Biochemical Activities of Highly Purified, Catalytically Active Human APOBEC3G: Correlation with Antiviral Effect

Yasumasa Iwatani, Hiroaki Takeuchi, Klaus Strebel, and Judith G. Levin

Laboratory of Molecular Genetics, National Institute of Child Health and Human Development, and Laboratory of Molecular Microbiology, National Institute of Allergy and Infectious Diseases, Bethesda, Maryland 20892

Received 21 December 2005/Accepted 4 April 2006

APOBEC3G (APO3G), a cytidine deaminase with two zinc finger domains, inhibits human immunodeficiency virus type 1 replication in the absence of Vif. Here, we provide a comprehensive molecular analysis of the deaminase and nucleic acid binding activities of human APO3G using a pure system containing only one protein component, i.e., highly purified, catalytically active enzyme expressed in a baculovirus system. We demonstrate that APO3G deaminates cytosines in single-stranded DNA (ssDNA) only, whereas it binds efficiently to ssDNA and ssRNA, about half as well to a DNA/RNA hybrid, and poorly to double-stranded DNA and RNA. In addition, the base specificities for deamination and binding of ssDNA are not correlated. The minimum length required for detection of APO3G binding to an ssDNA oligonucleotide in an electrophoretic mobility shift assay is 16 nucleotides. Interestingly, if nucleocapsid protein and APO3G are present in the same reaction, we find that they do not interfere with each other's binding to RNA and a complex containing the RNA and both proteins is formed. Finally, we also identify the functional activities of each zinc finger domain. Thus, although both zinc finger domains have the ability to bind nucleic acids, the first zinc finger contributes more to binding and APO3G encapsidation into virions than finger two. In contrast, deamination is associated exclusively with the second zinc finger. Moreover, zinc finger two is more important than finger one for the antiviral effect, demonstrating a correlation between deaminase and antiviral activities.

Human immunodeficiency virus type 1 (HIV-1) requires a virus-encoded gene product, Vif (originally called “Sor” or “A”) (8, 43), to replicate efficiently in primary human lymphocytes, monocytes, and certain T-cell lines (e.g., H9), which are commonly designated “nonpermissive cells.” In contrast, Vif is dispensable for replication in other T cells (e.g., SupT1), which are known as “permissive cells” (10, 36, 47). In early work on Vif, vif-deficient virions produced in nonpermissive cells were found to be significantly impaired in their ability to complete reverse transcription (12, 41, 47) as well as 100- to 1,000-fold less infectious than the wild type (WT) (8, 9, 43).

More recently, Sheehy et al. (38) identified APOBEC3G (APO3G), originally termed CEM15, as the cellular factor that restricts replication of vif-deficient HIV-1. APO3G expressed in nonpermissive cells is incorporated into vif-deficient virus particles, but its presence in WT virions is dramatically reduced (15, 20, 26, 27). In nonpermissive cells, Vif counteracts APO3G activity by targeting the inhibitor for degradation via the ubiquitination-proteosome pathway (5, 27, 28, 39, 42, 52), although other mechanisms may also be involved (20, 42).

APO3G is a member of a large family of cellular cytidine deaminases, which includes APOBEC1 (APO1) and activation-induced cytidine deaminase (AID) (19, 48). All of the APO proteins have one or two zinc finger domains containing the conserved motif (H/C)X23E(X)23-25CXXC (3, 48). APO3G has two zinc finger domains, which are important for the biochemical and biological activities of the protein. The APO enzymes catalyze deamination of cytosine residues in single-stranded DNA (ssDNA) and/or RNA to uracil. APO3G preferentially deaminates the terminal cytidine residue of the (T/C)CC sequence in ssDNA (2, 15, 24, 26, 44, 51, 54). This activity has the potential to greatly inhibit virus replication (6, 7, 11, 16, 21).

Until now, most of the studies on APO3G have been performed primarily in cell-based systems or with enzyme derived from virus lysates. To analyze the molecular properties of this protein and to investigate the mechanism of antiviral activity, we chose a more biochemical approach and set out to obtain highly purified, active enzyme. In earlier work it was reported that partially purified APO3G expressed in Escherichia coli has deaminase activity (15) (also see reference 2). However, at about the same time, another group was unable to detect this activity with a bacterially derived enzyme (54). Moreover, in our experience, we found that despite extensive efforts, purification of human APO3G expressed in E. coli failed to yield soluble, active protein, apparently because it does not fold properly (data not shown). Instead, we used a baculovirus expression system, and under these conditions we succeeded in producing enzymatically active glutathione S-transferase (GST)-tagged APO3G as well as untagged APO3G, which could each be purified to ~95%. Partially purified APO3G expressed as a GST fusion protein in baculovirus has been used in two previous studies (19, 44).

The availability of highly purified enzymatically active APO3G has made it possible for the first time to conduct a comprehensive molecular analysis of APO3G deamination and nucleic acid binding activities in the absence of other viral components that are present in viral lysates and without contaminating enzymes from the expression system. In this study, we show that while ssDNA is the exclusive substrate for deami-
nation, APO3G binds efficiently to both ssDNA and ssRNA, approximately half as well to a DNA/RNA hybrid, and poorly to double-stranded DNA (dsDNA) and dsRNA. We also report that the nucleotide specificities for deamination and binding of ssDNA are not correlated. Finally, using purified wild-type (WT) and zinc finger mutant proteins, we demonstrate that the two zinc finger domains play differing roles with respect to nucleic acid binding, deamination, and antiviral activity and speculate on the implications of our findings for the mechanism of the antiviral effect.

MATERIALS AND METHODS

Oligonucleotides. Oligonucleotides (oligos) used in this study were purchased from Integrated DNA Technologies or Lofstrand and are listed in Table 1 according to their use in particular experiments. For electrophoretic mobility shift assays (EMSAs) and deamination assays, DNA and RNA oligos were first gel purified by denaturing polyacrylamide gel electrophoresis (PAGE). Where indicated, oligos were 5'-P-[γ-32P]ATP (3,000 Ci/mmol) (GE Healthcare) using T4 polynucleotide kinase (Ambion), as described previously (13).

