Biochemical Characterization of APOBEC3H Variants: Implications for Their HIV-1 Restriction Activity and mC Modification

Jiang Gu[1], Qihan Chen[1], Xiao Xiao[1,4], Fumiaki Ito[1], Aaron Wolfe[1,4] and Xiaojiang S. Chen[1,2,3,4]

*1 - **Molecular and Computational Biology Program,** Departments of Biological Sciences and Chemistry, University of Southern California, Los Angeles, CA 90089, USA*
*2 - **Center of Excellence in NanoBiophysics,** University of Southern California, Los Angeles, CA 90089, USA*
*3 - **Norris Comprehensive Cancer Center,** University of Southern California, Los Angeles, CA 90089, USA*
*4 - **Genetic,** Molecular and Cellular Biology, Keck School of Medicine, University of Southern California, Los Angeles, CA 90089, USA*

***Correspondence to Xiaojiang S. Chen:** Molecular and Computational Biology Program, Departments of Biological Sciences and Chemistry, University of Southern California, Los Angeles, CA 90089, USA. xiaojiac@usc.edu*
http://dx.doi.org/10.1016/j.jmb.2016.08.012
Edited by M.F. Summers

Abstract

APOBEC3H (A3H) is the most polymorphic member of the APOBEC3 family. Seven haplotypes (hap I–VII) and four mRNA splicing variants (SV) of A3H have been identified. The various haplotypes differ in anti-HIV activity, which is attributed to differences in protein stability, subcellular distribution, and/or RNA binding and virion packaging. Here, we report the first comparative biochemical studies of all the A3H variants using highly purified proteins. We show that all haplotypes were stably expressed and could be purified to homogeneity by *Escherichia coli* expression. Surprisingly, four out of the seven haplotypes showed high cytosine (C) deaminase activity, with hap V displaying extremely high activity that was comparable to the highly active A3A. Furthermore, all four haplotypes that were active in C deamination were also highly active on methylated C (mC), with hap II displaying almost equal deamination efficiency on both. The deamination activity of these A3H variants correlates well with their reported anti-HIV activity for the different haplotypes, suggesting that deaminase activity may be an important factor in determining their respective anti-HIV activities. Moreover, mC deamination of A3H displayed a strong preference for the sequence motif of T-mCpG-C/G, which may suggest a potential role in genomic mC modification at the characteristic “CpG” island motif.

Introduction

The seven APOBEC3 (A3) family members of DNA cytosine (C) deaminases, which are clustered on human chromosome 22 (reviewed in Ref. [1] and references therein), play a defensive role against endogenous retroelements and infectious retroviruses [2–6]. APOBEC3H (A3H) is the most divergent member of the A3 family, containing a Z3 Zn-coordinating domain that is phylogenetically distinct from the Z1- and Z2-type domains of other A3 proteins [7,8]. A3H is also the most polymorphic of the A3 family [7–9]; A3H mRNA can undergo alternative splicing to generate four splicing variants (SVs; Fig. 1a): SV154, SV182, SV183, and SV200 [4,10,11], and seven distinct human A3H haplotypes (hap I–VII) have been identified, which are composed of various combinations of five single nucleotide polymorphisms (Fig. 1b and c) [4,10,11].

Evidence indicates that the seven A3H haplotypes possess different antiviral activities. Only hap II, hap V, and hap VII are reported to effectively restrict Vif-deficient HIV-1 [4,10–12], and previous literature suggests that the anti-HIV activity of A3H can be through both deaminase-dependent and -independent pathways [10,13]. Previous studies indicate that the different anti-HIV activities of A3H variants can be attributed to the following factors: variation in protein stability of the haplotypes [4,14–16], different subcellular localizations [17,18], different binding affinities to RNA [18], and relative levels of virion packaging [12].

Additionally, A3H has been shown to be found in different oligomeric forms, and it can oligomerize in cells [16,19,20].

The distribution of A3H haplotypes in the human population is correlated with geographical location [4,10]. Evidence indicates that the more stable A3H haplotypes can suppress HIV replication effectively enough to delay infection [21], and a higher frequency of the highly active and stable hap II is present in Africa, possibly a result of the long-term presence of the HIV viral pathogen endogenous to the region [4,22–24]. On the other hand, a majority of people of non-African descent carry A3H alleles for an inactive form of the A3H haplotypes, and the variants of Vif extracted from HIV-1 strains infecting these populations are less effective at degrading stable A3H alleles prior to adaptation [4,10,15,21]. In fact, it was shown that adaptive changes in viral Vif sequences could be attributed to the absence or presence of the different antiviral A3H haplotypes [15,20,21,24–27]. These observations provide strong evidence for a significant role of A3H in immune defense against HIV infection.

