Published online 11 October 2006 *Nucleic Acids Research, 2006, Vol. 34, No. 19* ***5683–5694***
doi:10.1093/nar/gkl721

Evolutionarily conserved and non-conserved retrovirus restriction activities of artiodactyl APOBEC3F proteins

Stefán R. Jónsson[1,2,3,4], Guylaine Haché[1,2,3], Mark D. Stenglein[1,2,3], Scott C. Fahrenkrug[3,5], Valgerdur Andrésdóttir[4] and Reuben S. Harris[1,2,3,*]

[1]Department of Biochemistry, Molecular Biology and Biophysics, University of Minnesota, [2]Institute for Molecular Virology, [3]Arnold and Mabel Beckman Center for Transposon Research, Minneapolis, MN 55455, USA, [4]University of Iceland, Institute for Experimental Pathology, Keldur v/Vesturlandsveg, 112 Reykjavík, Iceland and [5]Department of Animal Sciences, University of Minnesota, St Paul, MN 55108, USA

Received August 4, 2006; Revised September 16, 2006; Accepted September 18, 2006

ABSTRACT

The APOBEC3 proteins are unique to mammals. Many inhibit retrovirus infection through a cDNA cytosine deamination mechanism. HIV-1 neutralizes this host defense through Vif, which triggers APOBEC3 ubiquitination and degradation. Here, we report an APOBEC3F-like, double deaminase domain protein from three artiodactyls: cattle, pigs and sheep. Like their human counterparts, APOBEC3F and APOBEC3G, the artiodactyl APOBEC3F proteins are DNA cytosine deaminases that locate predominantly to the cytosol and can inhibit the replication of HIV-1 and MLV. Retrovirus restriction is attributable to deaminase-dependent and -independent mechanisms, as deaminase-defective mutants retain significant anti-retroviral activity. However, unlike human APOBEC3F and APOBEC3G, the artiodactyl APOBEC3F proteins have an active N-terminal DNA cytosine deaminase domain, which elicits a broader dinucleotide deamination preference, and they are resistant to HIV-1 Vif. These data indicate that DNA cytosine deamination; sub-cellular localization and retrovirus restriction activities are conserved in mammals, whereas active site location, local mutational preferences and Vif susceptibility are not. Together, these studies indicate that some properties of the mammal-specific, APOBEC3-dependent retroelement restriction system are necessary and conserved, but others are simultaneously modular and highly adaptable.

INTRODUCTION

Expression of many of the human APOBEC3 (A3) proteins has been shown to inhibit the infective potential and mobility of a broad and growing number of retroviruses and retrotransposons [reviewed by (1–4)]. Humans encode seven A3 proteins, *Homo sapiens* (Hs) HsA3A, HsA3B, HsA3C, HsA3DE, HsA3F, HsA3G and HsA3H, in tandem on chromosome 22 (2,5–7). HsA3A has recently been shown to inhibit the mobility of both long terminal repeat (LTR) and non-LTR retrotransposons (8–10). HsA3B can also inhibit L1 and Alu retrotransposition, as well as the replication of SIV and to a lesser extent HIV-1 (8–13). HsA3C potently inhibits SIV, but it has shown little activity against other substrates (11,13). HsA3DE was recently shown to possess weak antiviral activity (14). HsA3G was the first member of this family to be associated with HIV restriction (15). HsA3F and HsA3G are both capable of potently inhibiting a variety of exogenous and endogenous retroelements (9,10,12,13,15–25). Finally, although HsA3H elicited DNA cytosine deaminase activity, it was unable to restrict SIV or HIV-1 replication (26). Several simian (e.g. chimpanzee), one carnivore (cat) and one rodent (mouse) APOBEC3 protein have also been shown to possess retroelement restriction activities [e.g. (20,25,27–30)]. APOBEC3 proteins from other mammals have yet to be examined.

Three themes appear to be emerging from these studies. First, the A3 proteins deaminate cytosines to uracils (C→U) within single-strand DNA (ssDNA). This property enables the A3 proteins to target the cDNA replication intermediates of all of the aforementioned retroviruses and retrotransposons. Second, retroelement restriction is mediated by at least two distinct mechanisms—by retroviral cDNA cytosine deamination (the hallmark activity of this family of proteins) and by a

*To whom correspondence should be addressed. Tel: +1 612 624 0457; Fax: +1 612 625 2163; Email: rsh@umn.edu

deamination-independent mechanism that is not well understood [e.g. (12,31–34)]. Finally, many of the retroelements (especially the retroviruses) that have been examined in detail can evade A3-dependent restriction. For instance, the virion infectivity factor (Vif) of HIV-1 and SIV recruits a cellular ubiquitin ligation complex to purge cells of A3G, the Bet protein of several different foamy viruses appears to directly bind and neutralize A3G and other retroviruses have simply evolved to exclude A3 proteins from nascent virions (e.g. MLV and HTLV) (13,28–30,33,35–39).

Strong evidence indicates that the conflict between host A3 proteins and invasive retroelements is ancient. Phylogenetic analyses based on representative vertebrate genome sequences indicate that the A3 proteins are at least as old as the mammalian lineage, because rodents encode one and primates seven A3 proteins (2,6). Other vertebrates, such as birds and fish do not have A3 proteins *per se*, but they do encode activation-induced deaminase (AID), an A3 orthologue that uses DNA cytosine deamination to trigger immunoglobulin gene hypermutation and isotype switch recombination [recently reviewed by (40–42)]. Comparative studies of A3 proteins from humans and non-human primates, New World monkeys, such as the tamarin and the woolly monkey, have demonstrated that the mammalian A3 proteins have been under a strong positive selection for at least 33 million years (26,43,44). The strong and likely ongoing positive selection and the unparalleled A3 gene expansion from one in rodents to seven in primates combine to suggest that the A3 proteins form a highly flexible and adaptable innate host defense system, which may very well be capable of readily adapting to new and potentially invasive retroelements.

The evolutionary gap between rodents and primates is ~90–100 million years [e.g. (45)]. This large genetic distance enables some comparative studies, but it limits others. Therefore, to close some of this distance, to enable more extensive comparative studies, and to examine potentially novel A3-virus conflicts, we have cloned and characterized A3 proteins from representative artiodactyls: cattle (*Bos taurus*; Bt), sheep (*Ovis aries*; Oa) and pigs (*Sus scrofa*; Ss). These studies focused on artiodactyl A3 proteins that were predicted to be similar in size (ca. 400 amino acids) and domain organization (two conserved zinc-binding deaminase domains) to the mouse *Mus musculus* (Mm) A3 protein and the well-characterized HsA3F and HsA3G proteins. The results of our studies emphasize the importance of DNA cytosine deamination in retrovirus restriction, but they also highlight the existence of a conserved deaminase-independent restriction mechanism. Moreover, non-conserved properties, such as Vif susceptibility, active site location and local mutational preferences combine to suggest mechanistic flexibilities that the APOBEC3 proteins might employ while adapting to new and potentially threatening genetic challenges.

MATERIALS AND METHODS

Artiodactyl A3F cDNA sequences and expression plasmids

Escherichia coli expression plasmids were derived from pTrc99A (GE Healthcare). The SsA3F expression construct was made by PCR amplifying a cDNA [GB# BI346898 (46)] using 5′-NNNNGAGCTCAGGTACCACCATGGATCCTCAGCGCCTGAGAC and 5′-NNNNGTCGACTCATCTCGAGTCACTTCTTGATG, digesting the PCR product with KpnI and SalI, and ligating it into a similarly digested pTrc99A. The BtA3F expression construct was made in the same manner by PCR amplifying a cDNA [GB# BE684372 (47)] using oligonucleotides 5′-NNNNGAGCTCAGGTACCACCATGCAACCAGCCTACCGAGGC and 5′-NNNNGTCGACCTAAATTGGGGCCGTTAGGAT. OaA3F was obtained by first amplifying a fragment by degenerative PCR using oligonucleotides 5′-TWYRTVTCCTGGAGCCCCTG and 5′-CCRKWWGTWGTAGAGGCGR, which were made to conserved regions of APOBEC3 proteins using template cDNA from sheep macrophages. The amplified fragment was sequenced and used to make OaA3F-specific oligonucleotides for 3′ RACE to obtain the remainder of the OaA3F coding sequence, 5′-AACCAGGTCTATGCTGGGACT and 5′-CTGGGGATGTACCAGAATGTG were used with an oligo dT primer 5′-AAGCAGTGGTAACAACGCAGAGTACT$_{30}$VN. The OaA3F expression plasmid was made by amplification from cDNA of sheep macrophages using oligonucleotides 5′-NNNNGAGCTCAGGTACCACCATGCCCTGGATCAGCGACCAC and 5′-NNNNGTCGACCTAAGTCGGCGCCGTCAGGAT in the same manner as the BtA3F and SsA3F constructs. E→Q zinc-binding domain mutants were constructed using standard site-directed mutagenesis protocols (Stratagene). The following oligonucleotides were used: SsA3F (E87→Q) 5′-CCGACCCGCCCTGCCACGCCCAGCTCTGCTTCCTCTCTTGT and 5′-ACAAGAGAGGAAGCAGAGCTGGGCGTGGCAGGGCGGGTCGG; SsA3F (E267→Q) 5′-ACAAGAAAAAGCGACATGCACAAATTCGTTTTATTGACAAG and 5′-CTTGTCAATAAAACGAATTTGTGCATGTCGCTTTTTCTTGT; BtA3F (E80→Q) 5′-GTGGGACTCGCTGCCACACCCAACTCC;GCTTCCTGTCTTGG and 5′-CCAAGACAGGAAGCGGAGTTGGGTGTGGCAGCGAGTCCCAC; BtA3F (E260→Q) 5′-ACAAGAAGCAGCGGCATGCACAAATTCGCTTTATTGACAAG and 5′-CTTGTCAATAAAGCGAATTTGTGCATGCCGCTGCTTCTTGT; OaA3F (E67→Q) 5′-CTGGGACTCACTGCCACAGCCAACGCCGCTTCCTGTCTTGG and 5′-CCAAGACAGGAAGCGGCGTTGGCTGTGGCAGTGAGTCCCAG; OaA3F (E247→Q) 5′-ACAAGAAGCAGCGGCATGCACAAATTCGCTTTATTGACAAG and 5′-CTTGTCAATAAAGCGAATTTGTGCATGCCGCTGCTTCTTGT. Eukaryotic expression plasmids were derived from pcDNA3.1(+) (Invitrogen) by subcloning the KpnI/SalI-flanked A3 cDNAs from the *E.coli* expression plasmids.

