

17. Harrich D & Hooker B (2002) Mechanistic aspects of HIV-1 reverse transcription initiation. *Rev. Med. Virol.* 12, 31-45.

18. Freed EO (2001) HIV-1 replication. *Somat. Cell Mol. Genet.* 26, 13-33.

19. Yu Q *et al.* (2004) APOBEC3B and APOBEC3C are potent inhibitors of simian immunodeficiency virus replication. *J. Biol. Chem.* 279, 53379-53386.

20. Yu Q *et al.* (2004) Single-strand specificity of APOBEC3G accounts for minus-strand deamination of the HIV genome. *Nat. Struct. Mol. Biol.* 11, 435-442.

21. Bishop KN *et al.* (2008) APOBEC3G inhibits elongation of HIV-1 reverse transcripts. *PLoS. Pathog.* 4, e1000231.

22. Lochelt M *et al.* (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.

23. Muckenfuss H *et al.* (2006) APOBEC3 proteins inhibit human LINE-1 retrotransposition. *J. Biol. Chem.* 281, 22161-22172.

24. Mariani R *et al.* (2003) Species-specific exclusion of APOBEC3G from HIV-1 virions by Vif. *Cell* 114, 21-31.

25. Loewen N *et al.* (2003) FIV Vectors. *Methods Mol. Biol.* 229, 251-271.

26. Perkovic M *et al.* (2008) Species-specific inhibition of APOBEC3C by the prototype foamy virus protein Bet. *J. Biol. Chem.*

27. Zielonka J *et al.* (2010) Vif of feline immunodeficiency virus from domestic cats protects against APOBEC3 restriction factors from many felids. *J. Virol.* 84, 7312-7324.

28. Simm M *et al.* (1995) Aberrant Gag protein composition of a human immunodeficiency virus type 1 vif mutant produced in primary lymphocytes. *J. Virol.* 69, 4582-4586.

Supplementary Material
-
CHAPTER I

Model Structure of APOBEC3C Reveals a Binding Pocket Modulating RNA Interaction Required for Encapsidation

Benjamin Stauch, Henning Hofmann, Mario Perković, Martin Weisel, Ferdinand Kopietz, Klaus Cichutek, Carsten Münk, Gisbert Schneider

Supplementary Table S1. PCR primer sequences.

outer primers	
BSBPK-hu3C-*f or*	5'-TAAGCGGGATCCTATCTAAGAGGCTGAACATG-3'
hA3C-mut-3'-*out*	5'-TCGAGCGGCCGCCACTGTGCT-3'
Pocket mutants	
BSBPK-**K22A**	5'-CTTCCAATTTGCAAACCTATGGG-3'
BSBPK-**K22A**-*rc*	5'- CCCATAGGTTTGCAAATTGGAAG-3'
BSBPK-**T92A**	5'-CACCTGGTACGCATCTTGGAGCC-3'
BSBPK-**T92A**-*rc*	5'-GGCTCCAAGATGCGTACCAGGTG-3'
BSBPK-**R122A**	5'-CTTCACCGCCGCCCTCTACTACT-3'
BSBPK-**R122A**-*rc*	5'-CTTCACCGCCGCCCTCTACTACT-3'
BSBPK-**N177A**	5'-ATTAAAAACCGCCTTTCGACTTC-3'
BSBPK-**N177A**-*rc*	5'-GAAGTCGAAAGGCGGTTTTTAAT-3'
Dimerization mutants	
BSBPK-**K51A**	5'-TGTCTCCTGGGCGACGGGCGTCT-3'
BSBPK-**K51A**-*rc*	5'-AGACGCCCGTCGCCCAGGAGACA-3'
BSBPK-**F55A**	5'-GACGGGCGTCGCCCGAAACCAGG-3'
BSBPK-**F55A**-*rc*	5'-GACGGGCGTCGCCCGAAACCAGG-3'
BSBPK-**W74A**	5'-CTTCCTCTTCGTTCTGCGACG-3'
BSBPK-**W74A**-*rc*	5'-CGTCGCAGAACGCAGAGAGGAAG-3'
RT-PCR	
7SL- forward	5´- ATCGGGTGTCCGCACTAAG-3´
7SL-reverse	5´- CACCCCTCCTTAGGCAACCT-3´
5.8 S-forward	5´- CGACTCTTAGCGGTGGATCAC-3´
5.8 S-reverse	5´- AAGCGACGCTCAGACAGGCGTA-3´
Cloning of pGEX-p2NCp1-SIVtan	
fXmaI-p2:NC:p1	5´- GGCCCGGGTGTAGAAGCAATGCAACAAATGC -3´
rXhoI-p2:NC:p1:	5´- GGCTCGAGTCATAAAAAATTGCGCGGCTTTG -3´

Additional Material and Methods

Model refinement and evaluation.