Plasmid construction. To construct a plasmid for APO3G expression in a baculovirus system, a DNA fragment of the complete APO3G open reading frame was amplified by PCR from pcDNA-APO3G (20), using primers JL650 and JL629, and was then inserted into the pET41a (+) vector (Novagen) between the PshAI and Xhol restriction sites. The resulting plasmid is called pET-GST-APO3G-WT. An XbaI and Xhol fragment was cloned into pFastBac1 (Invitrogen) and named pFastBac-GST-APO3G-WT. For construction of zinc finger mutants with a Cys-to-Ser change, two DNA fragments, each with a single mutation, were made by PCR using the following primers: for C100S, JL667 and JL668; for C291S, JL648 and JL649. The DNA fragments were inserted into the pFastBac-GST-APO3G-C100S, pFastBac-GST-APO3G-C291S, and pFastBac-GST-APO3G-C100S/C291S were inserted into the pDNA-APO3G-WT previously cut with AflIII and Xhol. In all cases, the sequences of both the insert and the boundary regions were verified by DNA sequencing by the University of Maryland Biopolymers Laboratory.

Purification of human APO3G and mutant proteins. The pFastBac-GST-APO3G plasmids, containing either WT or zinc finger mutations, were used to make recombinant bacmid DNA in E. coli DH10Bac, according to the protocol supplied by the manufacturer for the Bac-to-Bac Baculovirus Expression System (Invitrogen). To obtain recombinant viruses, Sf21 cells were transfected with each bacmid DNA by using Cellfectin (Invitrogen); high-titer viruses were isolated by centrifugation at 20,000 × g for 30 min and were bound to glutathione Sepharose High Performance resin (GE Healthcare) for 3 h at 4°C. The cells were disrupted by sonication in lysis buffer (25 mM HEPES [pH 7.0], 500 mM NaCl, 1% Triton X-100, 10 mM CaCl2, 1 mM EDTA, 5 mM 2-mercaptoethanol [2-ME], 10% glycerol, 1 mM phenylmethylsulfonyl fluoride, complete protease inhibitor [Roche]). DNase I (5 U/ml) and RNase A (40 μg/ml) were added, and the entire mixture was kept on ice for 1 h. The soluble fraction was isolated by centrifugation at 20,000 × g for 30 min and was bound to glutathione Sepharose High Performance resin (GE Healthcare) for 3 h at 4°C.

To facilitate dissociation of residual nucleic acids from the protein, the resin was first washed with lysis buffer containing 1 M NaCl without glycerol or protease inhibitors and then with lysis buffer containing 500 mM NaCl without protease inhibitors. The bound GST-APO3G proteins were eluted with glutathione buffer (50 mM Tris-HCl [pH 7.4], 10 mM NaCl, 40 mM reduced glutathione, 5% glycerol, 10 mM ZnCl2), and the eluate was dialyzed against Enterokinase buffer (Novagen). To obtain APO3G without a GST tag, the dialyzed solution was diluted (so that the APO3G concentration was about 10 μM) and then incubated in the presence of recombinant Enterokinase (5 U/ml) at 21°C for 12 h. Although most of the cleaved APO3G was precipitated, only the soluble fraction (approximately 10 to 30% of the total protein) was collected for the next step. In contrast to APO3G, GST-tagged APO3G proteins remained highly

**TABLE 1. Oligonucleotides used in this study**

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence*</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>JL650</td>
<td>5'-GAT GAA GCC TCA TCT CAG AAA CAC AGT G-3'</td>
<td>Construction of pET-GST-APO3G WT</td>
</tr>
<tr>
<td>JL629</td>
<td>5'-CCA TGA CTC GAG TTA CTA GGT TTC CTG ATT CTG GAG AAT GGC-3'</td>
<td>Construction of pET-GST-APO3G WT</td>
</tr>
<tr>
<td>JL667</td>
<td>5'-AGC CCC TGC ACA AAG TCG ACA AGG GAT ATG-3'</td>
<td>Construction of C100S mutation</td>
</tr>
<tr>
<td>JL668</td>
<td>5'-CAT ATC CCT TGT CGA CTT TGT GCA GCA GGC GCT-3'</td>
<td>Construction of C100S mutation</td>
</tr>
<tr>
<td>JL648</td>
<td>5'-AGC CCC TGC TTC AGC AGT GCA CAG GAA ATG GCT-3'</td>
<td>Construction of C291S mutation</td>
</tr>
<tr>
<td>JL649</td>
<td>5'-AGC CAT TTC TGC ACT GCT GAA GCA GGG GCT-3'</td>
<td>Construction of C291S mutation</td>
</tr>
<tr>
<td>JL650</td>
<td>5'-AAA AGA AGA GGG GCT ACC GCT-3'</td>
<td>Deamination assay</td>
</tr>
<tr>
<td>JL651</td>
<td>5'-ATT ATT ATT ATT ATT ATT ATT CUG CGG ATT TAT TTA TTT A-3'</td>
<td>Deamination assay</td>
</tr>
<tr>
<td>JL652</td>
<td>5'-ATT ATT ATT ATT ATT ATT ATT CCC AGG ATT TAT TTA TTT A-3'</td>
<td>Deamination assay</td>
</tr>
<tr>
<td>JL653</td>
<td>5'-TAA ATA AAT AAA TTC-3'</td>
<td>Deamination assay</td>
</tr>
<tr>
<td>JL654</td>
<td>5'-auu auu auu auu auu auu ccg auu uau uaa uuu a-3'</td>
<td>Deamination assay and EMSA</td>
</tr>
<tr>
<td>JL655</td>
<td>5'-aaa aga aaa ggg ggg acu gg-3</td>
<td>Deamination assay and EMSA</td>
</tr>
<tr>
<td>JL656</td>
<td>5'-AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA A-3'</td>
<td>Deamination assay and EMSA</td>
</tr>
<tr>
<td>JL657</td>
<td>5'-AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA A-3'</td>
<td>Deamination assay and EMSA</td>
</tr>
<tr>
<td>JL658</td>
<td>5'-TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT-3'</td>
<td>Deamination assay and EMSA</td>
</tr>
<tr>
<td>JL659</td>
<td>5'-CCC CCC CCC CCC CCC CCC CCC CCC CCC CCC CCC CCC C-3'</td>
<td>Deamination assay and EMSA</td>
</tr>
<tr>
<td>JL660</td>
<td>5'-UUU UUU UUU UUU UUU UUU UUU UUU UUU UUU UUU UUU UUU UUU UUU UUU-3'</td>
<td>Deamination assay and EMSA</td>
</tr>
<tr>
<td>JL670</td>
<td>5'-TAT TAT TAT TAT CCC GGG CAT GAT TTA-3'</td>
<td>Deamination assay and EMSA</td>
</tr>
<tr>
<td>JL671</td>
<td>5'-TTA TTA TTC CGC GGA TTT ATT TTT-3'</td>
<td>Deamination assay and EMSA</td>
</tr>
<tr>
<td>JL672</td>
<td>5'-TTA TTA TTC CGC GGA TTT ATT TTT-3'</td>
<td>Deamination assay and EMSA</td>
</tr>
<tr>
<td>JL673</td>
<td>5'-TAT GCC GGG CAT TT-3'</td>
<td>Deamination assay and EMSA</td>
</tr>
<tr>
<td>JL674</td>
<td>5'-TAT GCC GGG CAT TT-3'</td>
<td>Deamination assay and EMSA</td>
</tr>
<tr>
<td>JL675</td>
<td>5'-aaa auu uau gac ugg cua gaa cgg gag-3'</td>
<td>Deamination assay and EMSA</td>
</tr>
</tbody>
</table>