Additional DNA constructs

HsA3F, HsA3G and HsAID expression plasmids were described previously (31,48). MmA3 was made in the same manner as the artiodactyl constructs using oligonucleotides 5′-NNNNGAGCTCGGTACCACCATGGGACCATTCTGTCTGGGA and 5′-NNNNGTCGACATCAAGACATCGGGGGTCCAAGCTG. E→Q zinc-binding domain mutants were constructed using standard site-directed mutagenesis protocols (Stratagene). The following oligonucleotides were

used: HsA3G (E67→Q) 5′-CTTAAGTACCACCCACAGATGAGATTCTTCC and 5′-GGAAGAATCTCATCTGTGGGTGGTACTTAAG; HsA3G(E259→Q) 5′-CCTTGAAGGCCGCCATGCACAGCTGTGCTTCCTGG and 5′-CCAGGAAGCACAGCTGTGCATGGCGGCCTTCAAGG; HsA3F(E67→Q) 5′-CTGAGCACCACGCACAAATGTGCTTCCTC and 5′-GAGGAAGCACATTTGTGCGTGGTGCTCAG; HsA3F(E251→Q) 5′-CCCATTGTCATGCACAAAGGTGCTTCCTC and 5′-GAGGAAGCACCTTTGTGCATGACAATGGG; MmA3(E73→Q) 5′-CCAGTATAAAAAGCAGATTTGAGCGTGGATGTTGTCCTTG and 5′-CAAGGACAACATCCACGCTCAAATCTGCTTTTTATACTGG; MmA3(E290→Q) 5′-CTTATCAAGGAAGAGGATTTGTGCATGCTGTTTGCCTTTC and 5′-GAAAGGCAAACAGCATGCACAAATCCTCTTCCTTGATAAG. Eukaryotic expression plasmids were derived from pcDNA3.1(+) (Invitrogen) and pEGFP-N3 (Clontech) by inserting KpnI and SalI or SacI and SalI-flanked A3 cDNAs from the *E.coli* expression plasmids. The coding sequences of HsAID, HsA3F, HsA3G and MmA3 used in this study are identical those represented by GenBank accession nos NP_065712, NP_660341, NP_068594 and AAH03314, respectively. The coding sequence of HsA3B has two amino acid substitutions, W228L and D316N (12), which distinguish it from GenBank NP_004891.

Amino acid alignments and phylogenetic analyses

Protein alignments and phylogenetic analyses of the conserved, zinc-binding deaminase cores were done using the Clustal W software (49). A rooted phylogenetic tree was constructed using the neighbor-joining method. The GenBank accession numbers of the sequences used in these comparisons were AAI01974 (HsA1), NP_112436 (MmA1), NP_037039 (RnA1), NP_006780 (HsA2), XP_538909 (CfA2), NP_001029527 (BtA2), NP_033824 (MmA2), XP_217334 (RnA2), XP_528775 (PtAID), NP_065712 (HsAID), NP_033775 (MmAID), XP_575660 (RnAID), NP_001033771 (BtAID), NP_663745 (HsA3A), NP_004891 (HsA3B), NP_055323 (HsA3C), Q96AK3 (HsA3DE), NP_660341 (HsA3F), NP_068594 (HsA3G), NP_861438 (HsA3H), AAT44392 (PtA3G), ABD72578 (PtA3H), AAH03314 (MmA3) and NP_001028875 (RnA3).

Rif^R mutation assays

The intrinsic DNA cytosine deaminase activity of the A3 proteins was assayed by quantifying the accumulation of Rif^R mutants in *ung*-deficient *E.coli* [e.g. (31,48)]. Briefly, dilutions of A3- or vector control-expressing single colonies were grown to saturation in a rich bacterial growth medium containing 100 μg/ml ampicillin and 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG). Aliquots were spread to plates containing rifampicin (100 μg/ml; Sigma) to select for Rif^R mutants. Dilutions were spread to plates containing ampicillin to determine the number of viable cells. Mutation frequencies were calculated as the number of Rif^R mutants per 10^7 viable cells. For each experimental condition, 8 or 12 independent transformants were analyzed.

Fluorescence microscopy

Cells were visualized on a Zeiss Axiovert 200 microscope at 400× total magnification. A total of 7500 HeLa cells were seeded on LabTek chambered coverglasses (Nunc). After 24 h of incubation, these cells were transfected with 200 ng of the pEGFP-N3-based DNA constructs. After an additional 24 h of incubation, images of the live cells were collected. HsA3B, HsA3F and HsA3G–eGFP fusion constructs were reported previously (12).

Green fluorescent protein (GFP)-based retrovirus infection assays

293T cells were grown in DMEM supplemented with 10% FBS, penicillin and streptomycin. Viruses were produced by lipid-mediated transfection (Fugene; Roche) of 50% confluent 293T cells with the following plasmids. HIV-GFP: 0.22 μg of CS-CG (50), 0.14 μg of pRK5/Pack1(Gag-Pol), 0.07 μg of pRK5/Rev, 0.07 μg of pMDG (VSV-G Env) and 0.5 μg of an A3 expression plasmid or an empty vector control, as described previously (31,51). In some experiments, an HIV-1 Vif or a ΔVif control plasmid was included (0.5 μg). MLV-GFP was produced similarly, except the proviral plasmid was pM3P-GFP and a MLV gag/pol construct was used (18,51). After 48 h of incubation, the virus-containing supernatants were clarified by low speed centrifugation, filtered (0.45 μm), and quantified using a reverse transcriptase activity-based enzyme-linked immunosorbent assay (ELISA) (Cavidi Tech). Reverse transcriptase-normalized supernatants were applied to fresh 293T cells, and infection was allowed to proceed for 72 h. Infectivity (GFP fluorescence) was measured by flow cytometry (FACSCalibur, BD Biosciences). For experiments requiring the recovery of retroviral DNA for hypermutation analyses, the viral supernatants were treated with 50 U/ml DNase (Sigma) prior to 293T cell infection.

Retroviral DNA sequence analyses

Genomic DNA was incubated with DpnI to remove possible contaminating input CS-CG plasmid DNA. GFP was amplified using high-fidelity Phusion polymerase (Finnzymes) and CS–CG-specific primers, 5′-CGTGTACGGTGGGAGGTCTA and 5′-TTGGTAGCTGCTGTGTTGCT. PCR products were cloned and sequenced as described previously (51). Mutational analyses were performed using Sequencher software (Gene Codes Corp.).

GenBank accession numbers

The OaA3F cDNA sequence has been assigned GenBank DQ974645. The GenBank EST entries for BtA3F (BE684372) and SsA3F (BI346898) have been updated to include the full cDNA sequence (DQ974646 and DQ974647, respectively).

RESULTS

The double deaminase domain APOBEC3F proteins of artiodactyls

All known A3 proteins have either one or two conserved, zinc-binding deaminase domains, consisting of amino acids

HXE-$X_{23–28}$-$PCX_{2–4}C$ (X can be any amino acid) (2,5–7). The histidine and the two cysteines coordinate zinc and the glutamate participates directly in the C→U deamination reaction. NCBI BLAST searches using the human and mouse A3 deaminase domains as query polypeptides revealed several artiodactyl ESTs, which suggested the presence of at least one A3 protein in cattle and pigs. Corresponding cDNA clones were obtained, sequenced and shown to encode A3 proteins with two putative zinc-binding, cytosine deaminase domains (Figure 1A and Supplementary Figures S1 and S2; Materials and Methods). The orthologous sheep double domain A3 cDNA sequence was obtained using a combination of degenerate PCR and RACE (Figure 1A and Supplementary Figures S1 and S2). All three of these A3 proteins were similar in size to the 373 amino acid HsA3F protein, except the pig A3 protein, which was slightly longer due to a unique C-terminal, serine-rich extension (Figure 1A and Supplementary Figures S1 and S2).

Nomenclature standards dictate that protein names should be assigned based on the closest human orthologue [e.g. (52,53)]. However, amino acid comparisons showed that the N- and the C-terminal artiodactyl deaminase domains were most similar to different human A3 proteins (data not shown). To resolve this ambiguity and to facilitate name assignments, we named the artiodactyl double domain proteins after 'the human double domain A3 protein with the highest degree of active site identity' (see below for experimental demonstrations). The active sites of these artiodactyl A3 proteins were 56–62% identical to HsA3F (Figure 1B and Supplementary Figure S1). With the exception of HsA3E, which has equivalent identity (but no demonstrated activity), all of the other human A3 proteins had less identity. Thus, we have named the double domain deaminase proteins from cow, sheep and pig after HsA3F and, hereafter, will refer to them as BtA3F, OaA3F and SsA3F, respectively. Amino acid alignments of the active deaminase domains plus five residues on each side showed that the cow and sheep A3F active sites are 78% identical (Figure 1B and Supplementary Figures S1 and S2). Both the cow and the sheep proteins shared a lower level of identity with the pig protein (56%). These identity differences are consistent with the estimated times of divergence from their last common ancestor, as cows and sheep were estimated to have diverged from each other 14–20 million years ago, whereas pigs diverged from this lineage approximately 35–55 million years ago [e.g. (54)].