MODELLER 9v4 implements the dynamic programming algorithm by a variable gap penalty (1): Based on the PDB template structures, gaps were penalized differentially and favored in solvent-exposed loops and surface regions compared to secondary structure elements and core regions and therefore placed in a preferential structural context. Penalties were given as a vector with values previously suggested as optimal for the sequence identity found in the alignment (1).

REDUCE (2) was applied to add hydrogens to all initial models obtained by MODELLER 9v4 (24, 42). The models were subsequently energy-minimized, using the CHARMM22 forcefield (3), generalized Born implicit solvent model and a gradient descent algorithm as implemented in MOE 2006.08 (Chemical Computing Group, Montreal, Canada): first hydrogens only, then sidechains, finally backbone atoms, tethered by a force constant of 0.1. PROCHECK v.3.5 (4) was used to evaluate both the minimized models and the template structures. The models with fewest violations of backbone angles were equilibrated by Molecular Dynamics (MD) simulation. The MD simulations were carried out by NAMD Scalable Molecular Dynamics (5) in a water sphere with harmonic boundary conditions, CHARMM22 force field, and Langevin Dynamics in a canonic NVT ensemble, 2 fs timestep. After initial minimization for 10,000 discrete steps, the simulation was run for 20 ns at 310 K. The resulting trajectory was visualized and evaluated using VMD Visual Molecular Dynamics (6). The ANOLEA software (27), implementing a knowledge-based residue-wise pseudo-potential, was utilized to generate ”energy” profiles of the final model structures of A3C and the respective PDB templates, using five-residue window averaging. Differences between superposed structures were quantified as root mean square deviation (RMSD) of Cα atoms. Pictures of protein models were generated using UCSF Chimera v1 (7).

Additional references

1. Madhusudhan MS, Marti-Renom MA, Sanchez R & Sali A (2006) Variable gap penalty for protein sequence-structure alignment. *Protein Eng Des Sel* 19, 129-133.
2. Word JM, Lovell SC, Richardson JS & Richardson DC (1999) Asparagine and glutamine: using hydrogen atom contacts in the choice of side-chain amide orientation. *J. Mol. Biol.* 285, 1735-1747.
3. McKerell AD *et al.* (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B.* 102, 3586-3616.
4. Laskowski RA, MacArthur MW, Moss DS & Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* 26, 283-291.
5. Phillips JC *et al.* (2005) Scalable molecular dynamics with NAMD. *J. Comput. Chem.* 26, 1781-1802.
6. Humphrey W, Dalke A & Schulten K (1996) VMD: visual molecular dynamics. *J. Mol. Graph.* 14, 33-38.
7. Pettersen EF *et al.* (2004) UCSF Chimera-a visualization system for exploratory research and analysis. *J. Comput. Chem.* 25, 1605-1612.

