JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

pubs.acs.org/jcim

Identification of Novel *S*-Adenosyl-L-Homocysteine Hydrolase Inhibitors through Homology-Model-Based Virtual Screening, Synthesis, and Biological Evaluation

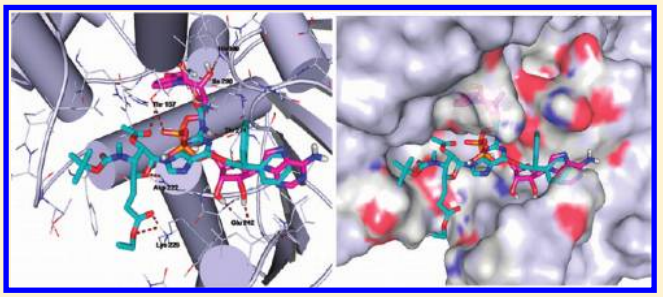
Prashant Khare,[†,#] Amit K. Gupta,[‡,#] Praveen K. Gajula,[‡] Krishna Y. Sunkari,[‡] Anil K. Jaiswal,[†] Sanchita Das,[†] Preeti Bajpai,[§] Tushar K. Chakraborty,*[,‡] Anuradha Dube,*[,†] and Anil K. Saxena*[,‡]

[†]Parasitology Division and [‡]Medicinal and Process Chemistry Division, C.S.I.R.-Central Drug Research Institute, Lucknow, India

[§]Integral University, Lucknow, India

Supporting Information

ABSTRACT: The present study describes a successful application of computational approaches to identify novel *Leishmania donovani* (*Ld*) AdoHcyase inhibitors utilizing the differences for *Ld* AdoHcyase NAD^+ binding between human and *Ld* parasite. The development and validation of the three-dimensional (3D) structures of *Ld* AdoHcyase using the *L. major* AdoHcyase as template has been carried out. At the same time, cloning of the *Ld* AdoHcyase gene from clinical strains, its overexpression and purification have been performed. Further, the model was used in combined docking and molecular dynamics studies to validate the binding site of NAD in *Ld*. The hierarchical structure based virtual screening followed by the synthesis of five active hits and enzyme inhibition assay has resulted in the identification of novel *Ld* AdoHcyase inhibitors. The most potent inhibitor, compound **5**, may serve as a "lead" for developing more potent *Ld* AdoHcy hydrolase inhibitors as potential antileishmanial agents.



INTRODUCTION

The neglected tropical diseases (NTDs) affect more than one billion people (one-sixth of the world's population).[1] *Leishmania donovani* (*Ld*) has been characterized as one of the most challenging NTD causing agent by the World Health Organization as around 350 million people are affected by leishmaniasis, and around 2 million new cases occur every year.[2] Surprisingly, the NTD drug discovery pipeline is still almost dry. Low returns of investments have discouraged drug companies from investing in research projects devoted to the discovery of novel drug candidates for NTDs.[3] No effective vaccines are available against *Leishmania* infection,[4] and treatment relies mainly on chemotherapy; yet, the available range of effective drugs is highly limited.[5]

In order to generate an adequate armory of drugs to treat visceral leishmaniasis, new and effective drug targets are required to combat this disease. Enzymes or metabolites present in the parasite but absent from their mammalian host are considered as ideal targets for rational drug design. The reversible conversion of *S*-adenosyl-L-homocysteine (AdoHcy) to adenosine (Ado) and homocysteine (Hyc)[6] involves a unique and only known enzyme *S*-AdoHcyase[7] (EC 3.3.1.1.). It occurs downstream of the *S*-adenosylmethionine (AdoMet) dependent transmethylation enzymes,[8] which are responsible for a wide variety of important biological functions. Since, AdoHcy is a product inhibitor of all *S*-adenosylmethionine (AdoMet) dependent methyltransferases, the catalytic activity of AdoHcyase is critical in eukaryotic cells to maintain the normal cellular level of AdoHcy and to permit the numerous transmethylation reactions required in normal cell functions to proceed.[8] Inhibition of AdoHcyase in eukaryotic cells causes an increase in the cellular level of AdoHcy and inhibition of AdoMet-dependent methyltransferases that result in various pharmacological effects (e.g., antiviral, antiparasitic, antiarthritic, immunosuppressive, etc.).[9] Parasites such as *Leishmania donovani*,[10] *Trypanasoma cruzi*,[11] and *Plasmodium falciparum*[12] express their own AdoHcyase; distinct from their human counterparts, these parasitic enzymes are potential targets for developing antiparasitic agents.