* Uppercase and lowercase denote deoxyribo- and ribonucleotides, respectively. Note that the uppercase U is a deoxyuridine. The “p” and “(mC)” represent monophosphorylated modification at the 5’ end and 5-methyldeoxyuridine, respectively.
soluble, APO3G and the GST-tagged proteins were subjected to fractionation on a high-performance liquid chromatography-DEAE column (10 μm; 75 by 7.5 mm; Tosoh Bioscience), using a linear gradient from 20 to 1,000 mM NaCl with a flow rate of 1 ml/min over 50 min. The solutions used to make the gradient were Buffer A (50 mM Tris-HCl [pH 7.4], 20 mM NaCl, 10% glycerol, 0.5% Triton X-100, 1 mM 2-ME) and Buffer B, which is identical to Buffer A except that the NaCl concentration is 1,000 mM. APO3G was eluted after ~21 min. Protein concentrations were determined using a Coomassie Plus kit (Pierce). Note that the legends specify whether untagged APO3G (referred to as “APO3G”) or tagged APO3G (referred to as “GST-APO3G”) was used in a particular experiment.

Cytidine deamination assays. Two different methods were used to detect deaminase activity of APO3G: (i) a UDG-dependent assay and (ii) primer extension terminated with a dideoxynucleotide.

(i) UDQ-dependent assay. The 32P-labeled substrate (30 nM) was incubated in deamination buffer (10 mM Tris-HCl [pH 8.0], 50 mM NaCl, 1 mM EDTA, 1 mM dithiothreitol [DTT]) at 37°C for 90 min in the presence or absence of APO3G (or GST-APO3G) (0.4 to 1.2 μM), UDQ from E. coli (2 μL) (New England Biolabs) and UDQ buffer were added to the deamination reaction mixture, which was incubated at 37°C for 45 min. The reaction mixture was treated with 0.15 M NaOH at 37°C for 20 min, heated at 95°C for 5 min, and then promptly chilled on ice. After separation of the cleavage products by PAGE in a 15% denaturing gel, the radioactivity in the gel was quantified by scanning with a PhosphorImager (GE Healthcare) in the linear range of band intensity, followed by analysis with ImageQuant software.

(ii) Primer extension terminated with a dideoxynucleotide. As described above for the UDQ-dependent protocol, the substrate (30 nM) was incubated in deamination buffer at 37°C for 90 min in the presence or absence of APO3G (or GST-APO3G) (1.2 μM). The deaminated substrate was annealed to a 32P-5′-end-labeled primer (45 nM) (JL665) by heating at 85°C for 5 min and then gradually cooling to 25°C. An equal volume of primer extension cocktail (100 mM Tris-HCl [pH 8.0], 150 mM KCl, 14 mM MgCl2, 2 mM DTT, 1.0 U/μl of SUPERaseIn [Ambion]), 20 mM HIV-1 reverse transcriptase (RT) (Workington), 100 μM dideoxy ATP, and 1.2 μM each of dTP and dGTP was added to the annealed product, and the mixture was incubated at 37°C for 15 min. Reactions were terminated by freezing on dry ice, followed by addition of gel loading buffer II (Ambion). The products were separated by PAGE in a 15% denaturing gel. Radioactivity was quantified as described above.

Electrophoretic mobility shift assays (EMSAs) with APO3G. For formation of a duplex, the labeled oligo was heat annealed to its complement at a ratio of unlabeled to labeled oligo of 1.5:1. The labeled oligo (30 nM) was incubated with the indicated concentration of purified proteins in gel-shift reaction buffer (50 mM Tris-HCl [pH 8.0], 75 mM KCl, 7 mM MgCl2, 1 mM DTT) at 37°C for 15 min. The resulting complexes were separated in native gels (4 or 6%) at 4°C. Radioactivity in the gel was quantified as described above. The percentage of substrate that was shifted (% Shifted) was calculated by dividing the volume of the shifted bands by the total volume of each lane and then multiplying by 100.

Deamination constant values (KD) were estimated by plotting binding activity versus concentration of APO3G from EMSA results with a 20-nucleotide (nt) ssDNA, since essentially one shifted band is formed in this case. Briefly, the KD is defined as [P][A·P]/[A·P], where [A·P] is the APO3G protein, P is the oligo, and A·P is the APO3G-oligo complex. The total oligo concentration (PT) is defined as [P] = [A·P] + [P]. If the ratio of shifted bands, B, is defined as B = [A·P]/[PT], then by manipulating the dissociation constant equations, KD can be calculated as follows: KD = [A·P]/B = 1.