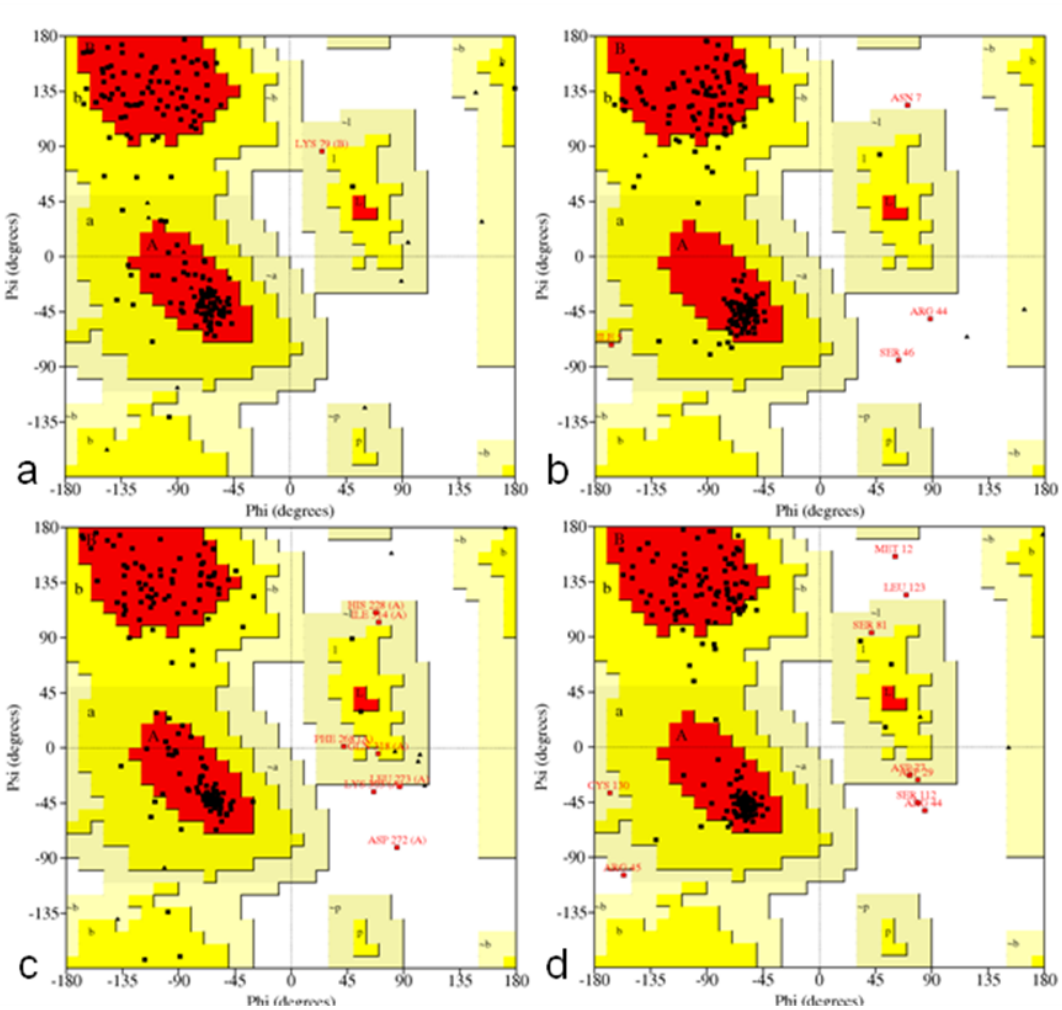


Figure S1. Ramachandran plots of (a) human APOBEC2 (A2), chain B (crystal structure), (b) A2-derived model structure of human APOBEC3C (A3C), (c) human APOBEC3G, C terminal domain (A3G-CD) and (d) A3G-CD-derived model structure of A3C. Stereochemical regions are colored red (most favourable), bright yellow (additionally allowed), pale yellow (generously allowed) and white (disallowed). Amino acids violating stereochemistry are indicated as red squares and numbered according to original PDB files / A3C amino acid sequence. Glycins are indicated as black triangles, all remaining amino acids as black squares.

(A2 crystal structure: 87.6% most favored, 11.8% additionally allowed, 0.6% generously allowed, 0% disallowed; A3G-CD solution structure: 84.8, 11.0, 2.4, 1.8%, respectively)

(A2-derived model: 83.7% most favorable, 14.0% additionally allowed, 0.6% generously allowed, 1.8% disallowed; A3G-CD-derived model: 84.9, 9.9, 2.3, 2.9%, respectively, with all generously and disallowed residues being in loop regions or the N-terminus modeled without template)

The global sequence identity of A3C and A2 is 30.2%, for secondary structure regions it is only marginally higher (32.8%). The global sequence identity of A3C and A3G-CD is 40.5%, but significantly higher (53.5%) for secondary structure regions.

Figure S2. Alignment of human APOBEC3C (A3C) to human APOBEC2 (A2) (a), human APOBEC3G, C-terminal domain (A3G-CD) (b) and human APOBEC3G, N-terminal domain (A3G-ND) (c). Secondary structure of A2 and A3G-CD as assigned by DSSP (H, helix; E, strand; else loop) are indicated above the sequences, if available. Identical amino acids are given in bold and denoted by ' | ', gaps in the respective sequence by '-'. The Zn2+-coordinating deaminase motif is highlighted (red box). A2 and A3C share 30.2% identical amino acids, A3G-CD and A3C 40.5%, A3G-ND and A3C 44.0%.