In order to facilitate the design of antiparasitic agents targeting parasite's AdoHcyase, it is imperative to overexpress the targeted parasitic enzyme and compare its structural and catalytic features with those of human (*Hs*) AdoHcyase. Earlier, the gene encoding AdoHcy hydrolase activity in *P. falciparum* was expressed in *E. coli*,[12] but the recombinant enzyme was not characterized. Similarly, the AdoHcyase from *T. vaginalis* was expressed as a fusion protein with a 63 His *N*-terminal tag;[13] however, its structural and catalytic features were not fully characterized.[14,15] Although the enzymatic and physicochemical properties of the purified recombinant *L. donovani* enzyme were characterized by Henderson et al. in 1992,[10] the first 3D

Received: December 9, 2011
Published: February 11, 2012

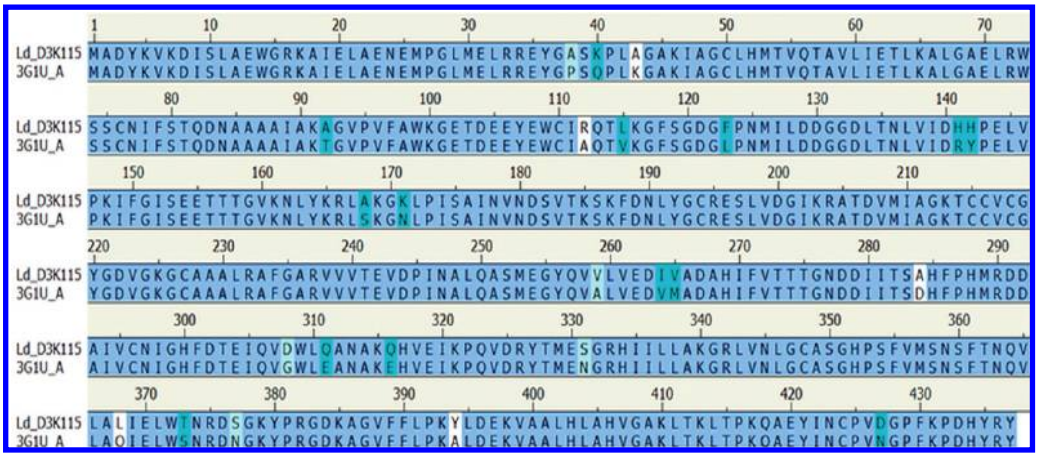


Figure 1. Multiple sequence alignment of *Lm* AdoHcyase protein (PDB ID: 3G1U) with *Ld* AdoHcyase sequence (Uniprot accession code: D3K115). The identical residues are shown in light blue color background while the residues with strong and weak similarities are shown in cyan and light sky blue colors. Non-matching residues are shown in white.

crystallographic structure of *L. major (Lm)* AdoHcyase protein was revealed by Siponen et al. in 2009 (PDB ID: 3G1U). However, the crystallographic structure of AdoHcyase protein for the most common and dominant *L. donovani* species is still unknown which is the major cause of visceral leishmaniasis (VL).

The hydrolysis of AdoHcy through AdoHcyase is NAD^+ dependent and involves two steps. First the NAD^+ binds with the AdoHcyase at the NAD^+ binding site, and thereafter, the enzyme hydrolyzes the substrate AdoHcy into Ado and Hcy. Earlier studies have shown that there are some remarkable differences in kinetic and thermodynamic parameters between the human and parasite enzymes.[8,10] *Ld* AdoHcyase has weaker binding affinity for NAD^+ and lower catalytic activity than those of the human enzyme.[10,16] The *Hs* AdoHcyase tightly binds one NAD^+ cofactor per subunit, which is essential for the catalysis. Therefore inhibitors have been designed to irreversibly reduce the NAD^+ form (active) to the NADH form (inactive). Several of these known inhibitors of the *Hs* AdoHcyase have been evaluated for their ability to inhibit parasite growth and replication.[10] It has been also reported earlier[10] and through our study that NADH showed ~98% inhibitory activity over *Ld* AdoHcyase enzyme. This inhibition is probably due to competitive binding of NADH to the NAD^+ binding site on the enzyme. Interestingly, NADH did not affect the activity of *Hs* AdoHcyase because the human enzyme binds NAD^+ tightly and the loose binding NAD^+ is unique for the *Ld* AdoHcyase. These differences provide the possibility of designing selective inhibitors of *Ld* AdoHcyase that will not inhibit human AdoHcyase and in this regard targeting the NAD^+ binding site instead of substrate binding site may be a novel approach. On the basis of this observation, the molecules were designed that may act as selective inhibitors of NAD^+ at its binding site of the *Ld* AdoHcyase than the human AdoHcyase for antileishmanial chemotherapies. Structure based virtual screening (SBVS) has been successfully utilized for the discovery of NCEs for many important targets.[17] In the present study, we report the development and validation of the three-dimensional (3D) structure of *Ld* AdoHcyase using the *Lm* AdoHcyase as a template. Simultaneously, cloning of the *Ld* AdoHcyase gene from clinical strains, its overexpression and purification were performed. This is followed by SBVS of a small focused virtual library of 8000 nucleoside analogues targeting the NAD^+ binding site of *Ld* AdoHcyase which resulted in five active hits followed by their synthesis. These hits were further subjected to the enzyme inhibition assay which led to the identification and understanding of mode of action of potential inhibitors targeting *Ld* AdoHcyase, and to best of our knowledge, this study represents the first report of inhibitors targeting the NAD^+ binding site of *Ld* AdoHcyase.