The HIV-1 nucleocapsid protein (NC) used in this work was the generous gift of Robert J. Gereick (SAIC-Frederick, Inc., Frederick, Md). The protein was prepared as described previously (50).

Virus production and analysis of virus infectivity. Virus production and analysis of virus infectivity were performed as reported previously (46). Briefly, to obtain virus particles, APO3G-negative HeLa cells were cotransfected with the vif-defective full-length molecular clone HIV-1 pNL4-3vif(−) and pcDNA3 APO3G or pcDNA3.1 (vector control). Virus infectivity was determined by single-cycle replication assays with LaSIV cells, obtained from the AIDS Research and Reference Reagent Program, National Institute of Allergy and Infectious Diseases, National Institutes of Health (catalog no. 5460); cells were originally obtained from J. W. Roos and J. E. Clements) (34). Infectivity was calculated by normalizing for the amount of input RT activity.

Incorporation of APO3G into virions. For immunoblot analysis of APO3G present in the cell and in virions, the following procedure was used. Cells were washed once with phosphate-buffered saline (PBS), suspended in PBS (200 μl per 5 × 106 cells), and then mixed with an equal volume of 2 × Laemmli buffer (Bio-Rad) containing 5% 2-ME. To prepare virus-associated proteins, cell-free-filtered supernatants from transfected HeLa cells (7 ml) were pelleted (75 min, 151,000 × g) through a 20% sucrose cushion (4 ml) in an SW41 rotor. The concentrated virus pellet was suspended in PBS (50 μl) and was mixed with an equal volume of 2× Laemmli buffer. Proteins were solubilized by heating for 10 to 15 min at 95°C. Cell lysates were normalized according to cell number, and virus lysates were normalized according to RT activity. Samples were subjected to sodium dodecyl sulfate (SDS)-PAGE; proteins were transferred to polyvinylidene difluoride membranes and reacted with anti-APO3G rabbit serum (20). To detect cap3 (CA) protein, anti-CA rabbit serum (obtained from the AIDS Research and Reference Reagent Program, National Institute of Allergy and Infectious Diseases, National Institutes of Health; catalog no. 4250) was used. The membranes were then incubated with alkaline phosphatase-conjugated secondary antibodies (Applied Biosystems), and the proteins were visualized by the Western-Star system (Applied Biosystems).

RESULTS

Enzymatic properties of purified APO3G expressed in a baculovirus system. To perform a detailed analysis of the biochemical properties of APO3G, the protein was expressed with an N-terminal GST tag in a baculovirus system and purified. Figure 1 shows the migration of purified GST-APO3G (~78 kDa) in an SDS–10% polyacrylamide gel (lane 2). Since the protein solubility is low, APO3G precipitated once the tag was removed. However, there was still sufficient soluble protein to allow further purification (lane 3). Based on Coomassie staining of the gel, the purity of APO3G (~46 kDa) (lane 3) and GST-tagged APO3G (lane 2) was estimated to be ~95%. The cytidine deaminase activity of APO3G was tested using a UDQ-dependent assay. To determine substrate specificity, we prepared ssDNA, dsDNA, and a DNA/RNA hybrid, each containing the same minus-strand DNA sequences (Fig. 2A). There are three regions containing potential deamination sites, which we designated regions I, II, and III (Fig. 2A). In the absence of HIV-1 NC (Fig. 2B), APO3G deaminated cytidines in all three of these regions in ssDNA (lane 4). This indicates that both CCC and TCC represent deamination motifs recognized by APO3G, in agreement with earlier results (15, 24, 26, 44, 51, 54). No deamination was detected in the absence of APO3G (lanes 1 to 3).

In contrast to the results with ssDNA, DNAs annealed to a complementary 20-nt DNA (lane 5) or RNA (lane 6) could not

FIG. 1. Purification of APO3G from a baculovirus expression system. Purified GST-APO3G and APO3G proteins were resolved by electrophoresis in an SDS–10% polyacrylamide gel and stained with Coomassie brilliant blue R-250. Lanes: 1, prestained SDS-PAGE standard (low range) molecular marker (Bio-Rad); 2, 5 μg of GST-APO3G; and 3, 3 μg of APO3G. Approximate molecular sizes are indicated on the left.
be deaminated in region II or III, which forms a dsDNA or a DNA/RNA hybrid, respectively. However, the dCs in single-stranded region I were deaminated, forming the same region I products as ssDNA. These results demonstrate that purified APO3G has single-strand-specific DNA deaminase activity. Although HIV-1 NC is a required cofactor for reverse transcription (25), the deamination pattern and efficiency of APO3G with these substrates were unaffected by the presence of NC (lanes 10 to 12). This suggests that NC does not interfere with the deamination activity of APO3G. The results also imply that ssDNA remains a target for APO3G even when the DNA is bound to NC.

Since the cytidine residue in the CpG sequence is frequently methylated in the eukaryote genome (33), it was also of interest to determine whether APO3G can deaminate dC with a methyl group at position 5, i.e., 5-methyldeoxycytidine (5mC). For this purpose, we developed an alternative deaminase assay in which primer extension is terminated by a dideoxynucleotide. We devised this assay because the putative deamination product of 5mC is thymidine, which is not a substrate for UDG.

Figure 2C demonstrates that in reactions containing the dC sample (positive control), the dideoxy ATP (ddATP) mixture, and GST-APO3G, a strong H1 band was generated in addition to a H5 product (lane 1). This indicates that deamination of dC to dU occurred predominantly at position 2. The appearance of a H5 product shows that a portion of the substrate was not deaminated, and therefore some extension could continue until the T residue at position 5 (lane 1). The +1 band appeared in the presence or absence of APO3G and is likely to be a pause product (lanes 1 and 2). In the absence of GST-APO3G, the predominant product was a H5 band (lane 2). In contrast, with the methylated substrate a H5 band was the major product detected in either the presence or absence of APO3G; the +1 pause product was also observed (lanes 3 and 4). Thus, there was no evidence for deamination of 5mC residues. The same conclusion was reached when the experiment was repeated with a dideoxy GTP (ddGTP) mixture (with dTTP and dATP) (data not shown).

