APOBEC3G is a single-stranded (ss) DNA deaminase that restricts replication of HIV-1 by inducing viral genome mutagenesis through deamination of cytosine to uracil on HIV-1 cDNA. APOBEC3G has polydisperse oligomeric states and deaminates ssDNA processively through jumping and sliding. APOBEC3G has a catalytically inactive N-terminal CD1 domain that mediates processivity and an active C-terminal CD2 domain that catalyzes deaminations. Here, we assess the determinants of APOBEC3G deamination efficiency mediated by the CD1 domain by comparing native APOBEC3G and two CD1 mutants, a monomeric mutant (F126A/W127A) and a clinical mutant associated with high viral loads (H186R). Biochemical assays on ssDNA or partially dsDNA and with a reconstituted HIV replication system demonstrate that both mutants of Apo3G have altered DNA scanning properties in either jumping (F126A/W127A) or sliding (H186R), which results in decreased abilities to induce mutagenesis during reverse transcription. The data reveal a functionality for Apo3G oligomers in deamination and provide the first biochemical characterization of the clinical mutant H186R. The data demonstrate that the balance between the jumping and sliding of Apo3G is needed for efficient mutational inactivation of HIV-1.

Received for publication, November 1, 2010, and in revised form, February 6, 2011 Published, JBC Papers in Press, February 7, 2011, DOI 10.1074/jbc.M110.199604

Yuqing Feng1 and Linda Chelico2
From the Department of Microbiology and Immunology, University of Saskatchewan, Saskatoon, Saskatchewan S7N 5E5, Canada
Biochemical Basis of Apo3G Deamination Efficiency

substrate (25, 26). The dimer interface, which includes Phe-126 and Trp-127 as well as Tyr-124 and Tyr-125, is located on loop 7 in the N-terminal half of Apo3G, termed the CD1 (25, 27, 28).

Apo3G catalyzes deaminations processively (24–26). Processivity appears to be mediated by the CD1 (25). The CD1 domain contains a deamination motif that is catalytically inactive but can bind nucleic acids (29, 30). The CD2 domain near the C terminus contains the active deaminase domain (29, 30). As a result, the CD1 domain contributes indirectly to catalysis by mediating the ssDNA scanning mechanism of Apo3G (25). Previous to these data (25), the CD1 was thought to only mediate incorporation of Apo3G into virions through RNA binding (29). The ssDNA scanning mechanism of Apo3G appears to be by facilitated diffusion that entails a three-dimensional search involving both sliding and hopping or jumping movements (24, 31). Compared with a one-dimensional mechanism involving only sliding movements, a three-dimensional processive mechanism has the potential to increase the searching efficiency over a large area of DNA because the enzyme can move nonlinearly and sample both closely and distantly spaced regions in succession (32–34). A consequence of this type of movement is that the deaminations catalyzed by Apo3G are stochastic (31).

The stochastic properties may have a beneficial role in vivo because the mutations induced by Apo3G in HIV-1 proviral DNA will always be different, thereby avoiding selective pressure on the virus. However, the restriction mechanism of Apo3G is still not a “sure fire” way to inactivate HIV-1. Recently, it has been demonstrated through experimental and theoretical work that if Apo3G does not induce a sufficient mutational load in the HIV-1 genome, the genetic changes may contribute to adaptive viral evolution by sublethal mutagenesis (35–39). All together, these data have raised the issue as to whether the mutagenesis-based restriction mechanism of Apo3G can be reliably developed as a novel HIV-1 therapy (37, 40).

Here, we have examined the deamination mechanisms of native Apo3G and two CD1 mutants, the monomeric mutant F126A/W127A (F/W mutant) (25) and the naturally occurring H186R mutant of Apo3G (41, 42). The H186R mutant is a polymorphism that has been associated with high viral loads and the C terminus contains the active deaminase domain (29, 30). As a result, the CD1 domain contributes indirectly to catalysis by mediating the ssDNA scanning mechanism of Apo3G (25). Previous to these data (25), the CD1 was thought to only mediate incorporation of Apo3G into virions through RNA binding (29). The ssDNA scanning mechanism of Apo3G appears to be by facilitated diffusion that entails a three-dimensional search involving both sliding and hopping or jumping movements (24, 31). Compared with a one-dimensional mechanism involving only sliding movements, a three-dimensional processive mechanism has the potential to increase the searching efficiency over a large area of DNA because the enzyme can move nonlinearly and sample both closely and distantly spaced regions in succession (32–34). A consequence of this type of movement is that the deaminations catalyzed by Apo3G are stochastic (31).

The stochastic properties may have a beneficial role in vivo because the mutations induced by Apo3G in HIV-1 proviral DNA will always be different, thereby avoiding selective pressure on the virus. However, the restriction mechanism of Apo3G is still not a “sure fire” way to inactivate HIV-1. Recently, it has been demonstrated through experimental and theoretical work that if Apo3G does not induce a sufficient mutational load in the HIV-1 genome, the genetic changes may contribute to adaptive viral evolution by sublethal mutagenesis (35–39). All together, these data have raised the issue as to whether the mutagenesis-based restriction mechanism of Apo3G can be reliably developed as a novel HIV-1 therapy (37, 40).