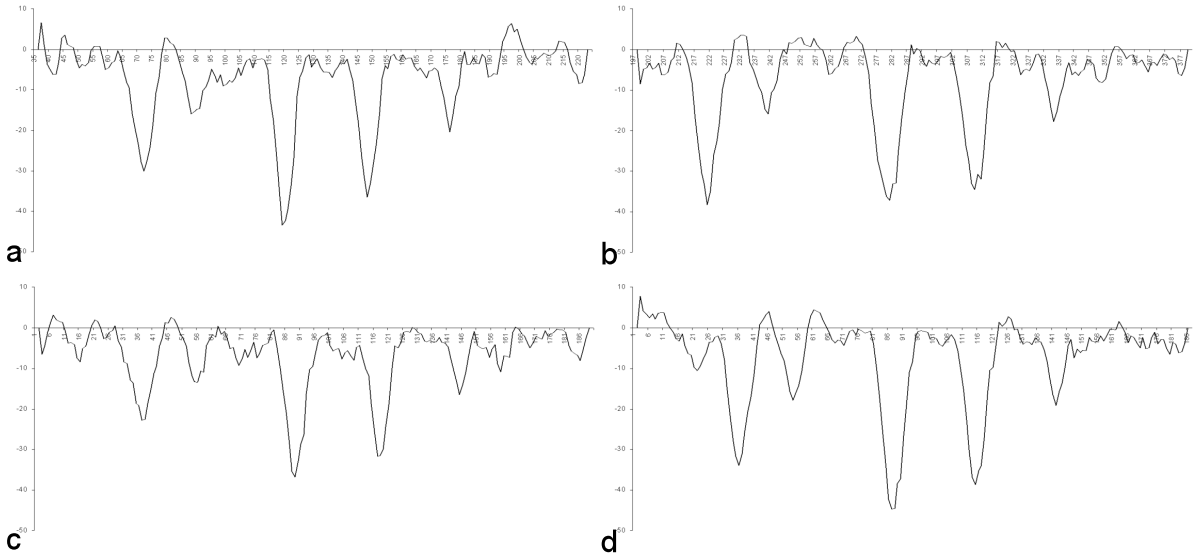


Figure S3. ANOLEA-profiles for (a) the A2 crystal structure, (b) the A3G-CTD crystal structure, (c) the A2-based model of A3C and (d) the A3G-CTD-based model structure of A3C. ANOLEA-"energies" represent a knowledge-based pseudo-potential and are plotted for each sequence position of the respective protein model.

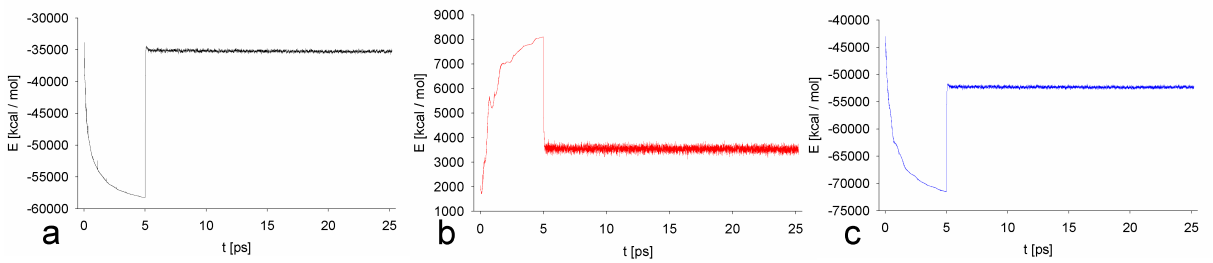


Figure S4. Equilibration of human APOBEC3C (A3C). Development of the total (a), van-der-Waals (b) and electrostatic energy (c) of the human A3C model over time. Energies are shown over the first 25 picoseconds (ps). During the course of the total 20 ns of simulation, energies are stationary (not shown).