Phylogenetic analyses of the artiodactyl A3F proteins

As described above, all A3 proteins have either one or two conserved, zinc-binding (Z) deaminase domains [e.g. Figure 1A]. These domains cluster into three distinct phylogenetic groups: Z1a, Z1b or Z2 (7). The human double Z domain proteins, HsA3F and HsA3G, have a Z1a/Z1a and a Z1a/Z1b organization, respectively, whereas the MmA3 protein has a Z1a/Z2 organization. Interestingly, all three of the artiodactyl A3F proteins have a Z1a/Z2 organization (Figure 1C).

The active sites of most of the human DNA cytosine deaminase-competent A3 proteins can be classed as Z1 [e.g. HsA3F is Z1a and HsA3G is Z1b (31,32)]. Recently, the only human protein with a Z2 designation, HsA3H, was also shown to possess DNA cytosine deaminase activity (26). It is not clear, which Z domain(s) of MmA3 is active (addressed below). Thus, the N-terminal Z1a or the C-terminal Z2 domain of the artiodactyl A3F proteins appeared to have the potential to catalyze DNA cytosine deamination.

The artiodactyl A3F proteins catalyze DNA cytosine deamination

To test whether the artiodactyl A3F proteins have the capacity to deaminate cytosines within ssDNA, the intrinisic mutator activity of these proteins was monitored using an *E.coli*-based mutation assay. Rif^R is attributable to base substitution mutations in the *E.coli* RNA polymerase B (*rpoB*) gene, and it occurs in approximately 1 of every 5 million bacterial cells. We have found previously that expression of several A3 family members, including HsAID, rat APOBEC1, HsA3C, HsA3F and HsA3G, can accelerate the accumulation of Rif^R mutations from a few- to several 100-fold (48,51,55). The mutator phenotype is accounted for by a pronounced C/G→T/A transition bias within *rpoB*. This assay therefore provides a robust measure of intrisic DNA cytosine deaminase activity.

Expression of each of the artiodactyl A3 proteins increased the Rif^R mutation frequency in *E.coli* from 3- to 7-fold, levels that were higher than those attributable to HsA3F but slightly lower than than those caused by HsA3G (Figure 2A). Curiously, expression of MmA3 failed to cause an *E.coli* mutator phenotype, despite the fact that it is clearly active and capable of deaminating the cDNA of a variety of retroelements [e.g. (20,27) and below]. It is not clear why MmA3 is inactive in this system, yet active in others.

Artiodactyl A3F DNA cytosine deamination preferences were examined by sequencing the *rpoB* gene of at least 100 independent Rif^R mutants (Figure 2B and C). In contrast to HsA3F and HsA3G, which preferentially deaminate cytosines at *rpoB* nucleotide positions 1721 and 1691, 5′-T<u>C</u> and 5′-C<u>C</u>, respectively (48,51), the artiodactyl A3F proteins showed less biased *rpoB* mutation spectra. OaA3F preferentially deaminated cytosine 1576, which is part of a 5′-G<u>C</u> dinucleotide. SsA3F also preferred cytosine 1576. However, SsA3F also clearly deaminated cytosine 1586, which is part of a 5′-A<u>C</u> dinucleotide. Interestingly, these two cytosines, C_{1576} and C_{1586}, are also preferred by HsAID [Figure 2B and (48,55)]. The fact that HsAID and SsA3F appear to share a 5′-purine-<u>C</u> deamination preference suggests that their common ancestor [perhaps an ancient AID (40)] may have had a similar target preference. However, several groups have reported that even subtle amino acid substitutions in HsA3 proteins can dramatically alter local DNA cytosine deamination preferences, and therefore it is not surprising that most of the present day A3 proteins show non-AID-like mutational preferences (13,31,56). For instance, BtA3F did not appear to have any prominent *rpoB* local mutation preference, as increased levels of C/G→T/A mutation were apparent at several sites. In conclusion, all three of the artiodactyl A3F proteins are capable of deaminating DNA cytosines to uracils, which triggers a corresponding shift in the pattern of C/G→T/A transition mutations within the *rpoB* mutation substrate.

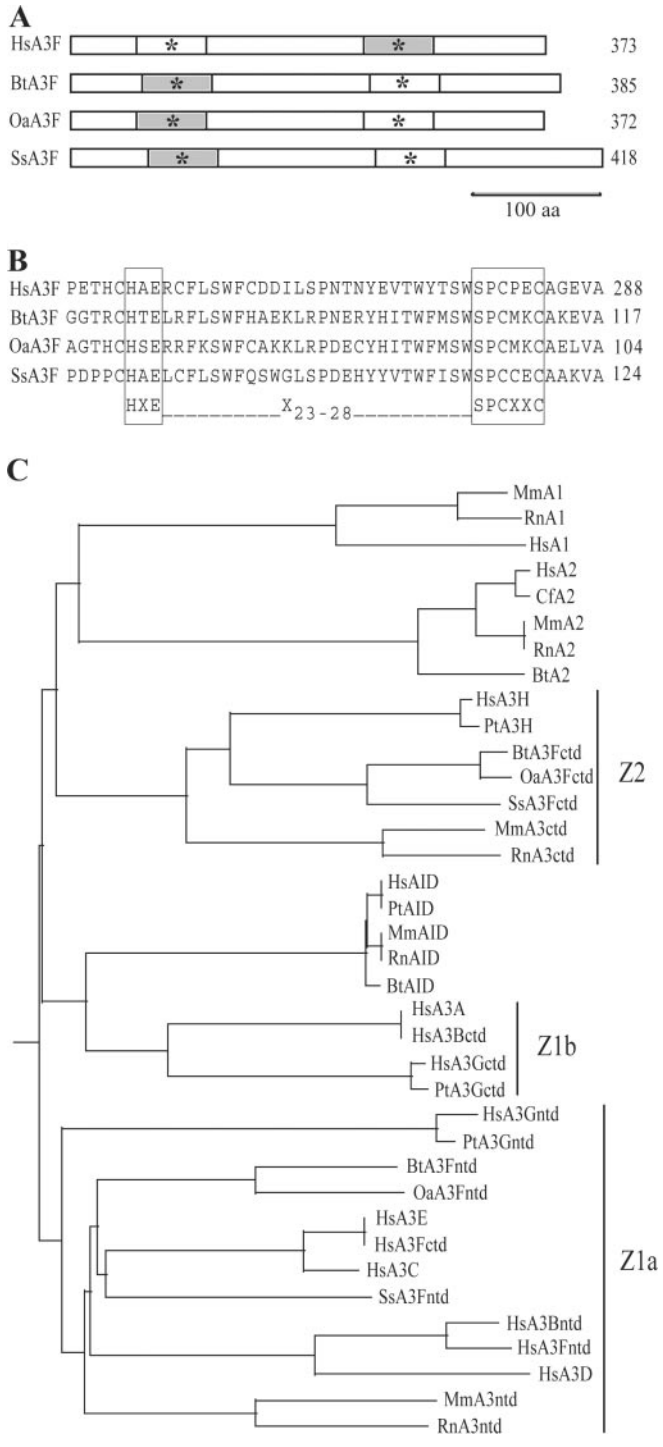

Figure 1. A comparison of artiodactyl A3F proteins. (**A**) A schematic of HsA3F, BtA3F, OaA3F and SsA3F. The conserved, zinc-binding deaminase domains are boxed (*) and those that are catalytically active are additionally shaded. The numbers on the right indicate the total polypeptide length. (**B**) HsA3F, BtA3F, OaA3F and SsA3F active site amino acid alignments [shaded regions from (A)]. The conserved HXE and PCXXC motifs are boxed. Amino acid positions are indicated on the right. (**C**) A neighbor-joining phylogenetic tree indicating the evolutionary relationship of several representative mammlian A3 family members. Branch lengths are proportional to the number of amino acid differences. Comparisons were done using the conserved Z domain amino acids, plus five additional residues on either side. The Z1a, Z1b and Z2 phylogenetic clusters are indicated. HsA3D and HsA3E represent the N- and C-terminal domains of HsA3DE. See the text for additional details. Bt, *B.taurus* (cow); Cf, *Canis familiaris* (dog); Hs, *H.sapiens* (human); Oa, *O.aries* (sheep); Pt, *Pan troglodytes* (chimpanzee); Rn, *Rattus norvegicus* (rat); Ss, *S.scrofa* (pig).