Here, we have examined the deamination mechanisms of native Apo3G and two CD1 mutants, the monomeric mutant F126A/W127A (F/W mutant) (25) and the naturally occurring H186R mutant of Apo3G (41, 42). The H186R mutant is a polymorphism that has been associated with high viral loads and decreased CD4 T-cell counts in HIV-1-infected patients (41, 42). The H186R mutant is still catalytically active and, in cell culture, can be incorporated into budding HIV-1 virions making the mechanism for its deficiency in HIV-1-infected patients unknown (41). This is the first analysis of how a monomeric mutant of Apo3G scans ssDNA and the first biochemical characterization of the clinical Apo3G mutant H186R. We characterized how these Apo3G forms deaminate ssDNA, partially dsDNA, and ssDNA being actively synthesized in a model in vitro HIV replication system. Both the F/W mutant and H186R have altered DNA scanning properties from native Apo3G. We propose a model that accounts for the functionality of a dual sliding and jumping scanning mechanism in ensuring Apo3G causes HIV-1 gene inactivation and avoids sublethal mutagenesis.
Biochemical Basis of Apo3G Deamination Efficiency

cloned into the pSP72 vector (Promega) using EcoRI and BglII sites, with T7 RNA polymerase and a pool of nucleotides containing fluorescein-12-UTP (Roche Applied Science). Reaction mixtures (80 µl), containing an F-labeled RNA or DNA (50 nM) in RT buffer and varying concentrations of 0–2000 nM native or mutant Apo3G, were incubated at 25 °C. Rotational anisotropy was measured using a Quantamaster QM-4 Fluorometer (Photon Technology International) with a single emission channel. Samples were excited with vertically polarized light at 494 nm, and both vertical and horizontal emission were monitored at 520 nm (2.5-nm band pass).

Model HIV Replication Assay—The pET-Blue-PPT-Prot vector was linearized with SmaI. A 423-nt RNA consisting of the HIV-1 PPT and segment of the protease gene fused to lacZα was produced and purified after incubation of the linearized DNA with T7 RNA polymerase, rNTPs, and buffer according to manufacturer’s instructions (Promega RibomAX Kit). A primer-template was formed by heat-annealing a synthetic DNA primer to the 3'-end of the RNA (supplemental Table S1). The reverse transcription/replication reaction contained 50 nM primer-template, 500 µM dNTPs, 2.4 µM HIV-1 RT (p66/p51), 3 µM NC, with or without Apo3G (0.2 µM) in RT buffer. The concentrations of reaction components were modeled from expected concentrations in a virion (60 nm radius) according to the literature and added relative to the genomic DNA with T7 RNA polymerase, rNTPs, and buffer according to manufacturer’s instructions (Promega RibomAX Kit). A primer-template was formed by heat-annealing a synthetic DNA primer to the 3'-end of the RNA (supplemental Table S1). The reverse transcription/replication reaction contained 50 nM primer-template, 500 µM dNTPs, 2.4 µM HIV-1 RT (p66/p51), 3 µM NC, with or without Apo3G (0.2 µM) in RT buffer. The concentrations of reaction components were modeled from expected concentrations in a virion (60 nm radius) according to the literature and added relative to the genomic RNA (22, 46–48). Reactions were incubated at 37 °C for 90 min. The reaction was stopped by the addition of 0.5% SDS and 2.5 mM EDTA, followed by proteinase K treatment to digest the proteins (0.5 g/ml, 30 min at 25 °C). The sample was then extracted with phenol/chloroform/isoamyl alcohol and buffer exchanged with a Micro BioSpin P6 column (Bio-Rad).

This cleaned reaction was subjected to a PCR using Pfu Cx (Stratagene), a polymerase that can use uracil as a template (Bio-Rad). This cleaned reaction was subjected to a PCR using Pfu Cx (Stratagene), a polymerase that can use uracil as a template (Bio-Rad). This cleaned reaction was subjected to a PCR using Pfu Cx (Stratagene), a polymerase that can use uracil as a template (Bio-Rad). This cleaned reaction was subjected to a PCR using Pfu Cx (Stratagene), a polymerase that can use uracil as a template (Bio-Rad). This cleaned reaction was subjected to a PCR using Pfu Cx (Stratagene), a polymerase that can use uracil as a template (Bio-Rad).

RESULTS

Apo3G is a polydisperse oligomeric protein (26). Oligomerization is important for virion incorporation of Apo3G (12, 27, 55, 56), but the utility of oligomers thereafter is unclear. Although it is accepted that Apo3G must transverse DNA/RNA hybrid regions by jumping or intersegmental transfer during HIV replication to search for ssDNA regions where deamination can occur (24, 57, 58), the structural determinants of the processive jumping and sliding mechanism are not known. Here, we address the key mechanistic features necessary for Apo3G to catalyze processive deaminations using the F/W mutant and H186R mutant of Apo3G for loss of function analysis.