So far, three APOBEC members, A3A, A3B, and AID, have also been shown to deaminate methylated C (mC), with A3A showing the strongest activity [28–31]. While the cellular functions of mC deamination for A3A and A3B have not been fully explained, the mC deamination activity associated with AID has been proposed as an alternative demethylation pathway for regulating methylation patterns in mouse germ cells [32] and for cell reprogramming in inducing pluripotent stem cells [33,34]. To date, no other APOBEC proteins outside of these three APOBEC members (A3A, A3B, and AID) have been shown to have significant levels of mC deamination activity.

Despite mounting studies on the different variants of A3H, there is currently no complete systematic study comparing the deamination and DNA substrate binding activities of all the naturally occurring A3H variants. It is also unknown whether A3H can deaminate mC. Here, we have described the expression and purification of all seven A3H haplotypes and four SVs in an *Escherichia coli* expression system. We have performed systematic comprehensive biochemical studies of these variants and show that four haplotypes of A3H (I, II, V, and VII) displayed high activity on both C and mC substrates, even though significant differences in overall activity levels exist among these variants. Surprisingly, mC deamination activity of hap II, the highest among all haplotypes, is comparable to normal C deamination by the same variant. This side-by-side biochemical study of the haplotypes of A3H reveals that deaminase activity correlates well with their differential anti-HIV activity shown by previous cellular studies. Furthermore, the highly active mC deamination ability of A3H on motifs containing -mCpG- suggests the possibility of modifying mC in CpG islands of the genome in multiple cell types that express A3H.

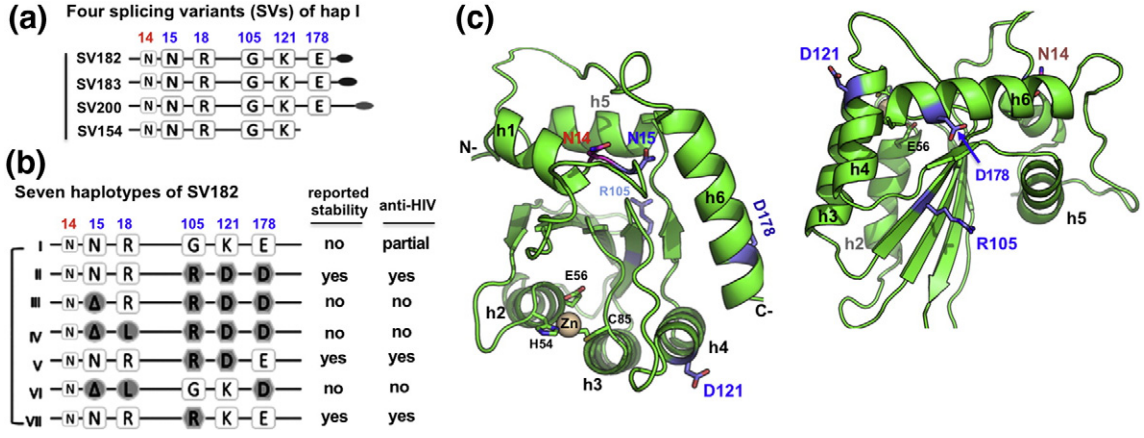

Fig. 1. The SVs and haplotypes of A3H. (a) Schematic overview of the four SVs based on hap I, with varying C-terminal extension. (b) The sequence variation of the seven haplotypes at the haplotype positions 15, 18, 105, 121, and 178 based on SV182. The position 14 (colored in red), a conserved N in all A3H variants, is also shown. The reported stability inside the cells and the anti-HIV activity for each haplotypes are listed on the right side based on literature. The weak anti-HIV activity of hap I is reported based on cell culture studies [10,11,17,18,27]. (c) Two views of a modeled A3H hap II structure based on A3C (PDB ID: 3VOW), which, like several previously published A3H models [13,18,57], shares the same core fold with all the known APOBEC structures. The locations of the haplotype residues (in sticks) are mapped to the modeled A3H structure. N15 at the end of helix 1 (h1) is buried, while its neighboring residue N14 is exposed. The R105 is on β4, the opposite side of h2, h3, and h4. D121 is at the beginning of h4, and D178 is near the end of h6.