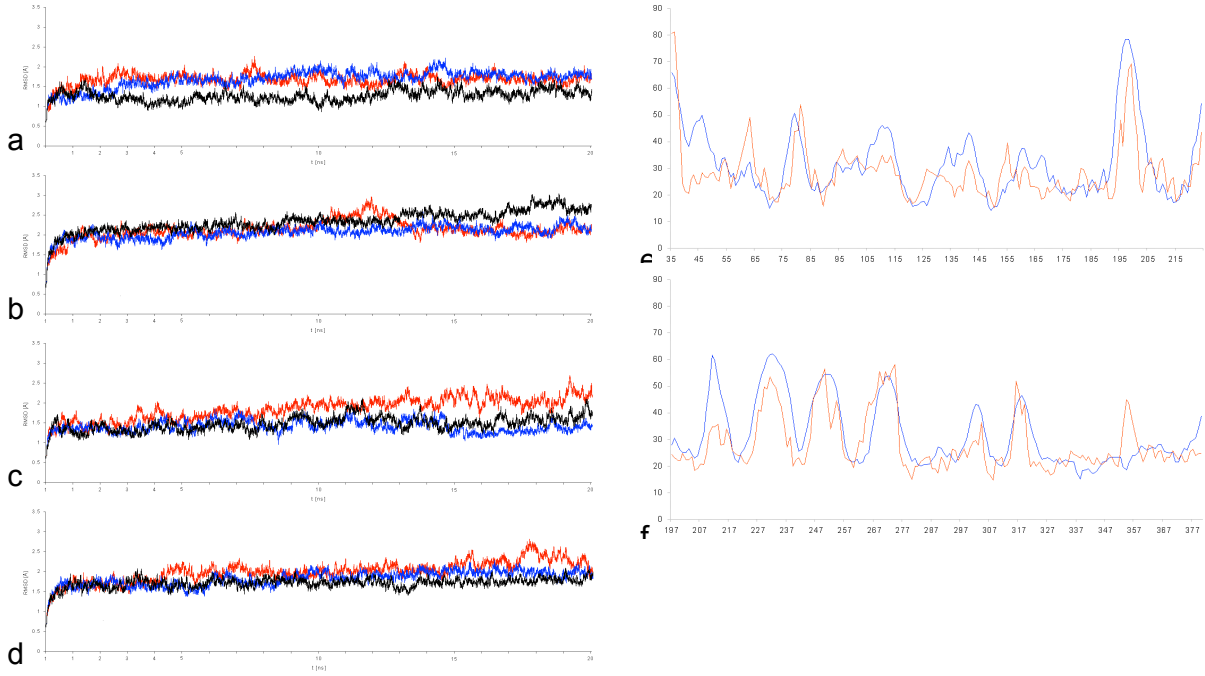


Figure S5. Evaluation of MD simulation of APOBEC structures. Development of the deviation of the conformation of the crystal structure of A2 (a), the A2-derived model of A3C (b), the crystal structure of A3G-CTD (c), and the A3G-CTD-derived model of A3C (d) with respect to the starting conformation over the course of the simulation, measured in Å root mean square deviation (RMSD). Simulations have been replicated three times (black, blue, red curve). Experimental (blue) and simulated B-factors (red) of C-alpha atoms of human APOBEC2, chain B (e) and human APOBEC3G, C-terminal domain (f).

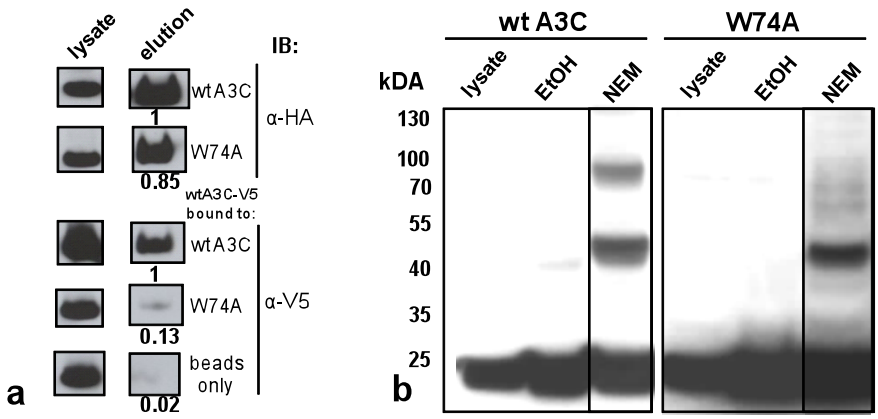


Figure S6. (a) Quantification of light units from immunoblot (IB) of bound V5-tagged wt A3C to indicated co-transfected HA-tagged wt A3C or W74A. Numbers indicate the signal intensity of V5-tagged wt A3C bound to W74A relative to the amount of V5-tagged wt A3C bound to HA-tagged wt A3C. Background interaction of V5-tagged wt A3C with anti-HA affinity beads was determined equally. (b) Immunoblot analysis of crosslinks of wt A3C or W74A in total cell lysate of transfected cells. Cleared cell lysate was treated with 50mM NEM for 2hrs at room temperature and separated on 4-12% SDS-PAGE. A3C wt or mutant were detected by an anti(α)-HA antibody.