Additionally, we found that methylation blocked deamination of an ssDNA containing the sequence TC(5mC)G (data not shown). Thus, if single-stranded regions of genomic DNA containing methylated CpG were to come in contact with APO3G, these DNA regions would be protected from APO3G deaminase activity.

We also investigated whether cytosines in ssRNA can be deaminated by APO3G (Fig. 2C). Since UDG does not react with uracils in RNA, our approach was to use the primer extension assay to address this question. As observed with the mC substrate, a H+2 termination product could not be detected with the ddATP mixture in the presence (lane 5) or absence (lane 6) of APO3G. Taken together, these results indicate that APO3G cannot significantly deaminate ssRNA in this system.

Characteristics of APO3G binding to nucleic acids. Detailed analysis of APO3G nucleic acid binding activity has not been
reported thus far. Using purified APO3G, we initially com-
pared the binding affinity of APO3G to DNA and RNA by
EMSA. As shown in Fig. 3A, as the concentration of APO3G
was raised an increase in the number of high-molecular-weight
complexes and a greater proportion of shifted material could
be detected; at the highest concentration, the percent shifted
value reached 91% for both the RNA and DNA samples. A
slightly higher amount of shifted bands was observed with the
40-nt RNA compared with the 40-nt DNA at lower APO3G
concentrations. However, these differences were small, sug-
gesting that the binding affinities of APO3G to RNA and DNA
were actually quite similar. Note that in each case, with high
concentrations of APO3G, four shifted bands with different
mobilities could be detected. We assume that these bands
represent four different nucleic acid-APO3G complexes. Inter-
restingly, in other EMSA experiments, APO3G bound with
similar efficiency to 33-nt DNA or RNA oligos containing the
biologically relevant HIV-1 SL-3 stem-loop sequence, which in
the RNA form represents the major viral RNA packaging
signal and is located at the 5’ end of the genome (35) (data not
shown).

To determine whether binding is specific for single-stranded
nucleic acids, EMSA was also performed with dsDNA, a DNA/
RNA hybrid, and dsRNA (Fig. 3B). These oligos had the same
DNA or RNA sequence as those used in Fig. 3A, in addition to
the appropriate 40-nt complement. The results showed that at
the highest concentration of APO3G (160 nM), the protein
bound approximately four- to fivefold more efficiently to the
dNA/RNA hybrid (55% shifted) (lane 10) than to either
dsDNA (lane 5) or dsRNA (lane 15) (13% and 11% shifted,
respectively). These data indicate that although APO3G bind-
ing to the DNA/RNA hybrid was significant, it was still ap-
proximately twofold lower than that observed with ssRNA or
ssDNA at comparable concentrations of APO3G (compare
Fig. 3A with B).

We also examined the nucleotide preference for APO3G
binding using 40-nt homopolymeric DNA oligos (Fig. 3C). The
data indicated that APO3G could bind efficiently to dT40 or
dU40 (lanes 15 and 20, 87% and 80% shifted, respectively, at
the highest APO3G concentrations) but with less efficiency to
dC40 (lane 10, 20% shifted) and dA40 (lane 5,
10% shifted).
(Note that we were unable to analyze dG40 because of its low
solubility.) Surprisingly, these results indicated that the nucle-
otide preferences for APO3G binding (dT or dU) differ from
the requirement for dC-containing motifs for deamination
(Fig. 2B). It is also of interest that at the highest APO3G
concentration, four different shifted bands were observed, as in
Fig. 3A. This finding demonstrates that the number of shifted
bands is independent of the oligo sequence.

The data of Fig. 3 showing formation of a greater number of
higher order complexes with increasing APO3G concen-
trations also suggest the possibility that when APO3G binds, it

![FIG. 3. Nucleic acid binding properties of APO3G analyzed by
EMSA. In all of the gels, the values for “% Shifted” (%) were calcu-
lated as described in the text. (A) The binding affinities of APO3G for
ssDNA (JL640) and ssRNA (JL654) were determined by EMSA. Forty-
nucleotide DNA and RNA oligos having the same sequence were used.
The APO3G concentrations were the following: lanes 1 and 6, 0 nM;
lanes 2 and 7, 20 nM; lanes 3 and 8, 40 nM; lanes 4 and 9, 80 nM; and
lanes 5 and 10, 160 nM. (B) The binding affinity of APO3G for
double-stranded oligos was tested with dsDNA (JL640/JL652), a
DNA/RNA hybrid (JL640/JL655), and dsRNA (JL654/JL655). One
oligo in each pair was radiolabeled: JL640 and JL654. The APO3G
concentrations for each set (lanes 1 to 5, 6 to 10, and 11 to 15) were the
same as the concentrations given for lanes 1 to 5 in the legend to panel
A. (C) The sequence specificity of APO3G binding was investigated with
the following 40-nt homopolymeric DNAs: dA40, JL657; dC40,
JL659; dT40, JL658; and dU40, JL660. The APO3G concentrations for
each set (lanes 1 to 5, 6 to 10, 11 to 15, and 16 to 20) were the same
as the concentrations given for lanes 1 to 5 in the legend to Fig. 3A. In
this figure as well as in succeeding figures, gel images from EMSA
reactions run on the same gel are enclosed within a box.)
covers a certain length of nucleic acid. To investigate the relationship between nucleotide length and APO3G binding, we measured binding to DNA oligos ranging in size from 14 to 40 nt (Fig. 4). When 24- and 22-nt DNAs were used, two shifted bands were detected with high concentrations of APO3G (lanes 9 and 10 and lanes 14 and 15, respectively), although the intensity of the uppermost band formed with the 22-nt DNA was significantly reduced compared with the analogous band seen with the 24-nt DNA. With the 20-, 18-, or 16-nt DNAs (lanes 19 and 20, 24 and 25, and 29 and 30, respectively), there was essentially only one shifted band. In contrast, a distinct shifted band could not be observed with the 14-nt DNA (lanes 32 to 35). These results demonstrate that APO3G binds nucleic acids that are 16 nt or greater in length.

