The Aicardi-Goutières syndrome gene and HIV-1 restriction factor SAMHD1 is a dGTP-regulated deoxynucleotide triphosphohydrolase

Rebecca D. Powell#, Paul J. Holland#†, Thomas Hollis, and Fred W. Perrino *
From the Department of Biochemistry, Wake Forest School of Medicine
Winston-Salem, North Carolina 27157

Running head: SAMHD1 degrades dNTPs

Address correspondence to:  Fred W. Perrino, PhD, Department of Biochemistry, Wake Forest School of Medicine, Winston-Salem, NC 27157, Tel. 336 –716–4349; Fax. 336–716–7671; E-mail: fperrino@wfubmc.edu

Background: Mutations in SAMHD1 cause Aicardi-Goutières syndrome. SAMHD1 is the HIV-1 restriction factor preventing replication in myeloid cells.

Results: The SAMHD1 HD-domain contains a dGTP-regulated deoxynucleotide triphosphohydrolase.

Conclusion: SAMHD1 catalytic activity reduces deoxynucleotide precursor levels restricting HIV-1 replication.

Significance: The HIV-2 Vpx accessory protein targets SAMHD1 for degradation in a viral strategy to control cellular deoxynucleotide levels for efficient replication.

The SAMHD1 protein is an HIV-1 restriction factor that is targeted by the HIV-2 accessory protein Vpx in myeloid-lineage cells. Mutations in the SAMHD1 gene cause Aicardi-Goutières syndrome, a genetic disease that mimics congenital viral infection. To determine the physiological function of the SAMHD1 protein, the SAMHD1 gene was cloned, recombinant protein was produced, and the catalytic activity of the purified enzyme was identified. We show that SAMHD1 contains a dGTP-regulated deoxynucleotide triphosphohydrolase. We propose that Vpx targets SAMHD1 for degradation in a viral strategy to control cellular deoxynucleotide levels for efficient replication.

Viruses target specific host cell molecules that would otherwise restrict viral propagation. The HIV-1 and other lentiviral genomes encode different sets of accessory proteins directed at eluding host adaptive and innate immune mechanisms to promote viral replication in specific cell environments (1,2). The Vpx and Vpr accessory proteins have been linked mostly to infection of myeloid lineage cells by HIVs (3). The Vpx accessory protein is found exclusively in HIV-2 and some SIVs, and Vpr is found in HIV-1 and HIV-2. Vpx targets a restriction factor that prohibits viral DNA synthesis in monocytic cells (4-6). Recently, the protein encoded by the SAMHD1 gene was identified as the HIV-1 restriction factor in cells of the myeloid lineage (7,8). SAMHD1 is thought to inhibit an early step in the viral life cycle by interfering with efficient synthesis of viral DNA but the mechanism is not known.

SAMHD1 was identified as an interferon-induced protein in human macrophages and dendritic cells (9-12) suggesting a function in the innate immune response. Mutations in SAMHD1 cause the severe neurodegenerative disorder Aicardi-Goutières syndrome (13), a genetic encephalopathy that closely mimics congenital viral infection and is characterized by inappropriate immune activation and aberrant interferon-α secretion (14-16). Likewise, mutations in the TREX1 exonuclease gene and in the three genes encoding the RNase H2 endonuclease cause Aicardi-Goutières syndrome, linking nucleic acid metabolism with this autoimmune disorder (17-19). TREX1 is the major 3’ exonuclease that degrades cellular single- and double-stranded DNA (20-26). In HIV-1 infected cells, TREX1 degrades nonproductive transcripts to clear excess viral cDNA, preventing innate immune activation (27,28). The RNase H2 recognizes and cleaves ribonucleotides present in RNA/DNA duplexes (29,30), and silencing RNASEH2A impairs HIV replication (31). This suggests a role in HIV infection. Thus, SAMHD1, TREX1, and RNase H2 participate in a cellular nucleic acid metabolic pathway that overlaps the
interferon-mediated innate antiviral and inflammatory responses.