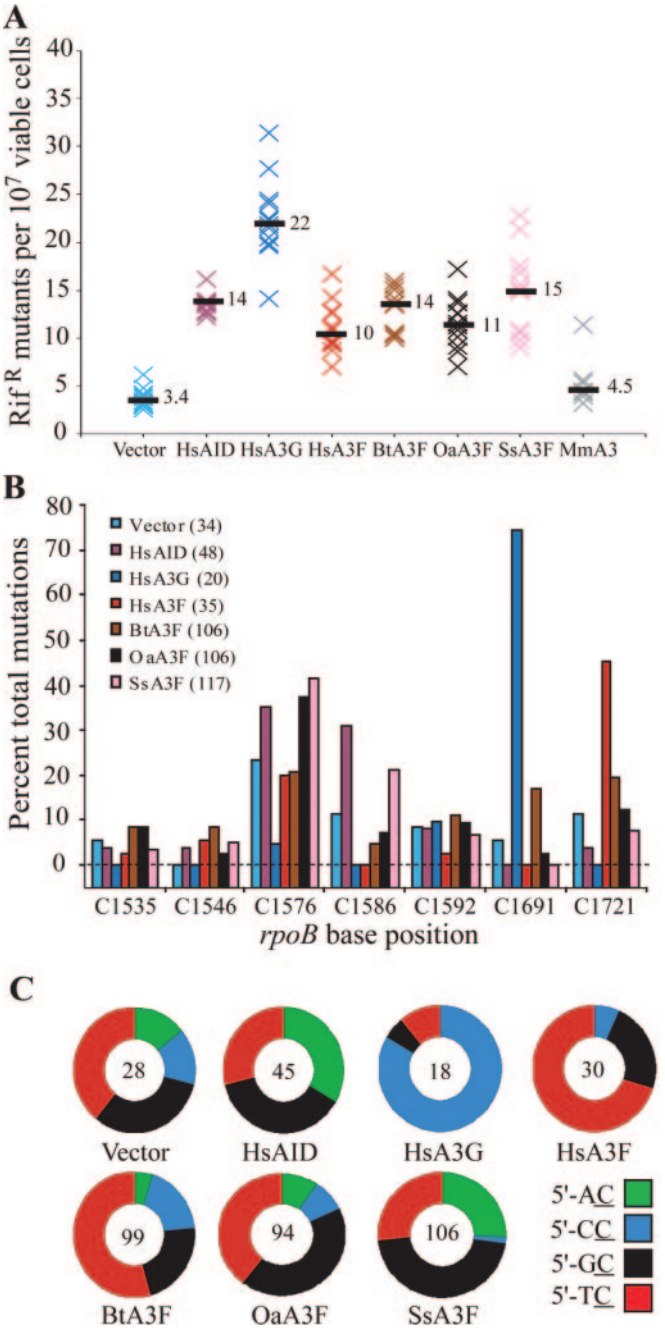

Figure 2. DNA cytosine deaminase activity of the artiodactyl A3F proteins in *E.coli*. (**A**) Rif^R mutation frequencies for 12 independent *E.coli* cultures expressing a vector control (light blue), HsAID (purple), HsA3G (dark blue), HsA3F (red), BtA3B (brown), OaA3B (black), SsA3F (pink) or MmA3 (gray). Each data point corresponds to the mutation frequency obtained from a single culture, and the median mutation frequency for each condition is shown (horizontal bar). (**B**) A histogram summarizing the C/G→T/A transition mutations detected in *rpoB*. Only cytosines that had greater than two mutations are shown. Apart from C_{1586}, all of the cytosines are located in the non-template strand of the *rpoB* gene. The number of independent Rif^R mutants that were sequenced is indicated in parentheses in the legend. For purposes of presentation the Y-axis extends below zero (a dotted line marks the actual base-line), and the histogram bars follow the color scheme (A). (**C**) Pie graphs depicting the frequency that each of the four dinucleotides was targeted by the indicated A3 protein. The total number of independent sequences analyzed is shown in the center of each graph, and the deaminated cytosine is underlined in the legend. The dinucleotide wedges of each pie are colored as indicated.

The artiodactyl A3F proteins are predominantly cytoplasmic

As an initial step toward understanding the potential retroelement targets of the artiodactyl A3F proteins, the sub-cellular distribution of these proteins was determined. A3-GFP constructs were transfected into HeLa cells and the sub-cellular localization of the A3 proteins was determined by live cell fluorescence microscopy. In agreement with prior work, HsA3B and a GFP-only control localized to the nucleus and cell-wide, respectively (12). In contrast, the artiodactyl A3F proteins and MmA3 appeared predominantly cytoplasmic (Figure 3). Many cells contained brightly fluorescing, punctate cytoplasmic aggregations, which may represent P bodies [e.g. (57)]. The significance of the cytoplasmic punctae remains to be determined. Nevertheless, the cytoplasmic localization pattern is nearly identical to that of HsA3F and HsA3G suggesting that the property of localizing to the cytoplasm is conserved and that the artiodactyl A3F proteins might function similarly to inhibit the replication of LTR-dependent retroviruses, such as HIV-1 or MLV [Figure 3; compare with (9,12,16,58)]. One should note, however, that the present data do not exclude the possibility that one or more of these A3 proteins might also possess the nucleo-cytoplasmic shuttling capability of AID, which also appears predominantly cytoplasmic but is clearly capable of entering and exiting the nuclear compartment where it triggers antibody gene diversification processes (59,60).

Retrovirus restriction by artiodactyl A3F proteins

A clear trend in the genetic conflict between A3 proteins and retroelements is that an A3 from a given host is either neutralized or avoided by retroelements that are specific to the host species. For instance, HIV-1 Vif counteracts both HsA3F and HsA3G, and SIV_{agm} Vif inhibits African green monkey A3G [e.g. (51,61–64). In many instances, cross-species comparisons enable potential species-specific mechanisms of neutralization to be avoided and the restrictive potential of A3 proteins to be studied. We therefore asked whether the artiodactyl A3F proteins could inhibit the infectivity of HIV-1 and MLV-based retroviruses. In these systems, a *GFP* gene embedded in proviral DNA provides a measure of both transfection efficiency (which correlates directly with virus production levels) and viral infectivity [e.g. (18,51)].

Expression of HsA3F and HsA3G caused 4- and 24-fold reductions in the infectivity of HIV-GFP, in agreement with previous studies [Figure 4A and (51,64)]. MmA3 was also capable of strongly inhibiting HIV-GFP. In comparison, expression of BtA3F, OaA3F or SsA3F caused 30-, 8- and 29-fold decreases in the infectivity of HIV-GFP, respectively (Figure 4A). These potent anti-HIV activities demonstrated that the artiodactyl A3F proteins possess at least one retrovirus restriction activity. These results further imply that the artiodactyl A3F proteins are able to specifically associate with the HIV-1 Gag–genomic RNA complex and thereby gain access to assembling virus particles (addressed further below).

Expression of MmA3 has little effect on the infectivity of MLV, presumably because MLV excludes (or simply avoids) this A3 protein [Figure 4B and (13,27,37,39)]. In contrast, HsA3F and HsA3G inhibit the infectivity of MLV-based

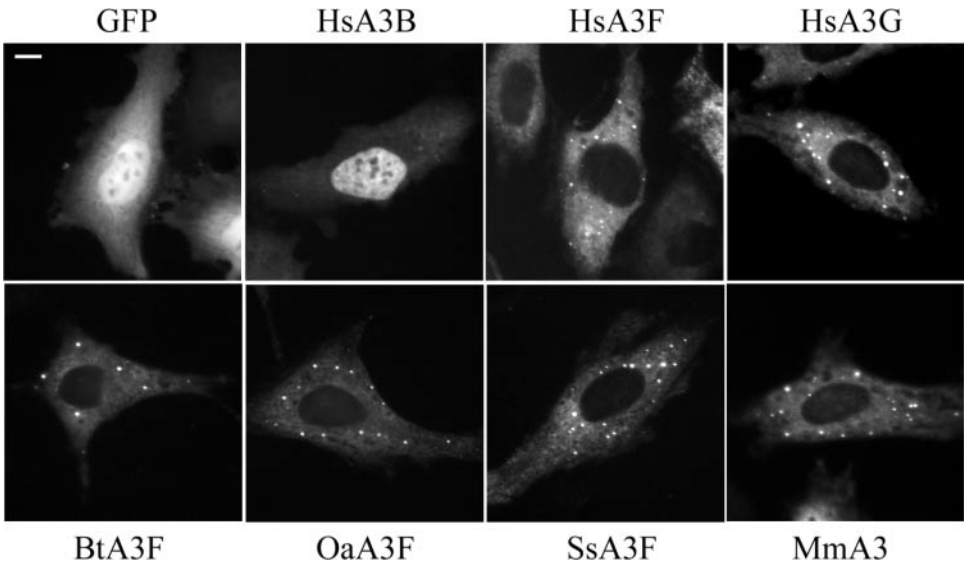

Figure 3. Sub-cellular distribution of the artiodactyl A3F proteins in comparison to the orthologous human and mouse A3 proteins. HeLa cells showing localization of the indicated, GFP-tagged A3 proteins or a GFP-only control. The scale bar indicates 10 μm.

retroviruses, but to a lesser extent than HIV lacking Vif [Figure 4B and, e.g. (18,51,64)]. Therefore, to ask whether the artiodactyl A3F proteins possess similar restriction potentials, the infectivity of MLV-GFP produced in the presence of these A3 proteins was monitored. Interestingly, similar to HsA3F or HsA3G, expression of the artiodactyl A3F proteins reduced the infectivity of MLV-GFP by 2- to 4-fold. Thus, the HIV-GFP and MLV-GFP infectivity data combined to suggest that the artiodactyl A3F proteins have a relatively broad retrovirus restriction potential.

The N-terminal zinc-binding, deaminase domain of the artiodactyl A3F proteins catalyzes C→U deamination, and this activity is necessary for full levels of retrovirus restriction

All of the double domain deaminases thus far characterized have catalytically competent C-terminal Z domains and apparently inert N-terminal Z domains (10,12,31,32,34). Most experiments have focused on HsA3B, HsA3F and HsA3G, DNA cytosine deaminases that have Z1a- or Z1b-type, C-terminal active sites (Figure 1). However, HsA3H, which has a single Z2-type deaminase motif, was also shown to possess DNA cytosine deaminase activity (26). Thus, it was possible that either the N- or the C-terminal Z domain (or both) of the artiodactyl A3F proteins would be catalytically active.