It seems reasonable to assume that when there is only a single shifted band, it may consist of a single uniform nucleic acid-APO3G complex. Using data from EMSAs with the 20-nt single shifted band, it may consist of a single uniform nucleic acid binding activity of APO3G, EMSA was performed with increasing concentrations of APO3G (lanes 9 and 10 and lanes 14 and 15, respectively), although the intensity of the uppermost band formed with the 22-nt DNA was significantly reduced compared with the analogous band seen with the 24-nt DNA. With the 20-, 18-, or 16-nt DNAs (lanes 19 and 20, 24 and 25, and 29 and 30, respectively), there was essentially only one shifted band. In contrast, a distinct shifted band could not be observed with the 14-nt DNA (lanes 32 to 35). These results demonstrate that APO3G binds nucleic acids that are 16 nt or greater in length.

In Fig. 5B, the reverse experiment was performed, i.e., a fixed amount of APO3G was bound to the DNA before addition of increasing amounts of NC (up to 1.75 nt/NC). At a given concentration of APO3G (20 or 80 nM), the relative amounts of bands shifted by APO3G were independent of NC concentration (lanes 6 to 15). Moreover, as seen in Fig. 5A, when NC and APO3G were both present in the reaction, the RNA bands were “super” shifted. This was especially evident with 80 nM APO3G (lanes 11 to 15). These results indicate that APO3G did not affect NC binding to the RNA oligo.

Interestingly, when the 33-nt SL-3 RNA sequence was used in the same assay, somewhat different results were obtained. In Fig. 5C, NC was bound to the RNA prior to addition of APO3G (lanes 6 to 10). As may be seen, SL-3 RNA was shifted more efficiently in reactions containing NC compared to reactions without NC (compare lanes 7 to 10 with lanes 2 to 5). This effect was most noticeable at the highest APO3G concentrations (compare lanes 9 and 10 with lanes 4 and 5). A similar phenomenon was observed with reactions in which the RNA was preincubated with APO3G before addition of NC (Fig. 5D). Thus, when fixed concentrations of APO3G and increasing concentrations of NC were added (lanes 7 to 10 and 12 to 15), there was a corresponding increase in the amount of the slowest migrating band. The increased binding efficiency was particularly striking when the pattern of shifted bands was compared at the 160 nM concentration of APO3G (lanes 11 to 15). In this case, the faster migrating bands seen in the reaction without NC (lane 11) were almost completely absent in the reactions with the highest concentrations of NC (lanes 14 and 15). The results with SL-3 RNA also suggest that the shifted bands contained both NC and APO3G.

Taken together, the data in Fig. 5 demonstrate that HIV-1 NC and APO3G do not interfere with each other’s binding to RNA and in fact form a complex containing RNA and both proteins. In addition, with SL-3 RNA, binding efficiency of APO3G was augmented when NC was present (see below).

Roles of the zinc finger domains in APO3G nucleic acid binding in vitro and its incorporation into virus particles. To assess the contribution of the zinc finger domains to the nucleic acid binding activity of APO3G, we first performed an EMSA in which APO3G was bound to ssDNA in the presence of the zinc chelating agent 1,10 phenanthroline. The results showed that zinc coordination is essential for APO3G binding to nucleic acids (data not shown). To investigate the function of each individual domain, we purified GST-APO3G containing the following zinc finger mutations with a Cys-to-Ser change:
first domain, C100S; second domain, C291S; and a double mutant, C100S/C291S.

The binding affinities at increasing protein concentrations of wild-type (WT) GST-APO3G and the three mutants were analyzed by EMSA (Fig. 6A) using a 40-nt ssDNA. The results showed that although the values for percentage of ssDNA shifted with WT (lanes 2 to 5) or C291S (lanes 12 to 15) were similar, at comparable concentrations binding activity of C100S (lanes 7 to 10) was reduced by approximately twofold. A drastic loss of binding activity was observed with the double mutant, C100S/C291S (lanes 17 to 20). Similar findings were made in EMSAs with a 20-nt ssDNA, and the $K_D$ values were estimated as follows: WT, $\sim 140$ nM; C100S, $\sim 320$ nM; C291S, $\sim 130$ nM; and C100S/C291S, $\sim 55$ $\mu$M. Note that the WT $K_D$ value for GST-APO3G was similar (approximately twofold higher) to that obtained for untagged APO3G with a 20-nt ssDNA ($\sim 76$ nM). Taken together, these findings indicate that both zinc fingers have the capacity to bind nucleic acid, illustrating the redundancy in zinc finger function with respect to binding. However, the contribution of the first zinc finger domain is more significant than that of the second domain.

To determine whether the binding capacity of APO3G WT and zinc finger mutants is correlated with the ability to encapsidate APO3G, we measured the activity of the purified WT and zinc finger mutant GST-APO3G proteins directly, using the UDG assay. As shown in Fig. 7A, WT and C100S had very similar levels of activity (lanes 1 and 2), whereas C291S and the double mutant had no detectable activity (lanes 3 and 4). The positive control shown in lane 6 was a dU-containing substrate, treated only with UDG. Note that some of the original material remained at the origin, indicating that UDG treatment was not complete.

Thus, taken together, the results of Fig. 6 demonstrate that the level of encapsidation of APO3G into virions (Fig. 6B) is correlated with the nucleic acid binding affinity measured in vitro (Fig. 6A). Moreover, although both zinc fingers have binding and encapsidation capabilities, the first zinc finger domain plays a more important role than the second zinc finger domain.