The catalytic activity of SAMHD1 is identified in this study. The SAMHD1 contains an N-terminal putative protein-interaction SAM (sterile alpha motif) -domain, and a C-terminal predicted phosphohydrolase HD (His…HisAsp…Asp) -domain. The SAMHD1 was cloned and bacterial-expressed recombinant enzyme was generated in sufficient quantities and purity to perform biochemical studies. The enzymatic properties indicate the SAMHD1 HIV-1 restriction factor is a dGTP-regulated deoxynucleotide triphosphohydrolase.

Experimental Procedures

Materials - The nucleotides were from Sigma and calf intestinal phosphatase was from Promega. The human and mouse SAMHD1 cDNAs (Invitrogen) were recovered by PCR and cloned into the pCDFDuet-1 (Novagen) plasmid. The HD-domain constructs were prepared by PCR and final constructs were verified by DNA sequencing.

Enzyme preparation - In expression studies, a His6-tag was cloned at the N- or C-terminus of the complete human and mouse SAMHD1 cDNAs and maximal yields of soluble enzymes were recovered using the mouse N-terminal His6-tag SAMHD1 constructs (data not shown). Mouse SAMHD1 plasmid constructs were transformed into Escherichia coli BL21 (DE3) Rosetta 2 cells (Novagen), grown to an A600=0.5 at 37°C, and quickly cooled on ice to 17°C. After induction with 1 mM isopropyl-β-D-thiogalactopyranoside, the cells were allowed to grow for 18 hr at 17°C. The SAMHD1 enzymes in cell extracts were bound to a nickel-NTA resin (Qiagen), washed, eluted, and purified to homogeneity using monoS chromatography (Fig. S1). TREX1 and RNase H2 were purified as described (22,30). Protein concentrations were determined by A280 using the molar extinction coefficient for complete mouse SAMHD1 ε = 63,800 M⁻¹ cm⁻¹ and SAMHD1 HD-domain (amino acids 117-627) ε= 58,050 M⁻¹ cm⁻¹.

Phosphohydrolase assays – Reactions (100 µl) were performed at 25°C in 96-well microplates initiated by enzyme addition and activity was monitored continuously at A410 using a Tecan Safire2™. Phosphatase reactions contained 50 mM HEPES (pH 7.0), 4 mM p-nitrophenyl phosphate, 5 mM of the indicated divalent metal ion, and 500 nM SAMHD1. Phosphodiesterase reactions (100 µl) contained 50 mM Tricine (pH 8.5), 4 mM bis-p-nitrophenyl phosphate or p-nitrophenyl 5’-thymidine monophosphate, 5 mM of the indicated divalent metal ion, and 500 nM SAMHD1 (32).

Nucleotide phosphohydrolase assays – The nucleotide phosphohydrolase reactions contained 20 mM Tris–HCl (pH 7.5), 2 mM dithiothreitol, 1 µM EDTA, 5 mM MgCl₂, the indicated nucleotides, and SAMHD1 enzyme. Reactions were incubated for the indicated times at 25°C and stopped by addition of EDTA to 5 mM final concentration. Protein was removed using 10,000 molecular weight cut-off concentrators (Millipore) and samples (10 µl) of the filtered solution was analyzed by ion-pair reverse-phase chromatography (33,34) using a CAPCELL PAK C18 column (Shiseido Fine Chemicals) on a Waters HPLC system. The column was equilibrated with buffer (20 mM sodium phosphate (pH 7.0), 5 mM tetra n-butylammonium phosphate, and 5% methanol) and eluted with a linear gradient of methanol from 5 to 60%. Products and reactants were quantified by measuring the A254. Quantification of the absorbance peaks was performed using the Empower software (Waters).