To begin to work-out the mechanism of retrovirus restriction by artiodactyl A3F proteins and to test whether the N- or the C-terminal (or both) Z domain of these proteins catalyzes DNA cytosine deamination, the conserved glutamate (E) of each active site was changed to glutamine (Q) and the resulting mutants were tested for HIV-GFP restriction activity. As reported previously, the glutamate of both the N- and the C-terminal Z domain of HsA3G contributed to inhibiting HIV-1 infectivity, but the C-terminal catalytic glutamate appeared to be more important [Figure 5A and (34,65,66)]. In contrast, both the N- and the C-terminal BtA3F Z domain E→Q mutants appeared to retain full levels of anti-HIV activity. Interestingly, the N-terminal OaA3F and SsA3F Z domain E→Q mutants were less able than the corresponding C-terminal domain mutants to inhibit the infectivity of HIV-GFP, a result particularly clear for SsA3F (Figure 5A). These data were essentially the inverse of the HsA3F and HsA3G E→Q mutant studies, and they suggested that the N-terminal, Z domain of these proteins catalyzes retroviral cDNA C→U deamination. MmA3 was clearly distinct, as both the N- and the C-terminal Z domain glutamates were required for HIV-GFP restriction (Figure 5A).

Although both the N- and the C-terminal Z domain E→Q mutants of the human and the artiodactyl A3 proteins showed significant levels of anti-retroviral activity, we surmised that bona fide catalytic site mutants should be unable to catalyze retroviral cDNA C→U deamination [although they may still inhibit retroviral infectivity; e.g. (32)]. Minus strand uracils template the incorporation of plus strand adenines, ultimately manifesting as retroviral plus strand G→A hypermutations. Therefore, to directly test which Z domain(s) catalyzes DNA cytosine deamination and to gain additional insight into the artiodactyl A3F retrovirus restriction mechanism, the *GFP* gene from the aforementioned HIV-GFP infectivity experiments was amplified by high-fidelity PCR, cloned and sequenced. HIV-GFP produced in the presence of a control vector showed a low base substitution mutation frequency, 0.00014 mutations per base, which is attributable to errors in RT and PCR (Figure 5B). In contrast, viruses produced in the presence of HsA3F, HsA3G, all three of the artiodactyl A3F proteins or MmA3 showed between 30- and 80-fold more base substitution mutations, which were almost exclusively plus strand G→A transition mutations (Figure 5B and Supplementary Figure S3). As described previously by Malim and co-workers, HsA3G with a C-terminal domain E→Q mutation failed to cause retroviral hypermutation, although this variant still significantly inhibited HIV-GFP infectivity [Figure 5 and Supplementary Figure S3

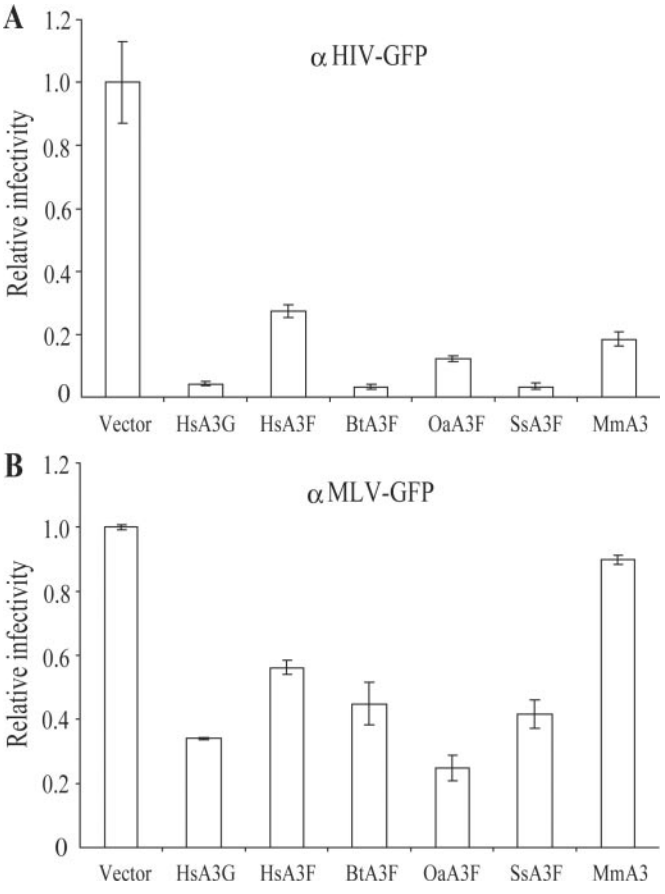

Figure 4. Retrovirus restriction activity of the artiodactyl A3F proteins. (**A**) Infectivity of HIV-GFP produced in the presence of a vector control or the indicated A3 protein. Data were normalized to the infectivity of HIV-GFP produced in the presence of a control vector, which was assigned a value of one. The mean and the SEM of three independent experiments are shown. HIV-1 Vif is not encoded by the proviral vector nor is it included *in trans* in these experiments (contrast with those shown in Figure 6). (**B**) Infectivity of MLV-GFP produced in the presence of the indicated constructs. Parameters are identical to those in (A).

and (32)]. The HsA3F C-terminal Z domain mutant was still able to modestly inhibit HIV-GFP infectivity, without obvious signs of retroviral hypermutation. Interestingly, E→Q substitutions in the N-terminal (but not the C-terminal) domain of all three of the artiodactyl A3F proteins abolished the accumulation of retroviral hypermutations (Figure 5B and Supplementary Figure S3). Thus, these data combined to demonstrate that the N-terminal Z domain of the artiodactyl A3F proteins is catalytic and that both deaminase-dependent and independent activities are required for full levels of retrovirus restriction. In support of this conclusion, the N-terminal Z domain glutamate of BtA3F is required for mutator activity in *E.coli*, whereas the C-terminal Z domain glutamate is dispensable (Figure 5C).

Retroviral hypermutation properties of artiodactyl A3F proteins

As described above, the *rpoB* mutation spectra of BtA3F, OaA3F and SsA3F suggested that these proteins would trigger retroviral hypermutation patterns biased toward

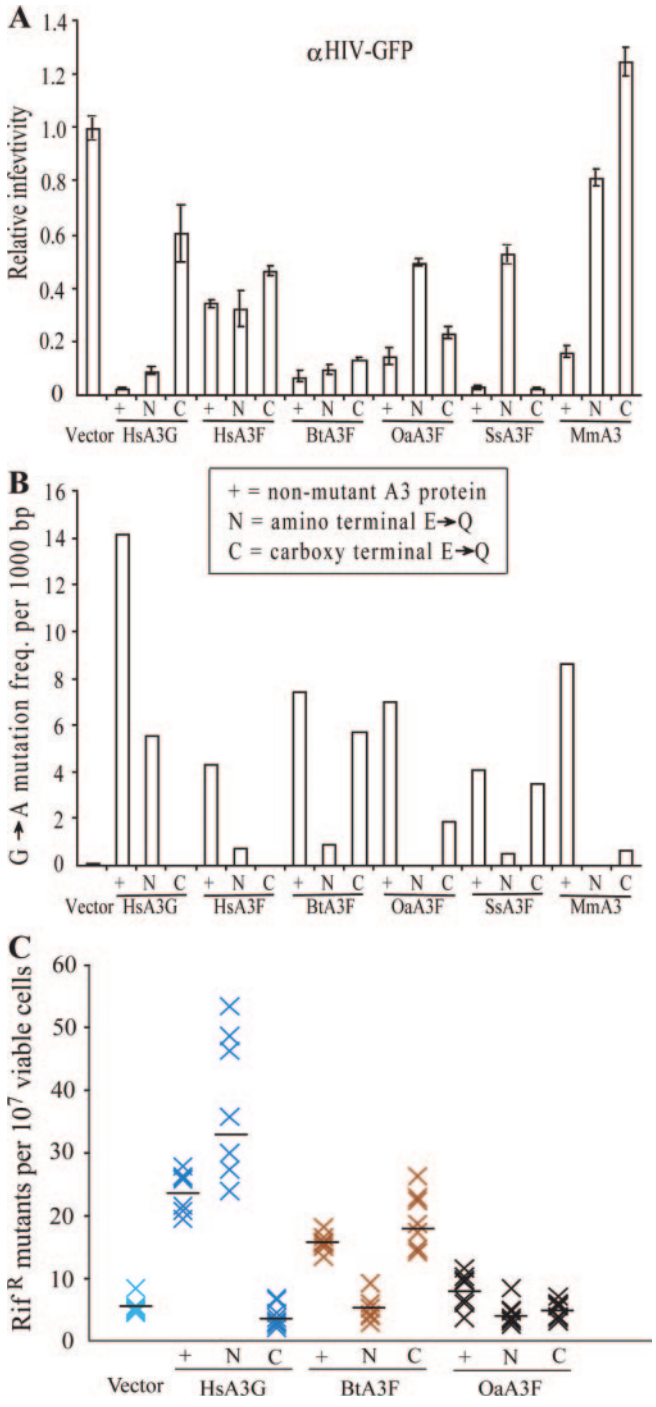

Figure 5. Relative contributions of the N- and C-terminal zinc-binding domains to HIV-GFP restriction. (**A**) Infectivity of HIV-GFP produced in the presence of a vector control or the indicated non-mutant (+) or mutant A3 protein containing an N-terminal E→Q substitution or a C-terminal E→Q substitution. Parameters are identical to those in Figure 4A. HIV-1 Vif is not included in these experiments (contrast with Figure 6). (**B**) Frequencies of plus strand retroviral G→A hypermutation observed in HIV-GFP DNA. (**C**) The N-terminal Z domain glutamate of the indicated artiodactyl A3F protein is required for mutator activity in *E.coli*, whereas the C-terminal Z domain glutamate is largely dispensable. Rif^R mutation frequencies for eight independent *E.coli* cultures expressing a vector control (light blue), HsA3G (dark blue), BtA3F (brown) or OaA3F (black). The addition sign (+), N and C refer to non-mutant, N-terminal Z domain E→Q mutant or C-terminal Z domain E→Q mutant proteins, respectively. Other parameters are identical to those in Figure 2A.