Pro cessivity and Polarity of Native and Mutant Apo3G—We examined if differences in the deamination mechanisms for native Apo3G and the mutants, F/W mutant and H186R, exist by using synthetic ssDNA substrates with two 5'-CCC deamination motifs spaced different distances apart. From these data, we can determine the degree of preference for deaminating C residues at the 5'-ssDNA end and if the mutants are able to deaminate two cytosine residues processively (31). These are hallmark features for native Apo3G (24).

Deamination of two 5'-CCC motifs 28 nt apart was tested on a synthetic ssDNA with a total length of 85 nt. The reactions were performed under “single hit” conditions (<15% substrate usage), which, by Poisson statistics, ensures that any given ssDNA substrate is acted on by at most one form of Apo3G (monomer or oligomer) (59). The specific activity of the native and mutant enzymes on this substrate differed, and the time course was varied for each enzyme to achieve <15% substrate usage. The specific activities were as follows: Apo3G, 8.7 pmol µg⁻¹ min⁻¹; F/W mutant, 22 pmol µg⁻¹ min⁻¹; and H186R mutant, 4.5 pmol µg⁻¹ min⁻¹.

Native Apo3G demonstrates a 3.5-fold preference for deamination of the C residue proximal to the 5'-ssDNA end (Fig. 1A,
Biochemical Basis of Apo3G Deamination Efficiency

deaminate each motif in a separate encounter with the ssDNA (uncorrelated) (24). The F/W and H186R mutants have similar processivity factors to the native Apo3G (Fig. 1A, compare processivity factors) but demonstrate an ~3-fold increase in the deamination polarity toward the 5′-ssDNA end (Fig. 1A, compare 5′C/3′C ratios). With these closely spaced CCC motifs, the processivity for the mutants appears to be comparable with the native enzyme.

Using a 118-nt substrate with two 5′CCC deamination motifs spaced further apart (61 nt), we tested if the polar and processive properties of the F/W and H186R mutants were retained. Native Apo3G prefers to deaminate the 5′-ssDNA end proximal CCC by a factor of ~3 (Fig. 1B, 5′C/3′C ratio) with a processivity factor of 8 (Fig. 1B, processivity factor). Similar to the deamination polarity on the shorter ssDNA substrate (Fig. 1A), the F/W mutant demonstrates a stronger preference than native Apo3G for C residues at the 5′-ssDNA end and is processive (Fig. 1B). The H186R mutant retains a strong preference for deamination of the 5′CCC motif (Fig. 1B, 5′C/3′C ratio of 13) and exhibits a processivity factor that is similar to native Apo3G (Fig. 1B, processivity factor of 8 for native and H186R).

To further test the mechanisms of how the Apo3G mutants scan ssDNA, we used a substrate with a dsDNA region located between the two 5′CCC motifs (Fig. 1C). This type of dsDNA acts as an obstacle to Apo3G because the enzyme cannot bind well to dsDNA and is predicted to result in a loss of the sliding component of ssDNA scanning (24, 26). Native Apo3G is still able to processively deaminate two C residues with a dsDNA region in-between, but with a 2-fold decrease in the processivity factor (Fig. 1, B and C, compare processivity factors of 8 and 4), as characterized previously (26). This suggests that the sliding component of native Apo3G has been blocked by the assay conditions and only the jumping component of the native Apo3G scanning mechanism can result in a processive deamination (24, 26). There is no preference for either the 5′- or 3′-proximal CCC motif because the dsDNA region creates about equal lengths of ssDNA regions on the 3′-side of the CCC motifs (Fig. 1C, sketch, 33 or 36 nt), and it is the availability of ssDNA in this region that contributes to the deamination polarity of Apo3G (24, 26). For the F/W mutant, there is almost a complete loss of processivity when the enzyme is confronted with a dsDNA block as indicated by the processivity factor of 1.5 and near absence of a double deamination band (Fig. 1C, F/W mutant 5′C & 3′C band). A nonprocessive enzyme would have a processivity factor of 1 (24). This suggests that there is a deficiency in the ability of the F/W mutant to successfully jump, and the enzyme uses a predominantly sliding mechanism. In contrast, the H186R mutant retains the same processivity factor regardless of whether a dsDNA block is present (Fig. 1, B and C, H186R, processivity factors are 8), which suggests that this mutant uses jumping more frequently than native and F/W mutant Apo3G.

We can estimate the relative distribution of jumping and sliding events that result in a deamination by calculating the
correlated deamination efficiency. This is the efficiency with which Apo3G deaminates a 3′C and moves on to deaminate a 5′C or vice versa (31). For the most efficient deamination (3′ → 5′), native Apo3G has a deamination efficiency of 48% on ssDNA (Fig. 1B, 5′C & 3′C/3′C). On partially dsDNA, this decreases to 29%. Therefore, the contribution of jumping is 60% (0.29/0.48) for native Apo3G (Fig. 1, B and C). Using the same calculation, the jumping frequency for the F/W mutant is only 14% (Fig. 1, B and C). In agreement with the F/W mutant exhibiting a predominantly sliding motion on ssDNA, the enzyme shows a 4-fold higher binding affinity for ssDNA than native Apo3G (Table 1). In contrast, the 3′ → 5′ processive efficiency of H186R is made up of 85% jumping and 15% sliding, which represents an ~1.5-fold increase in jumping, at the expense of sliding, from native Apo3G (Fig. 1, B and C, 5′C & 3′C/3′C). That the H186R mutant and native Apo3G have a similar binding affinity for ssDNA (Table 1) supports the hypothesis that the sliding movements are not decreased with H186R because of a deficiency in ssDNA binding but because of a structural change that has altered the balance between jumping and sliding.