Results

SAMHD1 is a Phosphohydrolase – The primary amino acid sequence of SAMHD1 identifies this protein as a member of the diverse superfamily of enzymes known or predicted to possess metal-dependent phosphohydrolase activities (35). The phosphohydrolase activity contained in SAMHD1 was detected using the purified recombinant mouse enzyme with an artificial chromogenic substrate bis-p-nitrophenyl phosphate. A robust metal-dependent catalytic activity was detected using the bis-p-nitrophenyl phosphate substrate with Mg²⁺ and Mn²⁺; modest activity was detected with Co²⁺; and no activity was detected with Zn²⁺, Ni²⁺, or Ca²⁺ (Fig. S1). Similar levels of activity were detected using the purified SAMHD1 HD-domain, therefore demonstrating the HD-domain is necessary and sufficient for full catalytic function in the absence of the SAM domain. The SAMHD1 activity is ~5-fold greater than the phosphohydrolase activities...
exhibited by the TREX1 and RNase H2 nucleases using the bis-p-nitrophenyl phosphate substrate (Fig. S1). The SAMHD1 exhibits no detectible activity using p-nitrophenyl phosphate and p-nitrophenyl 5'-thymidine monophosphate substrates, in contrast to the robust TREX1 activity with p-nitrophenyl 5'-thymidine monophosphate (data not shown). These results indicate SAMHD1 preference for phosphoester bond cleavage. Therefore, SAMHD1 was incubated with a variety of DNA and RNA polynucleotide substrates with no nuclease activity detected (data not shown). Thus, SAMHD1 is a metal-dependent phosphohydrolase with considerably different substrate specificity compared to the TREX1 and RNase H2 nucleases.

**SAMHD1 is a deoxynucleotide triphosphohydrolase** - SAMHD1 cleaves deoxynucleoside triphosphates at the α-phosphate position generating the deoxynucleoside and triphosphate products. SAMHD1 was incubated in a time course reaction containing dGTP, dATP, dTTP, and dCTP with Mg\(^{2+}\) and the products were fractionated by HPLC chromatography (Fig. 1A). SAMHD1 triphosphohydrolase activity is apparent by the appearance of four product peaks eluting from the column at positions corresponding to the deoxynucleosides with the four products accumulating over time. Quantification of the deoxynucleoside products indicated that SAMHD1 hydrolyzed each of the four nucleotides at similar rates of ~8 µM/min when all deoxynucleoside triphosphates are present in the reaction at 200 µM (Fig. 1B). In the presence of only dGTP SAMHD1 activity was similar at a rate of ~6 µM/min. However, the SAMHD1 activity was about 20-fold lower at ~0.3 µM/min when only dATP, dTTP, or dCTP were present in reactions (Fig. 1B). SAMHD1 activity was also measured in reactions with the individual dNTP substrates at a ten-fold higher concentration of 2000 µM (Fig. 1C). The SAMHD1 triphosphohydrolase activity of ~6 µM/min was similar for dGTP at both concentrations. The SAMHD1 activity in the presence of 2000 µM dNTPs resulted in measurable increases to ~1.4 µM/min with dATP, ~2.2 µM/min with dCTP, and ~1.8 µM/min with dTTP. Together, these data indicate SAMHD1 has a much higher affinity for dGTP relative to dATP, dTTP, and dCTP nucleotides and efficient cleavage of deoxynucleotides is dependent upon the presence of dGTP. The dGTP-dependence of SAMHD1 activity was demonstrated in reactions containing dATP, dTTP, or dCTP with dGTP (Fig. S2) and in reactions containing combinations of dATP, dTTP, and dCTP with dGTP (Fig. S3). Robust SAMHD1 hydrolysis was detected in all reactions containing one to four nucleotides only when dGTP was present in the reaction.