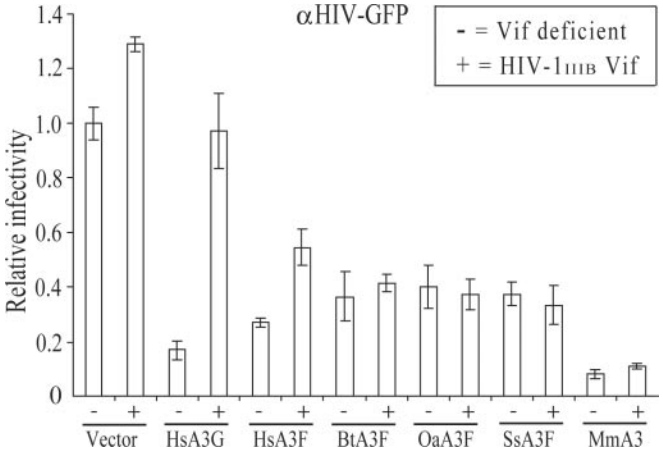

Figure 6. Artiodactyl A3F proteins are resistant to HIV-1 Vif. Parameters are identical to those used in Figure 4A, except for the inclusion of plasmids encoding HIV-1_{IIIB} Vif (+) or a HIV-1_{IIIB} ΔVif (−) control, which has translation stop codons at 33 and 34 amino acids (79).

5′-YC, 5′- GC, and 5′-RC, respectively (R = A or G; Y = C or T; Figure 2C). To test this prediction, we examined the types of base substitution mutations and the local retroviral cDNA deamination preferences attributable to expression of the artiodactyl A3F proteins (Figure 5B and Supplementary Figures S3 and S4). In terms of the dinucleotide mutation preferences, the base immediately 5′ of the targeted cytosine is a crucial target site determinant. HsA3F and HsA3G overwhelmingly preferred 5′-TC (84%) and 5′-CC (84%), respectively, whereas MmA3 preferred 5′-TC (61%) and 5′-CC (29%). Similar dinucleotide preferences were reported previously for these proteins [e.g. (13,18,51,64,67)]. Roughly paralleling the *E.coli* Rif^R mutation data, HIV-GFP sequences revealed that the cow and the sheep A3F proteins preferred a pyrimidine (Y) 5′ of the deaminated cytosine (93 and 79%, respectively; Supplementary Figure S4). Similarly, the pig A3F protein preferred 5′-GC (47%). This is notable because this is the only example of an A3 protein preferring 5′-purine-C [the immunoglobulin gene deaminase AID also has this preference; e.g. (68)]. In addition, the artiodactyl A3F proteins preferred a T at the −2 position, which is more similar to HsA3F than HsA3G (which prefers a C at this position).

The artiodactyl A3F proteins are fully resistant to HIV-1 Vif

HsA3F, HsA3G and chimpanzee A3G are neutralized by HIV-1 Vif [e.g. (51,64)]. However, many other monkey A3G proteins (e.g. African green monkey) and MmA3 are completely resistant (27,61–63,69). The full sets of interactions that govern the A3-Vif conflict have not been determined, and the artiodactyl A3F proteins are likely to prove useful in this regard. Therefore, HIV-GFP infectivity was monitored in the presence or absence of HIV-1 Vif and human, artiodactyl or mouse A3 proteins. As described previously, expression of HIV-1 Vif neutralized HsA3F and HsA3G (although the former to a lesser extent) and caused a proportional recovery of HIV-GFP infectivity [Figure 6; e.g. (13,51,64)]. Expression of HIV-1 Vif failed to enhance the infectivity of HIV-GFP produced in the presence of MmA3 or any of the artiodactyl A3F proteins. Thus, the artiodactyl A3F proteins were fully resistant to HIV-1 Vif.

DISCUSSION

Murine A3 and many of the human A3 proteins are capable of inhibiting the infectivity of a broad number of retroviruses and endogenous retroelements. Here, we have cloned and characterized a double domain A3 protein from three cloven hoof ungulates, the artiodactyls cattle, sheep and pigs. Artiodactyls are positioned between rodents and primates in most mammalian phylogentic trees. Thus, the artiodactyl A3F proteins can be considered evolutionary intermediates between the rodent and primate enzymes, and the studies presented here contribute to bridging this vast, 90–100 million year gap. The artiodactyl A3F data help delineate a number of conserved activities that define this group of proteins, including DNA cytosine deaminase activity, cytoplasmic subcellular localization and deaminase-dependent and independent retrovirus restriction activities. Moreover, the artiodactyl A3F proteins also help define non-conserved activities, the varying local DNA cytosine deamination preferences and the differential resistances to viral countermeasures, such as HIV-1 Vif.

One of the most intriguing non-conserved features of the artiodactyl A3F proteins is the fact that a DNA cytosine deaminase activity resides in the N-terminal Z domain, and not in the C-terminal Z domain like several of the human double domain A3 proteins. These data highlight the modular nature of the A3 proteins, which can clearly function as single domain proteins, as double domain proteins with a catalytically active C-terminal Z domain or as double domain proteins with a catalytically active N-terminal Z domain (e.g. HsA3H, HsA3G and SsA3F, respectively). We have not excluded the possibility that under some conditions or with other A3 proteins that both the N- and the C-terminal deaminase domains may be catalytically active. The mouse A3 protein offers an additional dimension of intrigue, as both the N- and the C-terminal zinc-binding deaminase domains are required for retrovirus restriction and for DNA cytosine deaminase activity (Figure 5 and Supplementary Figure 2). Prior studies have demonstrated that the N-terminal, zinc-binding domain of HsA3G mediates interactions with both HIV-1 Vif and Gag (70–76). The C-terminal domain of HsA3G dictates DNA cytosine deaminase activity, and both domains appear to contribute to dimerization activity (31,32,65). Thus, the division of activities between the N- and the C-terminal domains suggests that the A3 proteins are rapidly evolving as modular domains centered upon the conserved, zinc-binding motif. Such modularity may very well enable the A3 proteins to associate with each other to form highly adaptable and potent anti-retroelement defenses. For instance, in a multi-A3 protein complex, only one polypeptide needs to associate with the retroelement target in order to direct the restrictive potential of the other member(s) of the complex. Indeed, the first hints of such a combinatorial restriction potential were observed in retroviral hypermutation loads upon HsA3F and HsA3G co-expression (51).

Taken together with previous studies, it is clear that the mammalian A3 proteins can use both deamination-dependent

and -independent mechanisms to block the transmission of retroviruses and retrotransposons [e.g. (9,12,31–34,65,77)]. As discussed above, our studies indicate that the DNA cytosine deaminase activity can reside in either the N- or the C-terminal Z domain and that this domain appears to be the dominant contributor to restriction of HIV-based retrovirus substrates. Similar observations have been made previously for HsA3G (34,65,66). In contrast, other studies have indicated that either Z domain can mediate full (32,77) or partial (16) levels of HIV-1 restriction. These differential results may be attributable to differences in A3 protein expression levels. A resolution to this apparent paradox may be achieved by working-out the mechanism(s) of DNA deamination-independent restriction.

The studies presented here represent the first essential steps toward understanding the retroelement restriction activities of a third major branch of the vertebrate tree (the other two constituting rodents and primates). Ungulates, and specifically the artiodactyls cattle, sheep and pigs, provide humans with a number of benefits from food products to xenotransplantation possibilities. The possible conflict(s) between BtA3F and bovine immunodeficiency virus (BIV) and those between OaA3F and the sheep lentivirus, maedi-visna virus (MVV), will be of particular future interest. Ongoing studies have indicated that Vif-deficient MVV accumulates retroviral G→A hypermutations with an OaA3F-like dinucleotide mutation spectrum, suggesting that at least OaA3F is active *in vivo* [(78); S. Franzdóttir and V. Andrésdóttir, unpublished data]. Artiodactyl and human comparative studies may also contribute to understanding how HIV-1 uses Vif to neutralize human APOBEC3 proteins.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The authors thank R. LaRue for assistance with comparisons and nomenclature, T. Smith for providing cDNA clones and assistance with artiodactyl cDNA sequences, N. Martemyanova for providing expert technical assistance, M. Titus for microscopy facilities, N. Somia for sharing valuable reagents, and M. Murtaugh, P. Hackett, T. Smith and Harris laboratory members for helpful discussions. This work was supported by NIH grant AI064046. RSH is a Searle Scholar and a University of Minnesota McKnight Land Grant Professor. SRJ was the 2004–2005 Val Bjornson Icelandic Exchange Scholarship recipient and is supported in part by a grant from the Icelandic Research Fund. M.D.S. and G.H. were supported in part by a 3M Science and Technology Graduate Fellowship and an NSERC Graduate Studentship, respectively. Assistance with flow cytometry was provided by the University of Minnesota Comprehensive Cancer Center Flow Cytometry Core Facility. University of Minnesota Advanced Genetic Analysis Facility assisted with DNA sequencing. Funding to pay the Open Access publication charges for this article was provided by Spring Point Project.

Conflict of interest statement. None declared.