Apo3G-induced Mutation Spectra in the Protease Gene—To examine how jumping and sliding motions affect Apo3G deamination during HIV replication, we used an in vitro model HIV replication system to test the mutagenesis efficiency of native and mutant Apo3G. The in vitro system we developed uses an RNA molecule that is primed for first strand DNA synthesis with a synthetic primer (Fig. 2A). The RNA also has a polyuridine tract (PPT) that enables second strand synthesis to occur (data not shown). To the reaction we added NC, RT, and dNTPs, with or without Apo3G. The RNA produced contains 120 nt of the protease gene from HIV-1 containing the coding region for the active site and flap important for substrate recognition (60, 61). This region in the folded protein is a target for many protease inhibitors (62). Fused to the protease gene is the lacZ reporter sequence that is used to differentiate mutated and nonmutated DNA after ligation of the DNA into a plasmid and transformation of the DNA into E. coli (Fig. 2A). The lacZ region allows screening for mutated clones (white colonies) among the nonmutated clones (blue colonies).

The addition of Apo3G results in an ~16-fold increase in the population mutation frequency (frequency of white colonies) than with RT alone (Table 2, 0.94 versus 0.06). Sequences were analyzed in the HIV genome strand orientation where Apo3G-catalyzed deaminations result in G→A mutations on the coding strand. The mutation pattern of Apo3G in this in vitro system is biased 5′ → 3′ (Fig. 2B) in a manner similar to what has been reported to occur in vivo (20–21). This bias, which can be seen by visual inspection, is due to the 5′ cDNA region remaining single-stranded the longest and may also be enhanced by the 3′ → 5′ deamination bias of Apo3G (20, 21, 24). The distribution of Apo3G-induced mutations continues as a 5′ → 3′ gradient into the lacZ region (supplemental Fig. S1).

There are distinct hot spots for Apo3G deamination in the protease region that occur in over 60% of the clones (Fig. 2B). These are 5′AA CC TC (80% of clones, position 2393) and 5′T CC AA (64% of clones, position 2390). However, other 5′CC sites are not deaminated frequently, likely due to the proximity to the PPT that would result in this ssDNA region being quickly converted to dsDNA, i.e. position 2285–2289 that is immediately downstream of the PPT contains a 5′TCC TTC site that is not deaminated in any of the sequenced clones. However, a similar sequence (5′TCC T) at position 2317 demonstrated that Apo3G can deaminate both C residues in this sequence context with a frequency of 4% (Fig. 2B). Apo3G could induce mutagenesis at numerous sites per clone in the protease gene (Fig. 2C). Most of the mutations occurred at 5′GG or 5′GGG sites (Fig. 2D), which represent the preferred deamination site for Apo3G on the first strand cDNA (5′CC or 5′CCC).

The importance of using the protease gene as a substrate for Apo3G-catalyzed deaminations is that we can assess the level of Apo3G-induced restriction by examining the resulting amino acid changes in the protease-active site (Table 3). Based on a published mutagenesis study of the HIV-1 protease, we are able to infer whether Apo3G-induced mutations cause protease inactivation (Table 3) (63). The mutation spectra agree with genetic changes recovered from patient genomes (64). Notably, in this region of the protease gene, there is a Trp codon (5′TGG). In the cDNA, this is a deamination motif for Apo3G (5′CCA) and will always result in a stop codon (5′TAG) being formed if the second C is deaminated by Apo3G (Table 3, amino acid 42). The Apo3G-catalyzed deaminations resulted in 40% of the clones having a stop codon in the protease gene. Other Apo3G-induced mutations that also resulted in a loss of activity occurred more than 30% of the time at the protease amino acids 40, 51, and 52 (Table 3). However, many of the less preferred sites (~20% mutated clones) resulted in amino acid changes that would preserve protease activity at a comparable level to the wild-type enzyme. In particular, the Asp → Asn mutation at amino acid 30 (Table 3) is a known drug resistance mutation for the protease inhibitor nelfinavir (62) and occurred in 16% of the clones.

Role of Apo3G Oligomerization in Mutation Induction during Reverse Transcription—After establishing the expected mutational spectra in the HIV replication system, we tested the monomeric F/W mutant to determine whether the deficiency in jumping (Fig. 1C) would affect the mutational efficiency of the enzyme in a dynamic system. Based on population mutation frequencies (frequency of white colonies), the F/W monomeric form of Apo3G is 2-fold less able than native Apo3G to induce mutagenesis in the in vitro reverse transcription system (Table 2, 0.94 versus 0.42). In addition, for each clone, the mutational load induced by the F/W mutant in the protease gene was 4-fold less than native Apo3G (compare Fig. 2, B and C, with Fig. 3, A and B, and Table 2). No C targets were mutated more than 20% by the F/W mutant within the sequenced population, and mutations were concentrated nearest the 3′-end of the gene.