RESULT AND DISCUSSION

Generation of a Homology Model of *Ld* AdoHcyase. A plausible structural model of the *Ld* AdoHcyase has been generated using the 2.2 Å resolution crystal structure of *Lm* AdoHcyase protein (PDB ID: 3G1U) as a template. We already have submitted the complete gene sequence of *Ld* AdoHcyase to NCBI (accession code: GU353334.1) along with its amino acid sequence (Uniprot accession code: D3K115) collected from the clinical isolate 2001 of *Ld.* Multiple sequence alignment has been performed between our submitted target sequence (Uniprot accession code: D3K115) and template (PDB ID: 3G1U) using the Prime[18] module of the Schrodinger software package.[19] (Figure 1)

The global sequence identity and similarity between the *Ld* AdoHcyase (accession code: D3K115) and *Lm* AdoHcyase (PDB ID: 3G1U) amounts to 92.6% and 95.7% respectively suggesting that *Lm* AdoHcy hydrolase (PDB ID: 3G1U) may serve as a best candidate for selection as a template. Highly conserved residues were anchored, cocrystallized NAD ligand was retained, and the plausible model for *Ld* AdoHcyase was generated in the Prime[18] module of the Schrodinger software package.[19] The homology model was inspected to ensure that the side chains of the conserved residues were aligned to the template. The outlier residues were examined using the Ramachandran plot, and the loops with outlier residues were refined using the "refine loops" tool with the "serial-loop sampling" procedure of Prime.

Model Quality Assessment. Currently, there is not a single method able to consistently and accurately predict the three-dimensional structure of a protein. Similarly, there is no single method able to consistently and accurately predict the errors in a protein structure. Different methods use different approaches, thus they can complement to each other. Therefore, use of as many as possible different quality assessment (QA) methods may help in selecting the most

dx.doi.org/10.1021/ci2005862 | J. Chem. Inf. Model. 2012, 52, 777−791

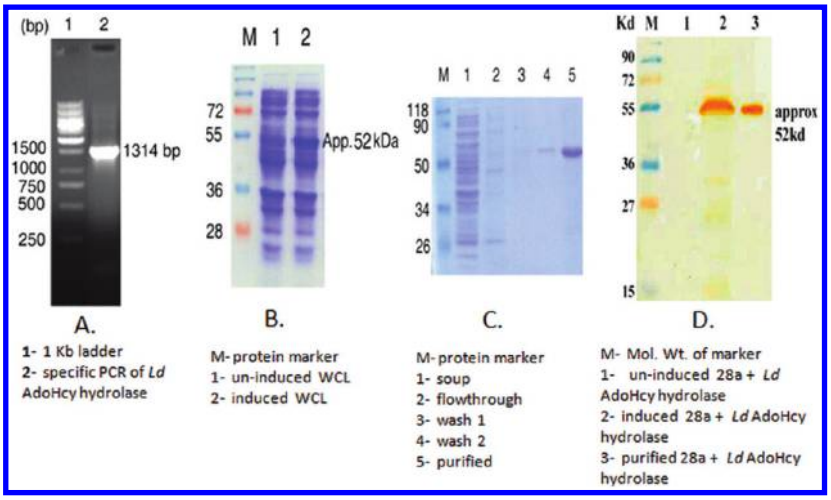


Figure 2. (A) Amplification of *Ld* AdoHcyase, (B) overexpression of *Ld* AdoHcyase, (C) purification of *Ld* AdoHcyase, and (D) Western blot confirmation of *Ld* AdoHcyase.