DANKE…

…Carsten für Deine großartige Unterstützung in dem Projekt, die Antworten auf (fast) alle Fragen, Dein unendliches Wissen über Viren, das hervorragende Arbeitsklima, Deine IMMER offene Tür, die Möglichkeit an so vielen Meetings teilzunehmen und, na klar, für 3C!

…Prof. Dr. Klaus Cichutek und Prof. Dr. Dieter Häussinger für die Möglichkeit in Ihren Laborräumen und mit Ihren Mitteln forschen zu dürfen.

…Gerd und Arnulf dafür, dass ihr während der 4 Jahre immer Zeit gefunden habt, meine Daten zu besprechen und kritisch zu beäugen, sowie für die unkomplizierte Übernahme der Begutachtung dieser Arbeit.

…an die Lab-Münk Stamm-„11“:
…JZ für 4 Jahre als „Bench-Nachbar des Vertrauens“, Ironie, tolle Touri-Tage in New York, wilde politische Theorien und Deine Hilfsbereitschaft. Und wenn’s ma wieder länger dauert – so en „Z-P“ hat uns nie geschadet!
…Mario für all Deine Hilfe und Ideen im Labor und Deine Gelassenheit. Hätte gerne öfter an deinem „Kopfkino“ teilgehabt!
…Dani deine Herzlichkeit, Deinen Perfektionismus und das großartige Korrekturlesen! Ach ja, ich werde Dein Tiramisu und Deine Kuchen vermissen!

…Gisbert und Benny für die sensationelle Zusammenarbeit und Euer sexy Modell!

…an alle „PEI-ler“: Andre, Andreas, Björn, Bodo, Egbert, Elea, Ferdi, Heide, Matthias, Julia, Marion, Ralf, Stani, Sylvia und, last but not least, Frau Schmidt und Frau Varga.

…an alle „Labor-Düsseldorfer“: Anand, Andreas K, Andreas Pf, Benjamin, Björn, Boris, Carina, Christian, Heiner, Katerina, Mariana, Melli, Saskia, Soraya, Ute und besonders der Wio.

…an die ODW - Giessen - Darmstadt - Düsseldorf - Buben und Mädels, die ihr mir immer wieder zeigt, dass es mehr als Arbeit gibt! Danke für all die schönen und schlechten Stunden, die Ihr mit mir geteilt habt. Es ist schön zu wissen, dass es Euch gibt, egal wo auf der Welt ihr euch auch grad rumtreibt! Danke!

…Mama, Papa, Kerstin. Für alles!

Lebenslauf

Persönliche Daten

Name:	Henning Hofmann
geboren:	03. Februar 1981
in:	Erbach/Odenwald

Schulausbildung

1986 – 1990	Grundschule, Bad König/Zell
1990 – 2000	Gymnasium Michelstadt
06/2000	Abitur

Universitätsausbildung

10/2000 – 09/2002	Grundstudium der Biologie an der Justus-Liebig Universität in Giessen
09/2002	Vordiplom Biologie
10/2002 – 06/2006	Hauptstudium der Biologie an der Technischen Universität Darmstadt
04/2005 – 12/2005	Anfertigung der Diplomarbeit unter Anleitung von Prof. Dr. H. Ulrich Göringer am Institut für Mikrobiologie und Genetik der Technischen Universität Darmstadt Titel: „TbRGG1 – ein oligo(U) und oligo(G) bindendes Protein aus *Trypanosoma brucei* – Echtzeitstudien durch Oberflächen-resonanz“
06/2006	Diplom Biologie
08/2006 – 10/2008	Beginn der Anfertigung der Dissertation unter Anleitung von Dr. Carsten Münk in der Abteilung Medizinische Biotechnologie des Paul-Ehrlich-Instituts in Langen
10/2008 – 08/2010	Weiterführung der begonnen Dissertation unter Anleitung von Prof. Dr. Carsten Münk in der Klinik für Gastroenterologie, Hepatologie und Infektiologie an der Heinrich-Heine-Universität in Düsseldorf (Umzug der Arbeitsgruppe) Titel: „The human cytidine deaminase APOBEC3C restricts retroviruses independent of editing - a biochemical and structural analysis”