### Biochemical Basis of Apo3G Deamination Efficiency

#### TABLE 1

| Apo3G             | ssDNA       | RNA         |
|-------------------|-------------|-------------|
| Native            | 335 ± 16    | 135 ± 3     |
| F/W mutant        | 92 ± 1      | 111 ± 23    |
| H186R             | 334 ± 64    | 43 ± 1      |

#### Apparent dissociation constant of native and mutant Apo3G for ssDNA and RNA

The apparent dissociation constant of native and mutant Apo3G for ssDNA and RNA is shown in Table 1. The H186R mutant has a similar binding affinity for ssDNA (Table 1) compared to the F/W mutant, indicating that the structural change affecting jumping does not significantly affect the binding affinity of Apo3G for ssDNA. This suggests that the observed change in deamination efficiency is due to the decreased jumping frequency in the F/W mutant, rather than a decrease in ssDNA binding.
corresponding to the 5’-region of cDNA that would have remained single-stranded the longest (Fig. 3A). The lacZ region was more heavily mutated than the protease gene, but overall the mutation frequency of the F/W mutant was 2-fold less than the native enzyme (compare supplemental Figs. S1 and S2). There were two clones that had over five mutations in the protease gene region, many of which were closely spaced (Fig. 3C, clone 3 (six mutations) and clone 12 (eight mutations)), suggesting that the F/W mutant was still able to act processively by sliding along ssDNA, in accordance with Fig. 1.

We scored the mutation clustering frequency induced by the F/W mutant and native Apo3G by plotting the number of clones with two or more mutations within a cluster of three to six G residues from eight sites in the protease-lacZ construct (Fig. 4). The F/W mutant induced clustered G→A mutations in a manner similar to native Apo3G (Fig. 4, p value 0.5). However, there is a greater standard deviation in the frequency of closely spaced mutations induced by the F/W mutant (Fig. 4), likely due to the inability to cause multiple clusters of mutations per
Regardless of the ability to slide, the F/W mutant-induced mutational load is less than native Apo3G (Table 2 and compare Figs. 3B and 2C), which suggests that jumping is used in conjunction with sliding for native Apo3G.

TABLE 3
Amino acid changes to the protease enzyme resulting from Apo3G-catalyzed deaminations

The phenotypic change to the enzyme was determined from mutagenesis studies on the protease enzyme in the literature (63). The term Active refers to the enzyme having wild-type levels of protease activity as defined by Loeb et al. (63).

| Protase amino acid position | Nucleotide change (underlined) | Amino acid change | Protease enzyme activity | Mutated clones (%) |
|-----------------------------|--------------------------------|-------------------|--------------------------|-------------------|
|                             | GAT → ΔAT                      | D → N             | Active                   | F228A/W127A H186R |
| 25                          | GGA → ΔA                       | G → K             | No activity              | 16                |
| 27                          | GGA → AA                       | G → R             | No activity              | 0                 |
| 29                          | GAT → ΔAT                      | D → N             | Active                   | 4                 |
| 30                          | GAT → ΔAT                      | D → N             | Active                   | 0                 |
| 34                          | GAA → ΔA                       | G → R             | No activity              | 4                 |
| 38                          | TGG → TAA                       | L ← I             | Active                   | 8                 |
| 40                          | GGG → ΔGG                       | G → R             | No activity              | 4                 |
| 42                          | GGI → ΔG                       | G → R             | No activity              | 0                 |
| 48                          | GGG → ΔG                       | G → R             | Active                   | 4                 |
| 49                          | GQA → ΔA                       | G → R             | No activity              | 0                 |
| 51                          | GQA → ΔA                       | G → R             | No activity              | 0                 |
| 52                          | GGT → ΔGT                      | G → S             | No activity              | 80                |

Because the F/W mutant binds the protease RNA with the same affinity as native Apo3G (Table 1), it is unlikely that the RNA present in the model HIV replication assay could trap the F/W mutant or inhibit its activity. In a separate deamination experiment with synthetic ssDNA, we found that the substrate usage and processivity of the F/W mutant are not affected by the presence of equimolar competitor RNA trap (data not shown). Rather, these data (Figs. 3 and 4) are consistent with Fig. 1, B and C, and suggest that the use of sliding alone as a DNA scanning mechanism is detrimental to the mutational efficiency of Apo3G.

The F/W mutant still retains a preference for deamination of 5’CC or 5’CC motifs, but the hot spots of the mutant differ from that of native Apo3G (Table 4 and compare Figs. 2B and 3A). The change in the deamination hot spots alters the frequency of amino acid changes in the protease region (Table 3). The most highly mutated site (20% of clones) due to F/W mutant deamination is at amino acid 42 (Trp → Stop). Amino acid 52, which is the most highly mutated by native Apo3G-catalyzed deaminations (80% of clones), is only mutated in 8% of the clones acted on by the F/W mutant (Table 3). The F/W mutant did not induce mutagenesis at amino acid 30 (Table 3), a drug resistance mutation for nelfinavir (62).