REFERENCES

1. Huthoff,H. and Malim,M.H. (2005) Cytidine deamination and resistance to retroviral infection: towards a structural understanding of the APOBEC proteins. *Virology*, **334**, 147–153.
2. Harris,R.S. and Liddament,M.T. (2004) Retroviral restriction by APOBEC proteins. *Nature Rev. Immunol*., **4**, 868–877.
3. Cullen,B.R. (2006) Role and mechanism of action of the APOBEC3 family of antiretroviral resistance factors. *J. Virol*., **80**, 1067–1076.
4. Bieniasz,P.D. (2004) Intrinsic immunity: a front-line defense against viral attack. *Nature Immunol*., **5**, 1109–1115.
5. Jarmuz,A., Chester,A., Bayliss,J., Gisbourne,J., Dunham,I., Scott,J. and Navaratnam,N. (2002) An anthropoid-specific locus of orphan C to U RNA-editing enzymes on chromosome 22. *Genomics*, **79**, 285–296.
6. Wedekind,J.E., Dance,G.S., Sowden,M.P. and Smith,H.C. (2003) Messenger RNA editing in mammals: new members of the APOBEC family seeking roles in the family business. *Trends Genet*., **19**, 207–216.
7. Conticello,S.G., Thomas,C.J., Petersen-Mahrt,S. and Neuberger,M.S. (2005) Evolution of the AID/APOBEC family of polynucleotide (deoxy)cytidine deaminases. *Mol. Biol. Evol*., **22**, 367–377.
8. Chen,H., Lilley,C.E., Yu,Q., Lee,D.V., Chou,J., Narvaiza,I., Landau,N.R. and Weitzman,M.D. (2006) APOBEC3A is a potent inhibitor of adeno-associated virus and retrotransposons. *Curr. Biol*., **16**, 480–485.
9. Bogerd,H.P., Wiegand,H.L., Doehle,B.P., Lueders,K.K. and Cullen,B.R. (2006) APOBEC3A and APOBEC3B are potent inhibitors of LTR-retrotransposon function in human cells. *Nucleic Acids Res*., **34**, 89–95.
10. Bogerd,H.P., Wiegand,H.L., Hulme,A.E., Garcia-Perez,J.L., O'Shea,K.S., Moran,J.V. and Cullen,B.R. (2006) Maintaining the integrity of the human genome: a cellular inhibitor of LINE-1 and Alu retrotransposition. *Proc. Natl Acad. Sci. USA*, **103**, 8780–8785.
11. Yu,Q., Chen,D., Konig,R., Mariani,R., Unutmaz,D. and Landau,N.R. (2004) APOBEC3B and APOBEC3C are potent inhibitors of simian immunodeficiency virus replication. *J. Biol. Chem*., **279**, 53379–53386.
12. Stenglein,M.D. and Harris,R.S. (2006) APOBEC3B and APOBEC3F inhibit L1 retrotransposition by a DNA deamination-independent mechanism. *J. Biol. Chem*., **23**, 16837–16841.
13. Bishop,K.N., Holmes,R.K., Sheehy,A.M., Davidson,N.O., Cho,S.J. and Malim,M.H. (2004) Cytidine deamination of retroviral DNA by diverse APOBEC proteins. *Curr. Biol*., **14**, 1392–1396.
14. Dang,Y., Wang,X., Esselman,W.J. and Zheng,Y.H. (2006) Identification of APOBEC3DE as another antiretroviral factor from the human APOBEC family. *J Virol*., doi: 10.1128/JVI.01123-06.
15. Sheehy,A.M., Gaddis,N.C., Choi,J.D. and Malim,M.H. (2002) Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature*, **418**, 646–650.
16. Mangeat,B., Turelli,P., Caron,G., Friedli,M., Perrin,L. and Trono,D. (2003) Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts. *Nature*, **424**, 99–103.
17. Zhang,H., Yang,B., Pomerantz,R.J., Zhang,C., Arunachalam,S.C. and Gao,L. (2003) The cytidine deaminase CEM15 induces hypermutation in newly synthesized HIV-1 DNA. *Nature*, **424**, 94–98.
18. Harris,R.S., Bishop,K.N., Sheehy,A.M., Craig,H.M., Petersen-Mahrt,S.K., Watt,I.N., Neuberger,M.S. and Malim,M.H. (2003) DNA deamination mediates innate immunity to retroviral infection. *Cell*, **113**, 803–809.
19. Schumacher,A.J., Nissley,D.V. and Harris,R.S. (2005) APOBEC3G hypermutates genomic DNA and inhibits Ty1 retrotransposition in yeast. *Proc. Natl Acad. Sci. USA*, **102**, 9854–9859.
20. Esnault,C., Heidmann,O., Delebecque,F., Dewannieux,M., Ribet,D., Hance,A.J., Heidmann,T. and Schwartz,O. (2005) APOBEC3G cytidine deaminase inhibits retrotransposition of endogenous retroviruses. *Nature*, **433**, 430–433.
21. Esnault,C., Millet,J., Schwartz,O. and Heidmann,T. (2006) Dual inhibitory effects of APOBEC family proteins on retrotransposition of mammalian endogenous retroviruses. *Nucleic Acids Res*., **34**, 1522–1531.
22. Dutko,J.A., Schafer,A., Kenny,A.E., Cullen,B.R. and Curcio,M.J. (2005) Inhibition of a yeast LTR retrotransposon by human APOBEC3 cytidine deaminases. *Curr. Biol*., **15**, 661–666.