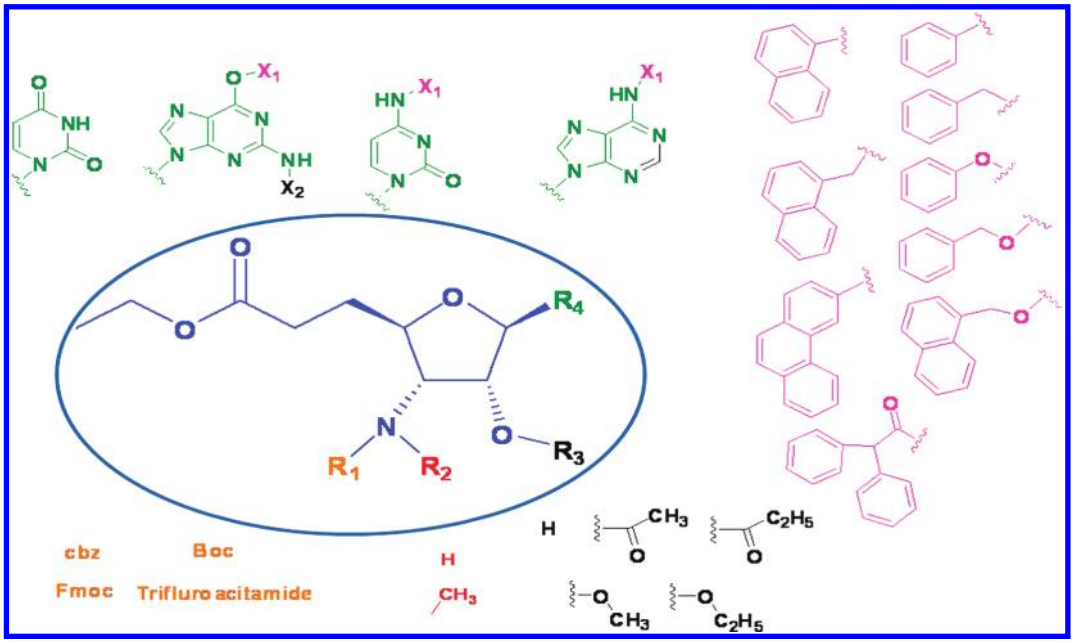


Figure 3. Protocore (encircled) used in the focused virtual library generation. R_1, R_2, R_3, R_4, X_1, and X_2 denote the attachment points where prepared reagents have been selectively attached. Orange, red, green, and pink colored functionalities were attached at R_1, R_2, R_4, and X_1 attachment points whereas black colored functionalities were attached at R_3 and X_2 attachment points.

plausible model. The resulting *Ld* AdoHcyase homology model was further subjected to different QA methods namely Procheck,[20] ProSA-web,[21] QMEAN,[22] and Errat.[23] The developed *Ld* AdoHcyase homology model has a very good stereochemical parameters and side chain packing densities within the acceptable range as there is no outlier residue (Procheck). A comparison plot between energy of amino acids with sequence positions of amino acids of *Ld* AdoHcyase homology model followed the similar trend as known PDBs (ProSA-web). The QMEAN score of 0.755 (estimated model reliability lies between 0 and 1), Z-score of −9.93, and Errat's overall protein quality factor of 95.14 (≥95 for highly significant models) further suggested the superior quality of the *Ld* AdoHcyase homology model. In order to relieve unfavorable steric clashes, this model was further energy minimized while their backbone was fixed using OPLS_2005 force field implemented in the minimize panel of the refinement tool.

Comparison of Model with Crystal Structure. The developed *Ld* AdoHcyase homology model was compared to its corresponding crystal structure (template) by calculating the root-mean square deviation (rmsd) of the α-carbons. The rmsd for the Cα helix atoms of *Lm* AdoHcyase protein (3G1U) with modeled *Ld* AdoHcyase was extremely low (0.017) which substantiated the accuracy of the developed homology model as the resulting *Ld* AdoHcyase model has highly symmetrical backbone conformations with that of the *Lm* AdoHcyase (3G1U).

Cloning, Overexpression, and Purification of *Ld* AdoHcyase. Specific forward and reverse primers were designed for our submitted complete gene sequence of *Ld* AdoHcyase (NCBI accession code: GU353334.1) and used for the amplification of complete open reading frame (ORF) of *Ld*

dx.doi.org/10.1021/ci2005862 | J. Chem. Inf. Model. 2012, 52, 777−791

AdoHcyase by polymerase chain reaction (PCR) using genomic DNA of *Ld* as template. PCR specifically amplified a fragment of 1314bp, the predicted size of *Ld* AdoHcyase (Figure 2A), which was subsequently cloned into pTZ57R/T (T/A) cloning vector and an expression vector pET 28a. Transformation of the construct 28a+ *Ld* AdoHcyase into *E. coli* strain BL21 (DE3) Rosetta expressed a soluble protein of expected size (Figure 2B). The protein was purified to homogeneity by Ni-NTA affinity chromatography (Figure 2C) which was confirmed by Western immunoblotting using anti-His antibody (Figure 2D). The details of cloning, overexpression, and purification of *Ld* AdoHcyase enzyme are described in the Experimental Section.