**Clinical H186R Apo3G Mutant**—We investigated how the decrease in DNA scanning by sliding (Fig. 1C) would influence the H186R-induced mutations during HIV replication. The H186R profile of mutation frequency and distribution differed from native Apo3G (Table 4 and compare Figs. 2B and 5A). The H186R mutant had a population mutation frequency (frequency of white colonies) of 0.56, which is ~2-fold less than...
Biochemical Basis of Apo3G Deamination Efficiency

![Graph showing frequency of clustered mutations induced by native and mutant Apo3G.](figure.png)

**FIGURE 4. Frequency of clustered mutations induced by native and mutant Apo3G.** In the 368-nt cDNA produced by reverse transcription of the protease gene and *lacZa* reporter gene, there are eight sites with at least three (and up to six) C residues. Within these sites, the frequency of clustered G→A mutations that would have resulted from Apo3G deamination was scored. **Horizontal bars** represent the average mutation cluster frequency. The data demonstrate that the native Apo3G and F/W mutant induce clustered mutations with a similar frequency (p value 0.528). However, the clustered mutations induced by the H186R mutant was significantly less than native Apo3G (p value 0.003).

| Site (nt) | Native | F/W mutant | H186R |
|----------|--------|------------|-------|
| 2357     | 0.40   | 0.10*      | 0.28* |
| 2358     | 0.12   | 0.20*      | 0.12  |
| 2364     | 0.40   | 0.16*      | 0.20* |
| 2381     | 0.20   | 0.03*      | 0.04* |
| 2383     | 0.16   | 0.05*      | 0.0*  |
| 2384     | 0.08   | 0.08       | 0.04* |
| 2390     | 0.64   | 0.12*      | 0.20* |
| 2391     | 0.16   | 0.04*      | 0.04* |
| 2393     | 0.80   | 0.08*      | 0.32* |

* Significant difference was designated as p ≤ 0.001 versus native Apo3G values.
* Significant difference was designated as p ≤ 0.01 versus native Apo3G values.
* Significant difference was designated as p ≤ 0.05 versus native Apo3G values.

native Apo3G (Table 2). Most frequently, H186R-catalyzed deaminations induced only 1–2 mutations in the protease gene (Fig. 5B), which results in a 2-fold decrease in the clone mutation frequency from native Apo3G (Table 2). The single sites were deaminated at random and clones with multiple mutations did not show clusters of deaminations as did native Apo3G suggesting that the enzyme is not scanning the DNA by sliding (compare Figs. 5C and 2D). This resulted in a significant decrease in clustered deaminations for H186R compared with native Apo3G (Fig. 4, p value 0.003). More H186R-induced mutations were recovered from the *lacZa* region, but the overall number of mutations per clone was ~2-fold less than native Apo3G (compare supplemental Figs. S1 and S3). These differences in the mutational spectra (Table 4) result in different codon changes in the protease region (Table 3). Many of the sites mutated extensively (>40%) from native Apo3G are mutated about equally (~20%) in the presence of H186R (Table 3). All together, these data underline the negative effects of the altered ssDNA scanning properties of the H186R mutant.

**TABLE 4**

| Distribution of mutations in the protease gene |
|------------------------------------------------|
| Site (nt) | Native | F/W mutant | H186R |
|----------|--------|------------|-------|
| 2357     | 0.40   | 0.10*      | 0.28* |
| 2358     | 0.12   | 0.20*      | 0.12  |
| 2364     | 0.40   | 0.16*      | 0.20* |
| 2381     | 0.20   | 0.03*      | 0.04* |
| 2383     | 0.16   | 0.05*      | 0.0*  |
| 2384     | 0.08   | 0.08       | 0.04* |
| 2390     | 0.64   | 0.12*      | 0.20* |
| 2391     | 0.16   | 0.04*      | 0.04* |
| 2393     | 0.80   | 0.08*      | 0.32* |

Rayleigh light scattering chromatograms generated from multicycle light scattering data show that the protein distributions of the monomer and dimer forms are similar for native Apo3G and H186R (Fig. 6, A and B). We also used AFM to demonstrate that H186R can oligomerize on ssDNA substrates in a manner similar to native Apo3G (Fig. 6, C–F). AFM can be used to determine the volume of proteins, which exhibits a linear relationship with molecular weight (53). The Apo3G native and H186R compositions shift from a distribution of monomers and dimers to a broad distribution of oligomers in the presence of ssDNA (Fig. 6, C–F). This is in contrast to the F/W mutant, which does not oligomerize in solution or when bound to DNA (25).

A prediction from the data for H186R is that the mutant is less efficient at deaminating closely spaced residues because this requires a sliding mechanism. As the distance between C residues decreases, the ability for H186R to act processively should decrease. On a ssDNA substrate with two 5’CCC residues spaced 13 nt apart, which is 2-fold less distance than the DNA shown in Fig. 1A (processivity factor 4), the processivity of H186R decreases 2-fold (Fig. 7A, processivity factor 2), whereas the WT processivity remains the same (Fig. 7A, processivity factor 6).