**Correlation between deaminase activity and the anti-HIV-1 effect of APO3G.** To clarify the effect of the zinc finger mutations on deaminase activity, we measured the activity of the purified WT and zinc finger mutant GST-APO3G proteins directly, using the UDG assay. As shown in Fig. 7A, WT and C100S had very similar levels of activity (lanes 1 and 2), whereas C291S and the double mutant had no detectable activity (lanes 3 and 4, respectively). The amounts of protein incorporated into virions were similar for the WT (lane 1) and C291S mutant (83% of WT) (lane 3), but encapsidation by the C100S mutant was approximately 55% of the WT level (lane 2). In contrast, the double mutation essentially eliminated incorporation of APO3G into virus particles (4% of WT) (lane 4). These data mirror the results shown in Fig. 6A. Note that the zinc finger mutations did not affect virus production, since each sample contained the same amount of CA (Fig. 6B, lower panel).

Thus, taken together, the results of Fig. 6 demonstrate that the level of encapsidation of APO3G into virions (Fig. 6B) is correlated with the nucleic acid binding affinity measured in vitro (Fig. 6A). Moreover, although both zinc fingers have binding and encapsidation capabilities, the first zinc finger domain plays a more important role than the second zinc finger domain.
The second zinc finger is solely responsible for deaminase activity. In light of these results, we wondered whether the antiviral effect of APO3G is a consequence of its deaminase activity. Until now, a consensus has not been reached. To address this issue, we investigated the antiviral activities of virions expressing WT or mutant APO3G. As shown in Fig. 7B, WT APO3G strongly suppressed virus infectivity (by 97%). In contrast, the double mutant had no antiviral effect (relative infectivity, 108%), in agreement with Newman et al. (31), presumably because the mutant protein is not encapsidated (Fig. 6B). The single mutants exhibited an intermediate response. Thus, the relative infectivity obtained with the C100S mutant was dramatically but not completely reduced (strong antiviral effect), with a measured value of 19% of the control. However, if the lower efficiency of C100S encapsidation is taken into account (Fig. 6B), the antiviral activity of this mutant approaches that of WT APO3G. In contrast, with the C291S mutant, the relative infectivity was considerable (67% of the control), consistent with a weak antiviral effect. These findings demonstrate that the second zinc finger domain is more important for antiviral activity than the first zinc finger domain and that a correlation exists between deaminase and antiviral activities.

**DISCUSSION**

Here, we present an in-depth molecular analysis of the nucleic acid binding and deaminase activities of APO3G, using highly purified, enzymatically active protein expressed in a baculovirus system. We find that the substrate specificity and nucleotide preferences differ for each of these activities. In addition, we demonstrate that the two zinc finger domains have different roles in nucleic acid binding, deamination, and antiviral activity. To our knowledge, this is the first detailed study of the properties of the purified protein. In addition, we show for the first time that purified GST-tagged and untagged APO3G have very similar activities.
2C). From a structural point of view this seems reasonable, since the methyl group at position 5 of cytosine might hinder the interaction of the amino group at position 4 with the catalytic center of APO3G. Moreover, this suggests that there is likely to be a strict spatial requirement to accommodate the base being deaminated.

In addition to 5mC residues in ssDNA, cytosine residues in dsDNA, dsRNA, a DNA/RNA hybrid, or in ssRNA are also refractory to deamination by APO3G (Fig. 2). These findings are consistent with previous reports (44, 51) and with the results of sequencing genomic RNA (51, 54). However, we cannot exclude the possibility that another factor(s) might be required for intracellular deamination of RNA substrates, as in the case of APO1 (4, 48).

The role of each zinc finger domain in deamination was tested directly, and the results were unequivocal (Fig. 7A). Thus, the second zinc finger domain is solely responsible for deaminase activity and the first zinc finger domain does not modify this activity. This conclusion is in accord with other work (14, 30, 31). In two studies, however, it was found that mutations in the catalytic Glu residues (40) or in zinc-coordinating residues (54) from either domain exhibited a similar reduction in deaminase activity. The explanation for these discrepancies with our findings and those of others (14, 30, 31) is not clear.

The nature of the nucleic acid binding properties of APO3G is a major issue that we address, since the binding capacity is essential for encapsidation of APO3G into virions (22, 30, 37, 45, 53) and in all likelihood for antiviral activity as well. These studies were performed in vitro using purified enzyme, but it is possible that, in vivo, other factor(s) that could affect the binding properties of APO3G might also be involved. In the experiments presented here, the EMSA results show that APO3G binds to ssRNA or ssDNA with similar efficiency (Fig. 3A) (51). We also find that APO3G has a relatively high affinity for a DNA/RNA hybrid (~50% of the binding activity seen with ssDNA or ssRNA) (Fig. 3B), although a lower binding efficiency was observed when a GST-APO3G protein expressed in E. coli was used (51). The ability to bind to a DNA/RNA hybrid might allow APO3G to remain transiently bound to the hybrid formed during minus-strand DNA synthesis. Thus, as RNase H degrades genomic RNA, APO3G would be positioned to readily deaminate the resulting ssDNA.

Additionally, we find that APO3G prefers to bind to dT or dU residues in homopolymeric ssDNAs and only poorly to similar oligos containing dC or dA residues (Fig. 3C). Surprisingly, the E. coli-expressed GST-APO3G was reported to have a preference for binding to C-rich ssDNA and RNA oligos (51). The possibility that the E. coli protein might not fold into the proper native conformation might account for this apparent discrepancy. Although in vivo data are not available, it is noteworthy that the results obtained here (Fig. 3C) and in other experiments showing a preference for binding to a dAT-rich oligo compared with a dA-rich oligo (data not shown) are consistent with the preferred binding of baculovirus-expressed GST-APO3G (19) and APO1 (1, 29) to AU-rich ssRNA. Importantly, as is the case for APO1 (1, 29), our results also demonstrate that the nucleotide preferences and substrate specificities of purified APO3G differ for deaminase and nucleic acid binding activities (compare Fig. 2 and 3).