Focused Virtual Library Generation. Earlier we have reported the successful utilization of focused virtual libraries containing unknown compounds for new lead discovery, synthesis, and further lead optimization.[24] The fragment-based method for the generation of focused virtual library has been used in the present study using the software CombiGlide[40] which allows the user to run fragment-based generation of a virtual library in a user-friendly interface. The NAD binding the "Protocore" (Figure 3) containing furan ring along with the ethyl propionate group mimicking the phosphate and pentose sugar part of the NAD^+ were considered for the library generation using the default parameter settings implemented in CombiGlide.[25]

Various fragment sets (Figure 3) were proposed for high affinity binding, illustrating the suitability of different structural motifs, in terms of five different nitrogenous bases along with different hydrophobic (aromatic/aliphatic), ionizable (positive/negative), and vector (donor/acceptor) features, at different positions relative to central protocore. The fragment set was prepared in the "Reagent Preparation Wizard" of the maestro panel of Schrödinger software, where among different substitution reactions with reactants, "R-I" type of reactants have been chosen for reagent preparation. The focused virtual library generation was carried out using the "Combinatorial Library Enumeration" wizard of the maestro panel where the virtually prepared reagents were selectively added to an attachment points "R_1", "R_2", "R_3", "R_4", "X_1", and "X_2" (Figure 3) which resulted in 8000 new nucleoside analogues. An extensive conformational search for each ligand of this library was performed using MacroModel[26] to ensure that concluding results were not biased by the selection of only one or two low-energy conformers. The enumerated focused virtual library was then used in the hierarchical SBVS panel for the identification of potential *Ld* AdoHcyase inhibitor targeting the NAD^+ binding site.

Structure Based Virtual Screening (SBVS). The SBVS paradigm evaluates the complementarities of the molecules to the binding site of target structure and can in principle discover new chemotypes, dissimilar to previous ligands that nevertheless fit the binding site well. Such chemotypes may provide new routes for the modulation of this key target. Similarly it may also help in the understanding of the mode of action of active compounds. The SBVS exercise was performed through Glide5.6[27] docking software. The *Ld* AdoHcyase protein complex was prepared using protein preparation wizard, and the NAD^+, complexed with *Ld* AdoHcyase protein, was selected for generation of a docking grid in the Glide5.6[27] docking software. Initially NAD^+ was docked at the defined grid using the automated "Extra precision" (XP) mode of Glide5.6 to ensure the correct binding mode of NAD^+ at *Ld* AdoHcyase. The comparative NAD^+ binding site analysis of *Ld* docked complex and *Lm* X-ray crystal structure of AdoHcyase revealed that there is no significant change in the binding mode of NAD^+ both in terms of NAD^+ conformation (Figure 4) and in

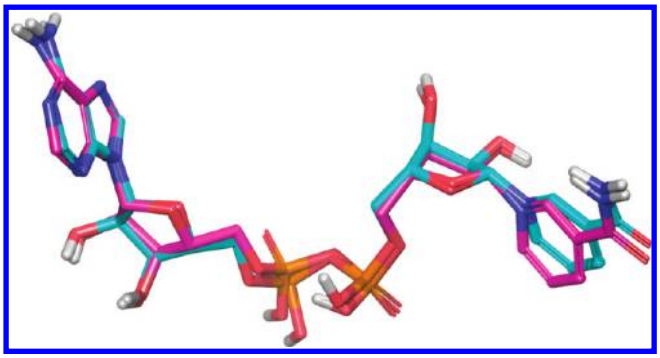


Figure 4. Comparison of binding pose for the NAD^+ in the cocrystallized structure of *Lm* AdoHcyase (pink colored carbon) and in the docked *Ld* AdoHcyase (cyan colored carbon).

the binding amino acid residues which could be easily understood by the high degree of homology for AdoHcyase between these two species (~95%).

In order to find out the novel *Ld* AdoHcyase inhibitors which may act as a competitive inhibitor of NAD^+, rigorous docking was carried out in a sequential manner in four substages, and the output of one substage served as the input of the next substage. Initially this focused chemical library of 8000 novel nucleoside analogues was screened by high-throughput virtual screening (HTVS) mode implemented in Glide5.6 where the 6360 molecules with greater than 3.5 docking score were further subjected to the SMART pattern filter algorithm implemented in Discovery Studio 2.1[28] for drug likeness and strategic pooling of compounds for high-throughput screening. Here, 4942 molecules passed this filter, and among the molecules that passed, 4290 molecules having log P value less than 5 were further subjected to the SP mode of rigid docking implemented in Glide 5.6. This mode (Glide-SP) is relatively more precise in terms of the docking and scoring of compounds. The resulting 1108 molecules with greater than 4.5 Glide-SP docking score value were further screened through the Glide extra precision (XP) mode of rigid docking which resulted in 157 molecules with a Glide-XP docking score value greater than 3.5. These 157 molecules were clustered into four clusters according to their nitrogenous base containing group (thymine and uracil were placed in the same group due to their structural similarity) and were further subjected to the docking and scoring function of another docking software GOLD version 3.1.[29] The NAD^+, complexed with *Ld* AdoHcyase protein was used for the selection of active site in the GOLD and the rigid docking exercise was performed. Interestingly the docking pose of these ligands in GOLD were the same as those of Glide which further substantiated the stability of docked poses; however, their scoring functions varied. Molecules having a GOLD fitness score greater than 60 in the case of adenine and guanine clusters and molecules having a GOLD fitness score greater than 30 in the case of uracil and cytosine clusters were screened which resulted in 96 active hits. These 96 hits were examined for their synthetic feasibility where the molecules having the highest GOLD fitness score with Boc protective group in each cluster were resulted in five active hits

dx.doi.org/10.1021/ci2005862 | J. Chem. Inf. Model. 2012, 52, 777–791

(Table 1). The whole workflow SBVS paradigm adopted in this study has been represented in Figure 5.