To ensure the RNA present in the HIV replication assay did not influence H186R processivity, we added an equimolar concentration of RNA trap to a deamination reaction with ssDNA containing two 5’CCC motifs spaced 28 nt apart (Fig. 7B). The competitive RNA trap does not decrease the processivity or substrate usage of H186R or native Apo3G (Fig. 7B) even though H186R has an 8-fold higher binding affinity for RNA than DNA (Table 1). This suggests that the jumping motion of H186R is the same as native Apo3G and that the enzyme maintains a close association with the DNA molecule and does not diffuse into the bulk solution (31). As a result, the changes in the mutation spectrum of H186R appear to be due to an intrinsic decrease in sliding from the enzyme as it is scanning DNA (Fig. 1C).

**DISCUSSION**

There has been a lack of understanding of the key mechanistic properties of Apo3G DNA scanning that allows for efficient deamination during HIV replication. To answer this question, we analyzed the Apo3G DNA scanning mechanisms using synthetic DNA substrates and an *in vitro* model HIV replication system. We demonstrated that the monomeric Apo3G F/W mutant cannot effectively deaminate HIV cDNA due to a deficiency in jumping (Figs. 1C and 3). These are the first data to implicate a functionality of oligomers in the mutagenesis process. We found the clinical Apo3G mutant H186R induced less mutagenesis of HIV cDNA due to decreased scanning by sliding (Figs. 1C, 4, and 5). These are the first data to biochemically characterize this clinical mutant and suggest a mechanism for the clinical phenotype (41, 42). Because H186R has the same oligomeric profile as native Apo3G (Fig. 6), the data suggest that oligomerization is necessary but not sufficient to restore the native DNA scanning mechanism. The F/W mutant and H186R induce 4- and 2-fold less mutagenesis per clone than native Apo3G (Table 2), respectively, which results in fewer
codon changes in the protease gene (Table 3). Overall, the data indicate that native Apo3G requires both jumping and sliding modes of DNA scanning (Figs. 1 and 2), which are mediated by the N terminus, for efficient deamination and avoidance of sublethal mutagenesis of HIV.

**Comparing Mutagenesis of the HIV-1 Protease-active Site by Native and Mutant Apo3G**—We identified 20 different mutations that can be introduced into HIV-1 protease (amino acids 25–52) by Apo3G-catalyzed deaminations (Table 3). Of these 20 mutations, 8 result in a catalytically active protease and 12 inactivate the protease, based on mutagenesis studies of HIV-1 protease (63). Key to this observation is that not all these sites are deaminated equally. Specific sites are more highly mutated due to spatial location or the presence of the preferred sequence context for Apo3G deamination, 5/H11032CC or 5/H11032CCC (20). The most highly preferred sites for deamination at protease codon positions 52, 51, and 42 would result in a loss of protease activity (Table 3), suggesting that if Apo3G has access to this region of the protease gene in vivo, the incurred deaminations would result in viral inactivation. Recently, it has been shown that a low level of mutagenesis (sublethal) may assist HIV-1 to evolve and evade host restriction strategies (36, 38). To consider if Apo3G can contribute to sublethal mutagenesis, we examined the clones that had only one mutation in the protease sequence. From the data in Fig. 2, which tested Apo3G at a concentration estimated to be found in virions (47), there were only three clones with inactivating mutations in the protease region (Table 3), suggesting that if Apo3G has access to this region of the protease gene in vivo, the incurred deaminations would result in viral inactivation and were randomly distributed between 4 mutational hot spots (codons 52, 51, 42, and 40) (supplemental Fig. S4). Notably, the one clone that did not inactivate HIV-1 protease was located at codon 30, an Asp → Asn mutation that can result in resistance to the protease inhibitor nelfinavir (62). As a result, these data cannot exclude that Apo3G contributes to sublethal mutagenesis that may give HIV-1 an evolutionary advantage. Apo3G has also been found to induce resistance to 3TC by inducing a Met → Ile mutation in RT (38). However, in the majority of clones, Apo3G was able to introduce multiple inactivating mutations to the protease (Table 3 and Fig. 2C).

The F/W mutant and H186R mutant induced less mutagenesis in the protease region than native Apo3G (compare Figs. 2B, 3A, and 5A and Tables 2 and 4). Although in the protease region the F/W mutant and H186R mutant were able to induce mutations that would inactivate the protease 75% of the time (Table 3), considering the decreased number of mutations in the whole HIV-1 genome, we conclude these mutant forms with altered DNA scanning properties will be less likely to cause viral inactivation simply due the decrease in mutations induced. In agreement with this, Bulliard et al. (28) found that the F126L or W127A mutants were 4- or 7-fold less effective at restricting HIV-1 replication than native Apo3G, even after the mutant deficiencies for packaging into HIV virions were corrected through a temporary HIV-Vpr fusion protein. Although the H186R mutant was found, in a cell-based assay, to inhibit HIV-1 infectivity at a level comparable with native Apo3G, the amount of H186R expressed in the transient cell-based expression system was not quantified (41). As a result, the level of H186R in these experiments could have been high enough to overcome the scanning deficiencies, but in our
experiments a level of H186R was used to mimic concentrations found in an HIV virion in vivo (Fig. 5). Because it has been demonstrated that the number of mutations induced by Apo3G correlates with HIV-1 inactivation (65), it is likely that the decreased ability of H186R to induce mutagenesis of HIV DNA would result in a decrease in its effectiveness in vivo. This hypothesis is consistent with the clinical phenotype of high viral loads and decreased CD4 T-cell counts in HIV-1-infected patients (41, 42), but we cannot exclude that other factors may be involved.