The deoxynucleoside triphosphate specificity of SAMHD1 was tested. SAMHD1 exhibits a modest but detectible level of triphosphohydrolase activity using GTP when incubated in the presence of all four ribonucleoside triphosphates with Mg\(^{2+}\) as the activating metal. However, hydrolysis of ATP, UTP, and CTP is not detected (data not shown). There is no evidence for generation of the mono- or diphosphate nucleotides, indicating SAMHD1 cleaves dNTPs to generate only the deoxynucleoside and triphosphate products. SAMHD1 exhibits no detectible activity using dGMP or dGDP. Also, SAMHD1 does not exhibit triphosphohydrolase activity of dATP using dGMP, dGDP, and dG\(^{2+}\) as the activator molecule.

**SAMHD1 is a dGTP-regulated deoxynucleotide triphosphohydrolase** – Binding of dGTP regulates SAMHD1 deoxynucleotide cleavage activity. The dGTP-regulation of SAMHD1 was demonstrated by incubating with 200 µM dATP, dTTP, or dCTP in the presence of increased concentrations of dGTP and quantifying the dA, dT, and dC products (Fig. 2). The SAMHD1 exhibits dramatically increased rates of dATP, dTTP, and dCTP hydrolysis upon addition of increased dGTP concentrations. The addition of only 150 nM dGTP resulted in half-maximal SAMHD1 activation to ~4.5 µM/min for hydrolysis of dATP and dTTP. Complete activation of SAMHD1 for dATP and dTTP was detected at 5 and 25 µM dGTP, respectively. The activation of SAMHD1 for dCTP hydrolysis required higher levels of dGTP with the half-maximal activity of ~4.7 µM/min detected at 25 µM dGTP and complete activation detected at 100 µM dGTP. These results demonstrate the differential activation of SAMHD1 by dGTP. The relatively low concentrations of dGTP required in these reactions to activate SAMHD1 triphosphohydrolase activity, particularly for dATP.
and dTTP, indicate that dGTP is an essential activator and suggest the presence of a dGTP-specific ligand-binding site. The SAMHD1 activity exhibits a hyperbolic response to increased concentrations of dGTP (Fig. 2). The dramatic increase in SAMHD1 deoxynucleotide triphosphohydrolase activity at relatively low dGTP concentrations points to a guanine-specific control mechanism.

**Discussion**

SAMHD1 is an HIV-1 restriction factor that prohibits efficient viral DNA synthesis. Our investigations of SAMHD1 catalytic function have revealed its deoxynucleotide triphosphohydrolase activity, uncovering the mechanism of viral restriction. The NMR structure of SAMHD1 SAM-domain has been determined (PDB 2E8O), confirming the presence of this protein-protein interaction domain in the N-terminal region. Sequence alignments identified SAMHD1 genes in diverse eukaryotes, but the lack of biochemical studies limited insight into physiological function. To initiate SAMHD1 functional studies a structural model was generated using ensemble fold recognition in the program Phyre (36,37). Using the complete human and mouse SAMHD1 sequences the apparent homologue *Enterococcus faecalis* EF1143 protein (PDB 3IRH) was identified (38). The *Efa* EF1143 protein is a deoxynucleotide triphosphohydrolase and the SAMHD1 HD-domain model suggests a high degree of sequence conservation at functional residues throughout the HD-domain (Fig. S4). The likely SAMHD1 active site metal-binding residues His-167, His-206, Asp-207, and Asp-311 (His-66, His-110, Asp-111, and Asp-183 in *Efa* EF1143 protein) are identified. Also, the proposed nucleophile-generating residues His-210, Asp-218, and His-233 (His-114, Glu-122, and His-129 in *Efa* EF1143 protein) are conserved. The EF1143 X-ray structure is a tetramer with four dGTP-specific regulatory sites identified (38). Importantly, the SAMHD1 model indicates residues Lys-116, Arg-145, Phe-165, Arg-451, Lys-455, and Lys-556 (Lys-14, Asn-36, Glu-41, Arg-44, Phe-64, Arg-326, Lys-330, and Lys-422 in *Efa* EF1143 protein) are in direct contact with the dGTP ligand. The model also predicts that residues Gln-149, His-162, Arg-164, Asp-319, His-370, and Tyr-374 (Gln-48, His-61, Arg-63, Asp-191, Tyr-239, and Tyr-243 in *Efa* EF1143 protein) are likely responsible for conferring deoxyribose substrate specificity. A high-resolution structure of SAMHD1 HD-domain will be required to validate this model.