Table 1. Summary of the Docking Scores for the Five Active Hits Gained from the SBVS

S.N.	Ligand Name	Nitrogenous base content	Glide-XP Score	GOLD Fitness score
1	1	Cytosine	-3.68197	54.1134
2	2	Cytosine	-4.75851	30.1852
3	3	Uracil	-4.71961	43.0192
4	4	Adenine	-3.51288	61.33
5	5	Guanine	-4.38775	65.9553

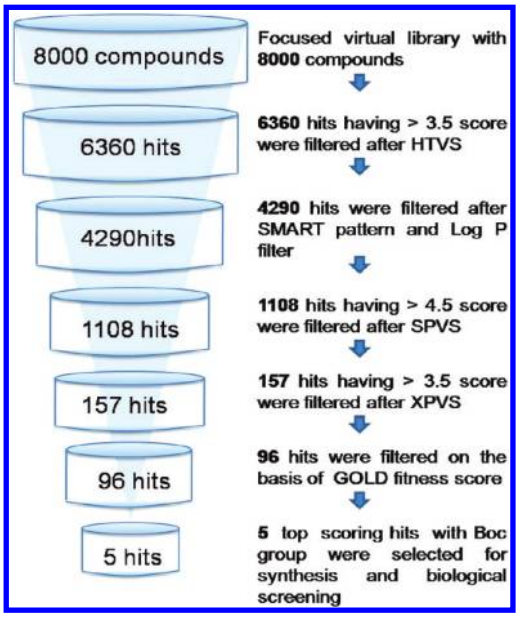


Figure 5. Flowchart of virtual screening procedure adopted in this study.

These five hits were further examined to low-temperature molecular dynamics simulations (MDS) using MacroModel to ensure the stable receptor−ligand interactions. The graphical plot analysis of Cα backbone rmsd and potential energy of protein ligand complex of 10 sampled structures for each ligand further substantiated the stability of the bound ligand into the NAD^+ binding site of *Ld* AdoHcyase. Further in order to check the solvent effects on binding poses of reported enzyme−ligand complexes, a graphical plot analysis for solvation energy of 10 samples of each enzyme−ligand complex of 100 ps MDS were performed in a similar environment where no drastic movement of the bound ligand and surrounding residues was observed due to a solvent effect during MDS run. The potential energy of all the five receptor ligand complexes reported in Table 1 of the manuscript except compound **1** were in their minimum energy state indicating their most favorable state. The potential energy difference between the minimum potential energy state of the MDS and docking output state for compound **1** was only 18.5 kcal/mol. In order to confirm that it is the most favorable binding state for compound 1 the minimum potential energy state of the compound **1**−enzyme complex was again subjected to the MDS study of 100 ps. The 2D-line plot analysis of 10 samples of the receptor−ligand complex of MDS run revealed the greater stability of this enzyme−ligand complex compared to that found earlier; however, no significant changes were observed in terms of bound ligand conformation as well as important amino acid interactions at the active site.

The binding site analysis of NAD^+ clearly deciphers that it covers a large binding cavity due to the large size of NAD^+ where the adenine moiety of the molecules occupies the hydrophobic cleft formed by the residues namely Ile-280, Val-243, Gly-219, Asn-247, Thr-241, and Asp-244 whereas two hydroxyl residues of pentose sugar attached with the adenine moiety makes the H bond with Glu-242 residue (Figure 6A). The phosphate group of the NAD^+ also makes the H bond with Val-223. It has also been observed that the three consecutive threonine residues 156, 157, and 158 along with His-300 play an important role in NAD^+ binding as they provide the anchorage to the pentose sugar attached with nicotinamide moiety of the molecule through H bonding. The nicotinamide moiety itself attached with Asn-345 and Ile-298 residues through H bonding (Figure 6A and B).