**Structure-based Model for Apo3G DNA Scanning**—The mutated residues of the F/W mutant (F126A/W127A) are located on loop 7 in the CD1 (supplemental Fig. S5), which mediates dimerization of Apo3G (25). The residues Phe-126 and Trp-127 may also interact directly with DNA and stabilize the DNA-protein interaction (25, 28). The data from the F/W mutant, which demonstrate that mutations in Phe-126 and Trp-127 cause a loss of jumping, suggest that an intact loop 7 is needed for efficient scanning and deamination of HIV cDNA (Fig. 3 and Table 2), probably because the enzyme encounters RNA/DNA hybrid regions on the cDNA from incomplete RNA degradation. The equivalent loop in the Apo3G CD2 domain (loop 72) has been shown to contain the determinants for deamination specificity and is thought to mediate an interaction with ssDNA (66–69). Because the CD1 domain has a role in determining the processivity and binding affinity of Apo3G for ssDNA (25), we suggest that loop 7 in the CD1 is a key structure that mediates the processive jumping mechanism of Apo3G.

His-186 is predicted to be located on helix 6 (supplemental Fig. S5). In the CD2 domain, the equivalent helix (helix 62) was
found to influence DNA binding and activity of Apo3G (66). It has been suggested that helix 62 is needed to interact with the negatively charged ssDNA backbone through residues Arg-374 and Arg-376 (66). Because His-186 is located in the center of helix 6 (supplemental Fig. S5), similar to Arg-374 and Arg-376 (66). Because His-186 is located in the center of helix 6 (supplemental Fig. S5), similar to Arg-374 and Arg-376 (66). Because His-186 is located in the center of helix 6 (supplemental Fig. S5), similar to Arg-374 and Arg-376 (66). Because His-186 is located in the center of helix 6 (supplemental Fig. S5), similar to Arg-374 and Arg-376 (66).

ACKNOWLEDGMENTS—We thank Dr. Brian Bandy for the use of the fluorometer, Jason Maley at the Saskatchewan Structural Sciences Centre for assistance with AFM, and Dr. Kerri Kobryn for critical review of the manuscript. The SEC-LX/LV/RI instrumentation was supported by National Institutes of Health Award SI0RR023748-01.

REFERENCES

1. Sheehy, A. M., Gaddis, N. C., Choi, J. D., and Malim, M. H. (2002) Nature 418, 646–650
2. Rehland, E. W., Stenglein, M. D., Shindo, K., Albin, J. S., Brown, W. L., and Harris, R. S. (2010) Nucleic Acids Res. 38, 4274–4284
3. 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) Cell 113, 803–809
4. Mangeat, B., Turelli, P., Caron, G., Friedli, M., Perrin, L., and Trono, D. (2003) Nature 424, 99–103
5. Zhang, H., Yang, B., Pomerantz, R. J., Zhang, C., Arunachalam, S. C., and Gao, L. (2003) Nature 424, 94–98
6. Conticello, S. G., Harris, R. S., and Neuberger, M. S. (2003) Curr. Biol. 13, 2009–2013
7. Marin, M., Rose, K. M., Kozak, S. L., and Kabat, D. (2003) Nat. Med. 9, 1398–1403
8. Sheehy, A. M., Gaddis, N. C., and Malim, M. H. (2003) Nat. Med. 9, 1404–1407
9. Stopak, K., de Noronha, C., Yonemoto, W., and Greene, W. C. (2003) Mol. Cell 12, 591–601
10. Yu, X., Yu, Y., Liu, B., Luo, K., Kong, W., Mao, P., and Yu, X. F. (2003) Science 302, 1056–1060
11. Alce, T. M., and Popik, W. (2004) J. Biol. Chem. 279, 34083–34086
12. Bach, D., Peddi, S., Mangeat, B., Lakkaraju, A., Strub, K., and Trono, D. (2008) Retrovirology 5, 54
13. Bogerd, H. P., and Cullen, B. R. (2008) RNA 14, 1228–1236
14. Burnett, A., and Spearman, P. (2007) J. Virol. 81, 5000–5013
15. Cen, S., Guo, F., Niu, M., Saadatmand, J., Deflassieux, J., and Kleiman, L. (2004) J. Biol. Chem. 279, 33177–33184
16. Drouais, M., Dussart, S., Courcouil, M., Bessou, G., Vigne, R., and Decroly, E. (2004) Biochem. Biophys. Res. Commun. 321, 566–573
17. Khan, M. A., Goila-Gaur, R., Opi, S., Miyagi, E., Takeuchi, H., Kao, S., and Strebel, K. (2007) Retrovirology 4, 48
18. Strebel, K., and Khan, M. A. (2008) Retrovirology 5, 55
19. Svorovskaia, E. S., Xu, H., Mibisa, J. L., Barr, R., Gorelick, R. I., Ono, A., Freed, E. O., Hu, W. S., and Pathak, V. K. (2004) J. Biol. Chem. 279, 35822–35828