In other experiments, we show that APO3G binding to ssDNA in the EMSA is efficient as long as the nucleic acid is 16 nt or longer (Fig. 4). Our findings also suggest that the single-shifted band observed by EMSA with the 20-, 18-, or 16-nt ssDNA (Fig. 4) consists of a single uniform complex with the nucleic acid and APO3G. In addition, from these results it would appear that APO3G covers approximately 10 to 15 nt. Interestingly, a dimeric form of APO3G is detected when an 18-nt ssDNA-APO3G complex is chemically cross-linked (Y. Iwatani and J. G. Levin, unpublished observation). This is in accord with data indicating that APO3G can form oligomers with other APO3G molecules (19, 30, 32, 40, 49) as well as with APO3B (19) or APO3F (49). Moreover, pull-down experiments have shown that APO3G oligomerization is extremely sensitive to RNase, implying that nucleic acid plays a role in this process (32). Structural studies will be necessary to determine the exact nature of nucleic acid complexes formed with APO3G in its monomeric and oligomeric forms.

NC and APO3G are both nucleic acid binding proteins, yet NC does not affect deamination of an ssDNA oligo by purified APO3G (Fig. 2A). Moreover, when NC and APO3G are both present in the same reaction, analysis by EMSA demonstrates that neither protein interferes with the other’s binding to RNA (either a 40-nt ssRNA oligo [Fig. 5A and B] or the biologically relevant 33-nt SL-3 RNA oligo [Fig. 5C and D]). Rather, a complex consisting of NC, APO3G, and RNA is formed. With the 40-nt ssRNA, binding does not appear to be cooperative, since NC does not affect the efficiency of APO3G binding and vice versa (Fig. 5A and B). Interestingly, in reactions with SL-3 RNA, NC promotes more efficient binding of APO3G to the stem-loop structure (Fig. 5C and D). This is presumably due to NC’s nucleic acid chaperone activity, which allows NC to destabilize secondary structures (reviewed in reference 25). As a consequence, the RNA becomes a more suitable substrate for APO3G binding. Taken together, these findings strongly support the results of other studies showing that an RNA binding or “bridging” activity is required for APO3G packaging by the NC domain in Gag (30, 37, 53). The data are also consistent with an alternative model suggesting that the Gag-APO3G interaction is mediated by protein-RNA interactions as well as by a specific protein-protein interaction (7).

APO3G binding to nucleic acids is dependent on zinc coordination, and mutational analysis indicates that there are differences in the functions of each domain (Fig. 6 and 7). How can we rationalize our observation that the first zinc finger domain is more important for nucleic acid binding (Fig. 6) (30) while deaminase activity is associated only with the second zinc finger domain (Fig. 7A) (14, 30, 31)? Amino acid sequence analysis provides some explanation for the greater binding activity of the first zinc finger domain. For example, the sequence motif of the first zinc finger (36 residues) contains eight positively charged (Arg, Lys, and His) amino acids, while zinc finger two (35 residues) has only two. Based on computer analysis, we estimate that the pI of the N-terminal half of APO3G (residues 1 to 188) is 9.63, whereas the value for the C-terminal domain (residues 189 to 384) is 5.90. These differences in amino acid charge would be expected to affect the nucleic acid binding activity of the two domains. In addition, it is known that aromatic amino acids found within nucleic acid binding proteins are often essential for binding activity (e.g.,
APO1 [29], APO3G [30], and NC [25]). Thus, in APO3G, the first zinc finger domain contains nine aromatic amino acids (enriched for Trp compared to zinc finger two), while the second zinc finger domain has seven. Taken together, these considerations reinforce the conclusion that the two zinc finger domains are not equivalent (18). Presumably, different folding of the two domains accounts for the exclusive presence of a functional catalytic center in the second zinc finger.

A critical aspect of APO3G's biological activity is its ability to strongly inhibit HIV-1 infectivity in the absence of Vif. We have asked a key question: Is the antiviral effect a consequence of APO3G's deaminase activity or is another mechanism involved? One approach has been to study the effects of point mutations in the zinc finger domains. A number of investigators have found that deamination is linked to antiviral activity (26, 30, 40, 54), although there has been disagreement over whether both zinc finger domains (26, 54) or primarily the second domain (30, 40) is involved. In contrast, in a recent study it was reported that point mutations in either the first or second domain had similar inhibitory effects on antiviral activity. These findings led to the proposal that the antiviral function of APO3G can be dissociated from deaminase activity and that some other mechanism, as yet undefined, is responsible for antiviral activity (31). These authors, who used untagged proteins for their work, suggested that conflicting data obtained by others might result from the use of epitope-tagged proteins, which could interfere with mutant APO3G expression and function (31). However, one of us (K. Strebel) has found that epitope-tagged and untagged WT or mutant APO3G have equivalent activities in a cell-based system (32). It is possible that these disparate results concerning antiviral activity reflect differences in experimental protocols, which would affect APO3G expression and/or encapsidation into virions.

In the present study, our results indicate that deamination is correlated with APO3G's antiviral effect. Thus, the second zinc finger domain, which contains the active site for deamination (Fig. 7A), is more important than domain one with respect to antiviral activity (Fig. 7B). However, we cannot rule out some contribution from the first zinc finger domain. If deamination indeed plays a crucial role in promoting the antiviral effect, it raises the important question of what mechanism is used by APO3G to block HIV-1 reverse transcription (12, 41, 47) and, ultimately, viral infectivity. Extensive G-to-A hypermutation in plus-strand DNA and possible degradation of DNA with abasic sites (by DNA polymerase) would affect APO3G expression and/or encapsidation into virions.

In summary, we have described the molecular properties of purified human APO3G and have demonstrated their relation to APO3G's antiviral activity. The results demonstrate that the second zinc finger domain, which we show contains the active site for deaminase activity, has a major role in the antiviral effect of APO3G and thus support a mechanism for antiviral activity that involves deamination.

ACKNOWLEDGMENTS

We thank Robert J. Crouch and Sergey Gaidamakov for assistance with the purification of APO3G as well as for helpful discussions. We are also indebted to Robert J. Gorelick for his generous gift of recombinant NC protein, Alan Rein for critical reading of the manuscript, and Tiyun Wu for gracious assistance with preparation of the figures. This research was supported by the Intramural Research Program of the NIH (NICHD and NIAID).

REFERENCES