The binding site analysis of three potent inhibitors (compound **1**, **5**, and **2**) demonstrated the nonspecific binding of these ligands at the NAD^+ binding site. However the inhibitory activity of these molecules varies according to the binding efficiency and tendency to mimic the NAD^+. The diphenylacetoxy group of compound **5** occupied the hydrophobic cleft formed by the residues Gly-211, Thr-275, Val-243, Glu-242, Gly-219, and Gly-276 similar to the adenine moiety of the NAD^+ (Figure 6C). The acetylamido group attached with nitrogenous base interacts with the Thr-274 through H bond whereas the acetyl group attached with tetrahydrofuran moiety of the molecule binds with Thr-157 residue through H bonding similarly as the diphosphate group of NAD^+ (Figure 6B). The binding surface analysis of compound **5** deciphers the dissimilar trend of binding for the aminoBoc and propionate groups attached with the tetrahydrofuran moiety in compound **5** with respect to NAD^+(Figure 6D). These functionalities along with tetrahydrofuran occupied the outer binding cleft made by the residues Asp-181, Ser-182, Thr-184, Gln-250, and Lys-225. The binding site analysis of compound **1** and **2** gave interesting findings as only two differences between these two molecules were the substitution of hydrogen of −OH and −NH group by acetyl and methyl group in the tetrahydrofuran moiety of the molecule. These changes interchange the orientation of cytosine base and propionate moiety of the compound with respect to each other. The aminoBoc moiety of both the molecules occupied the same hydrophobic pocket made by the residues Tyr-220, Gly-221 Gly-219, Glu-242, and Gly-276 which is similar to the hydrophobic pocket occupied by the adenosine moiety of NAD^+ (Figure 7A−C). The propanoate moiety of compound **2** occupied the binding cleft made by the residues Thr-156, Thr-157, Thr-158, Val-223, Ile-298, Gly-299, and His-300 where the carbonyl oxygen of propanoate moiety provides the additional anchorage to the molecule through H bonding with His-300 (Figure 7B). The cytosine moiety of the molecule follows the similar binding trend as with the adenine moiety of NAD^+ (Figure 7C). The orientation of cytosine and propionate moiety in compound **1** was found inverted with respect to compound **2** where the cytosine moiety compound **1** occupied the same binding pocket as occupied by the propionate moiety of compound **2** while the propionate moiety of compound **1** occupied the same binding pocket as occupied by the cytosine moiety compound **2** (Figure 7D). This inverse