23. Turelli,P., Mangeat,B., Jost,S., Vianin,S. and Trono,D. (2004) Inhibition of hepatitis B virus replication by APOBEC3G. *Science*, **303**, 1829.
24. Muckenfuss,H., Hamdorf,M., Held,U., Perkovic,M., Lower,J., Cichutek,K., Flory,E., Schumann,G.G. and Munk,C. (2006) APOBEC3 proteins inhibit human LINE-1 retrotransposition. *J. Biol. Chem.*, **281**, 22161–22172.
25. Delebecque,F., Suspene,R., Calattini,S., Casartelli,N., Saib,A., Froment,A., Wain-Hobson,S., Gessain,A., Vartanian,J.P. and Schwartz,O. (2006) Restriction of foamy viruses by APOBEC cytidine deaminases. *J. Virol.*, **80**, 605–614.
26. OhAinle,M., Kerns,J.A., Malik,H.S. and Emerman,M. (2006) Adaptive evolution and antiviral activity of the conserved mammalian cytidine deaminase APOBEC3H. *J. Virol.*, **80**, 3853–3862.
27. Mariani,R., Chen,D., Schrofelbauer,B., Navarro,F., Konig,R., Bollman,B., Munk,C., Nymark-McMahon,H. and Landau,N.R. (2003) Species-specific exclusion of APOBEC3G from HIV-1 virions by Vif. *Cell*, **114**, 21–31.
28. Kobayashi,M., Takaori-Kondo,A., Shindo,K., Abudu,A., Fukunaga,K. and Uchiyama,T. (2004) APOBEC3G targets specific virus species. *J. Virol.*, **78**, 8238–8244.
29. Russell,R.A., Wiegand,H.L., Moore,M.D., Schafer,A., McClure,M.O. and Cullen,B.R. (2005) Foamy virus Bet proteins function as novel inhibitors of the APOBEC3 family of innate antiretroviral defense factors. *J. Virol.*, **79**, 8724–8731.
30. Lochelt,M., Romen,F., Bastone,P., Muckenfuss,H., Kirchner,N., Kim,Y.B., Truyen,U., Rosler,U., Battenberg,M., Saib,A. *et al.* (2005) The antiretroviral activity of APOBEC3 is inhibited by the foamy virus accessory Bet protein. *Proc. Natl Acad. Sci. USA*, **102**, 7982–7987.
31. Haché,G., Liddament,M.T. and Harris,R.S. (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.
32. Newman,E.N., Holmes,R.K., Craig,H.M., Klein,K.C., Lingappa,J.R., Malim,M.H. and Sheehy,A.M. (2005) Antiviral function of APOBEC3G can be dissociated from cytidine deaminase activity. *Curr. Biol.*, **15**, 166–170.
33. Sasada,A., Takaori-Kondo,A., Shirakawa,K., Kobayashi,M., Abudu,A., Hishizawa,M., Imada,K., Tanaka,Y. and Uchiyama,T. (2005) APOBEC3G targets human T-cell leukemia virus type 1. *Retrovirology*, **2**, 32.
34. Navarro,F., Bollman,B., Chen,H., Konig,R., Yu,Q., Chiles,K. and Landau,N.R. (2005) Complementary function of the two catalytic domains of APOBEC3G. *Virology*, **333**, 374–386.
35. Yu,Y., Xiao,Z., Ehrlich,E.S., Yu,X. and Yu,X.F. (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.
36. Mehle,A., Goncalves,J., Santa-Marta,M., McPike,M. and Gabuzda,D. (2004) Phosphorylation of a novel SOCS-box regulates assembly of the HIV-1 Vif-Cul5 complex that promotes APOBEC3G degradation. *Genes Dev.*, **18**, 2861–2866.
37. Doehle,B.P., Schafer,A., Wiegand,H.L., Bogerd,H.P. and Cullen,B.R. (2005) Differential sensitivity of murine leukemia virus to APOBEC3-mediated inhibition is governed by virion exclusion. *J. Virol.*, **79**, 8201–8207.
38. Mahieux,R., Suspene,R., Delebecque,F., Henry,M., Schwartz,O., Wain-Hobson,S. and Vartanian,J.P. (2005) Extensive editing of a small fraction of human T-cell leukemia virus type 1 genomes by four APOBEC3 cytidine deaminases. *J. Gen. Virol.*, **86**, 2489–2494.
39. Abudu,A., Takaori-Kondo,A., Izumi,T., Shirakawa,K., Kobayashi,M., Sasada,A., Fukunaga,K. and Uchiyama,T. (2006) Murine Retrovirus escapes from murine APOBEC3 via two distinct novel mechanisms. *Curr. Biol.*, **16**, 1565–1570.
40. MacDuff,D.A. and Harris,R.S. (2006) Directed DNA deamination by AID/APOBEC3 in immunity. *Curr. Biol.*, **16**, R186–R189.
41. Petersen-Mahrt,S. (2005) DNA deamination in immunity. *Immunol. Rev.*, **203**, 80–97.
42. Longerich,S., Basu,U., Alt,F. and Storb,U. (2006) AID in somatic hypermutation and class switch recombination. *Curr. Opin. Immunol.*, **18**, 164–174.
43. Zhang,J. and Webb,D.M. (2004) Rapid evolution of primate antiviral enzyme APOBEC3G. *Hum. Mol. Genet.*, **13**, 1785–1791.
44. Sawyer,S.L., Emerman,M. and Malik,H.S. (2004) Ancient adaptive evolution of the primate antiviral DNA-editing enzyme APOBEC3G. *PLoS. Biol.*, **2**, E275.
45. Nei,M., Xu,P. and Glazko,G. (2001) Estimation of divergence times from multiprotein sequences for a few mammalian species and several distantly related organisms. *Proc. Natl Acad. Sci. USA*, **98**, 2497–2502.
46. Fahrenkrug,S.C., Smith,T.P., Freking,B.A., Cho,J., White,J., Vallet,J., Wise,T., Rohrer,G., Pertea,G., Sultana,R. *et al.* (2002) Porcine gene discovery by normalized cDNA-library sequencing and EST cluster assembly. *Mamm. Genome*, **13**, 475–478.
47. Smith,T.P., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T., Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C. *et al.* (2001) Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle. *Genome Res.*, **11**, 626–630.
48. Harris,R.S., Petersen-Mahrt,S.K. and Neuberger,M.S. (2002) RNA editing enzyme APOBEC1 and some of its homologs can act as DNA mutators. *Mol. Cell*, **10**, 1247–1253.
49. Thompson,J.D., Higgins,D.G. and Gibson,T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.*, **22**, 4673–4680.
50. Miyoshi,H., Blomer,U., Takahashi,M., Gage,F.H. and Verma,I.M. (1998) Development of a self-inactivating lentivirus vector. *J. Virol.*, **72**, 8150–8157.
51. Liddament,M.T., Brown,W.L., Schumacher,A.J. and Harris,R.S. (2004) APOBEC3F properties and hypermutation preferences indicate activity against HIV-1 *in vivo*. *Curr. Biol.*, **14**, 1385–1391.
52. Blake,J.A., Eppig,J.T., Bult,C.J., Kadin,J.A. and Richardson,J.E. (2006) The Mouse Genome Database (MGD): updates and enhancements. *Nucleic Acids Res.*, **34**, D562–D567.
53. Yu,M. and Irwin,D.M. (1996) Evolution of stomach lysozyme: the pig lysozyme gene. *Mol. Phylogenet. Evol.*, **5**, 298–308.
54. Reza Shariflou,M. and Moran,C. (2000) Conservation within artiodactyls of an AATA interrupt in the IGF-I microsatellite for 19-35 million years. *Mol. Biol. Evol.*, **17**, 665–669.
55. Petersen-Mahrt,S.K., Harris,R.S. and Neuberger,M.S. (2002) AID mutates *E.coli* suggesting a DNA deamination mechanism for antibody diversification. *Nature*, **418**, 99–103.
56. Langlois,M.A., Beale,R.C., Conticello,S.G. and Neuberger,M.S. (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.
57. Wichroski,M.J., Robb,G.B. and Rana,T.M. (2006) Human retroviral host restriction factors APOBEC3G and APOBEC3F localize to mRNA processing bodies. *PLoS. Pathog.*, **2**, e41.
58. Marin,M., Rose,K.M., Kozak,S.L. and Kabat,D. (2003) HIV-1 Vif protein binds the editing enzyme APOBEC3G and induces its degradation. *Nature Med.*, **9**, 1398–1403.
59. Ito,S., Nagaoka,H., Shinkura,R., Begum,N., Muramatsu,M., Nakata,M. and Honjo,T. (2004) Activation-induced cytidine deaminase shuttles between nucleus and cytoplasm like apolipoprotein B mRNA editing catalytic polypeptide 1. *Proc. Natl Acad. Sci. USA*, **101**, 1975–1980.
60. McBride,K.M., Barreto,V., Ramiro,A.R., Stavropoulos,P. and Nussenzweig,M.C. (2004) Somatic hypermutation is limited by CRM1-dependent nuclear export of activation-induced deaminase. *J. Exp. Med.*, **199**, 1235–1244.
61. Bogerd,H.P., Doehle,B.P., Wiegand,H.L. and Cullen,B.R. (2004) A single amino acid difference in the host APOBEC3G protein controls the primate species specificity of HIV type 1 virion infectivity factor. *Proc. Natl Acad. Sci. USA*, **101**, 3770–3774.
62. Schrofelbauer,B., Chen,D. and Landau,N.R. (2004) A single amino acid of APOBEC3G controls its species-specific interaction with virion infectivity factor (Vif). *Proc. Natl Acad. Sci. USA*, **101**, 3927–3932.
63. Mangeat,B., Turelli,P., Liao,S. and Trono,D. (2004) A single amino acid determinant governs the species-specific sensitivity of APOBEC3G to Vif action. *J. Biol. Chem.*, **279**, 14481–14483.
64. Wiegand,H.L., Doehle,B.P., Bogerd,H.P. and Cullen,B.R. (2004) A second human antiretroviral factor, APOBEC3F, is suppressed by the HIV-1 and HIV-2 Vif proteins. *EMBO J.*, **23**, 2451–2458.
65. Shindo,K., Takaori-Kondo,A., Kobayashi,M., Abudu,A., Fukunaga,K. and Uchiyama,T. (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.
66. Iwatani,Y., Takeuchi,H., Strebel,K. and Levin,J.G. (2006) Biochemical activities of highly purified, catalytically active human APOBEC3G: correlation with antiviral effect. *J. Virol.*, **80**, 5992–6002.
67. Yu,Q., Konig,R., Pillai,S., Chiles,K., Kearney,M., Palmer,S., Richman,D., Coffin,J.M. and Landau,N.R. (2004) Single-strand specificity of APOBEC3G accounts for minus-strand deamination of the HIV genome. *Nature Struct. Mol. Biol.*, **11**, 435–442.
68. Beale,R.C., Petersen-Mahrt,S.K., Watt,I.N., Harris,R.S., Rada,C. and Neuberger,M.S. (2004) Comparison of the differential context-dependence of DNA deamination by APOBEC enzymes: correlation with mutation spectra *in vivo*. *J. Mol. Biol.*, **337**, 585–596.
69. Xu,H., Svarovskaia,E.S., Barr,R., Zhang,Y., Khan,M.A., Strebel,K. and Pathak,V.K. (2004) A single amino acid substitution in human APOBEC3G antiretroviral enzyme confers resistance to HIV-1 virion infectivity factor-induced depletion. *Proc. Natl Acad. Sci. USA*, **101**, 5652–5657.
70. Conticello,S.G., Harris,R.S. and Neuberger,M.S. (2003) The Vif protein of HIV triggers degradation of the human antiretroviral DNA deaminase APOBEC3G. *Curr. Biol.*, **13**, 2009–2013.
71. Alce,T.M. and Popik,W. (2004) APOBEC3G is incorporated into virus-like particles by a direct interaction with HIV-1 Gag nucleocapsid protein. *J. Biol. Chem.*, **279**, 34083–34086.
72. Cen,S., Guo,F., Niu,M., Saadatmand,J., Deflassieux,J. and Kleiman,L. (2004) The interaction between HIV-1 Gag and APOBEC3G. *J. Biol. Chem.*, **279**, 33177–33184.
73. Svarovskaia,E.S., Xu,H., Mbisa,J.L., Barr,R., Gorelick,R.J., Ono,A., Freed,E.O., Hu,W.S. and Pathak,V.K. (2004) Human apolipoprotein B mRNA-editing enzyme-catalytic polypeptide-like 3G (APOBEC3G) is incorporated into HIV-1 virions through interactions with viral and nonviral RNAs. *J. Biol. Chem.*, **279**, 35822–35828.
74. Luo,K., Liu,B., Xiao,Z., Yu,Y., Yu,X., Gorelick,R. and Yu,X.F. (2004) Amino-terminal region of the human immunodeficiency virus type 1 nucleocapsid is required for human APOBEC3G packaging. *J. Virol.*, **78**, 11841–11852.
75. Zennou,V., Perez-Caballero,D., Gottlinger,H. and Bieniasz,P.D. (2004) APOBEC3G incorporation into human immunodeficiency virus type 1 particles. *J. Virol.*, **78**, 12058–12061.
76. Schafer,A., Bogerd,H.P. and Cullen,B.R. (2004) Specific packaging of APOBEC3G into HIV-1 virions is mediated by the nucleocapsid domain of the gag polyprotein precursor. *Virology*, **328**, 163–168.
77. Bishop,K.N., Holmes,R.K. and Malim,M.H. (2006) Antiviral potency of APOBEC proteins does not correlate with cytidine deamination. *J. Virol.*, **80**, 8450–8458.
78. Kristbjornsdottir,H.B., Andresdottir,V., Svansson,V., Torsteinsdottir,S., Matthiasdottir,S. and Andresson,O.S. (2004) The vif gene of maedi-visna virus is essential for infectivity *in vivo* and *in vitro*. *Virology*, **318**, 350–359.
79. Simon,J.H., Southerling,T.E., Peterson,J.C., Meyer,B.E. and Malim,M.H. (1995) Complementation of vif-defective human immunodeficiency virus type 1 by primate, but not nonprimate, lentivirus vif genes. *J. Virol.*, **69**, 4166–4172.