The HIV-2 Vpx targeting of SAMHD1 deoxynucleotide triphosphohydrolase to permit viral replication has a striking parallel in bacteriophage T7 restriction in bacteria. The *E. coli* mutant strain *OptA1* is unable to support replication of bacteriophage T7 gene 1.2 mutants (39). The *OptA1* strain is mutated in the upstream region of the *dgt* gene causing overproduction of the encoded dGTPase and a reduced cellular level of dGTP (40,41). Thus, the high triphosphohydrolase activity in the *OptA1* strain resulted in cellular nucleotide precursor levels insufficient to support bacteriophage T7 replication. The gene 1.2 protein inhibits the *E. coli* dGTPase, which raises nucleotide levels and permits T7 bacteriophage replication (42,43). Thus, targeting key nucleotide metabolism enzymes to control replication precursor molecule levels represents a mechanistic strategy in viral restriction that is perhaps more common than currently appreciated.

The finding of SAMHD1 deoxynucleotide triphosphohydrolase activity now links this Aicardi-Goutières syndrome enzyme with TREX1 and RNase H2 to various aspects of nucleic acid metabolism. Dysfunction of SAMHD1, TREX1 and RNase H2 enzymes in the same immune activation disease illustrates the critical nature of proper nucleotide metabolism, polynucleotide degradation, and RNA/DNA hybrid processing. Establishing SAMHD1 as an HIV-1 restriction factor and TREX1 in the degradation of HIV-1 transcripts to prevent innate immune activation places these enzymes at the interface of type I-interferon mediated autoimmunity and the antiviral response. Understanding SAMHD1, TREX1 and RNase H2 in cellular nucleic metabolism could provide new therapeutic strategies in autoimmune and viral disease.
REFERENCES


**FOOTNOTES**

*These authors contributed equally to this work.
† Current address: Institute for Bioscience and Biotechnology Research, University of Maryland, College Park, 9600 Gudelsky Drive, Rockville, MD 20850
*This work was supported, in part, by National Institutes of Health Grant GM069962 (to F. W. P.). This work was also supported by Alliance for Lupus Research Grant 179222 (to F. W. P.) and American Heart Association Grant 10GRNT3650033 (to T. H.).
*The abbreviations used are: dG, deoxyguanosine; dA, deoxyadenosine; dT, deoxthymidine; dC, deoxycytidine;
FIGURE LEGENDS

Figure 1. **SAMHD1 is a nucleotide triphosphohydrolase** – A, time course reactions containing 400 nM SAMHD1, 200 µM dGTP, dATP, dTTP, dCTP, and 5 mM MgCl₂ were performed in triplicate for the indicated times and products were fractioned by HPLC as described under “EXPERIMENTAL PROCEDURES.” A chromatogram at each time point from one of the reactions is shown. The indicated positions of elution of the substrates and products were determined in separate injections of each dNTP and nucleoside. B, the dG, dA, dC, and dT products generated at the indicated times from the three reactions in A and from triplicate time course reactions containing only dGTP (dG₀), only dATP (dA₀), only dTTP (dT₀), and only dCTP (dC₀) were quantified. C, the dG, dA, dT, and dC products from reactions containing 200 µM or 2000 µM of the indicated nucleotide were quantified and rates were calculated.

Figure 2. **SAMHD1 is a dGTP-regulated triphosphohydrolase** - Reactions containing 400 nM SAMHD1, 200 µM dATP, dTTP, or dCTP, 5 mM MgCl₂ and the indicated concentration of dGTP were performed for 20 min at 25°C. The reactions were fractioned by HPLC and the dA (●), dT (▼), and dC (○) products were quantified as described under “EXPERIMENTAL PROCEDURES.”