dx.doi.org/10.1021/ci2005862 | J. Chem. Inf. Model. 2012, 52, 777−791

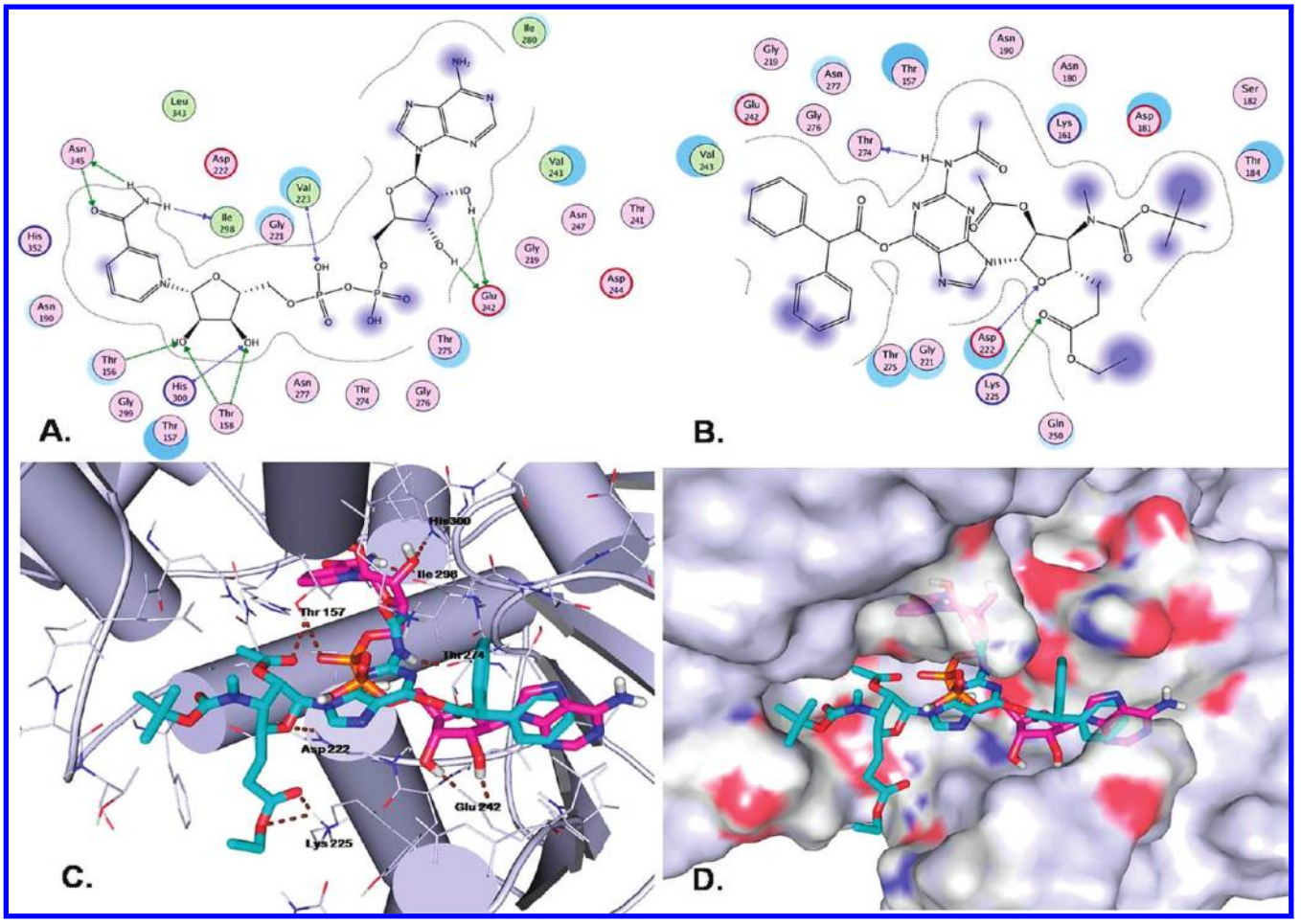


Figure 6. Binding pose analysis of NAD^+ and compound **5** at the NAD^+ binding site of *Ld* AdoHcyase. (A) 2D binding pose view of NAD^+. (B) 2D binding pose view of compound **5** (C) Comparative 3D binding pose view of compound **5** (cyan color) along with NAD^+ (pink color). (D) Binding cavity surface view of compound **5** (cyan color) along with NAD^+ (pink color).

orientation in the cytosine and propionate moiety of compound **1** was may be due to the strong H bonding of the hydoxyl group with Thr-274 which directs the inverse orientation of these two molecules.

In order to evaluate the inhibitory activity of the five active hits gained from SBVS paradigm as well as to validate the developed homology model, these five hits (Table 1) were synthesized and subjected to the enzyme inhibition assay.

Binding Free Energy Calculation. The binding free energies (ΔG_{bind}) for each protein−ligand complex were calculated using generalized Born surface area (MM/GBSA) methods implemented in the Schrodinger package. The complex, ligand, and binding free energies of the five protein−ligand complexes based on MM/GBSA calculation are summarized in Table 2. The highest stability of compound **5** was further described by its highest binding (ΔG_{bind}) energy value, which was found comparable to the NAD^+.

Synthesis. Synthesis of five active hits gained from SBVS paradigm (compounds **1**−**5**) (Figure 8) has been performed in their protected forms. Synthesis of **1** has been reported earlier.[30] For the synthesis of **2**−**5**, we started from the known compound **7** (Scheme 1) which was prepared from glucose diacetonide,[31] **6**, according to the reported procedure.[32] Reduction of azide using H_2 under atmospheric pressure in presence of Pd/C gave an intermediate amine, which on treatment with $(Boc)_2O$ furnished **8** in 85% yield. Compound **9** was obtained from compound **8** in two steps, involving reduction of $(Boc)_2O$ using LAH in THF,[33] followed by in situ treatment of amine with Fmoc-OSu and DIPEA in CH_2Cl_2 to give **9** in 80% yield over two steps. Selective deprotection of the 5,6-*O*-isopropylidine group of **9** using 77% aqueous AcOH, at 65 °C furnished diol **10** in 78% yield. Oxidative cleavage of the diol using $NaIO_4$ in THF:H_20 (2:1) afforded an aldehyde, which on Wittig olefination with $Ph_3P{=}CHCOOEt$ followed by reduction of the double bond using H_2 under atmospheric pressure in the presence of Pd/C in EtOAc afforded **11** in 83% yield. Acetonide deprotection of **11** with 50% aqueous TFA at 0 °C to room temperature furnished a lactol, which on acetylation with acetic anhydride afforded compound **12** in 85% yield over two steps. Next, compounds **13a**, **13b**, **13c**, and **13d** were synthesized from **12** under different conditions. For the syntheses of **13a**[34] and **13b**,[35] compound **12** was coupled with cytosine and uracil, respectively, in the presence of TMSOTf to give **13a** in 75% and **13b** in 72% yields. The modified nucleoside **13c** was obtained from **12** via $SnCl_4$ mediated glycosylation of unprotected adenine,[36] whereas **13d** was synthesized from **12** using a protocol developed by Zou and Robins.[36a,37] Finally deprotection[38] of Fmoc followed by in situ protection of the resulting amine with $(Boc)_2O$ afforded compounds **2**−**5** in 83−87% yields.

In summary, synthesis of compounds **2**−**5** have been achieved starting from D-glucose with overall yields of **2**,

dx.doi.org/10.1021/ci2005862 | *J. Chem. Inf. Model.* 2012, 52, 777−791

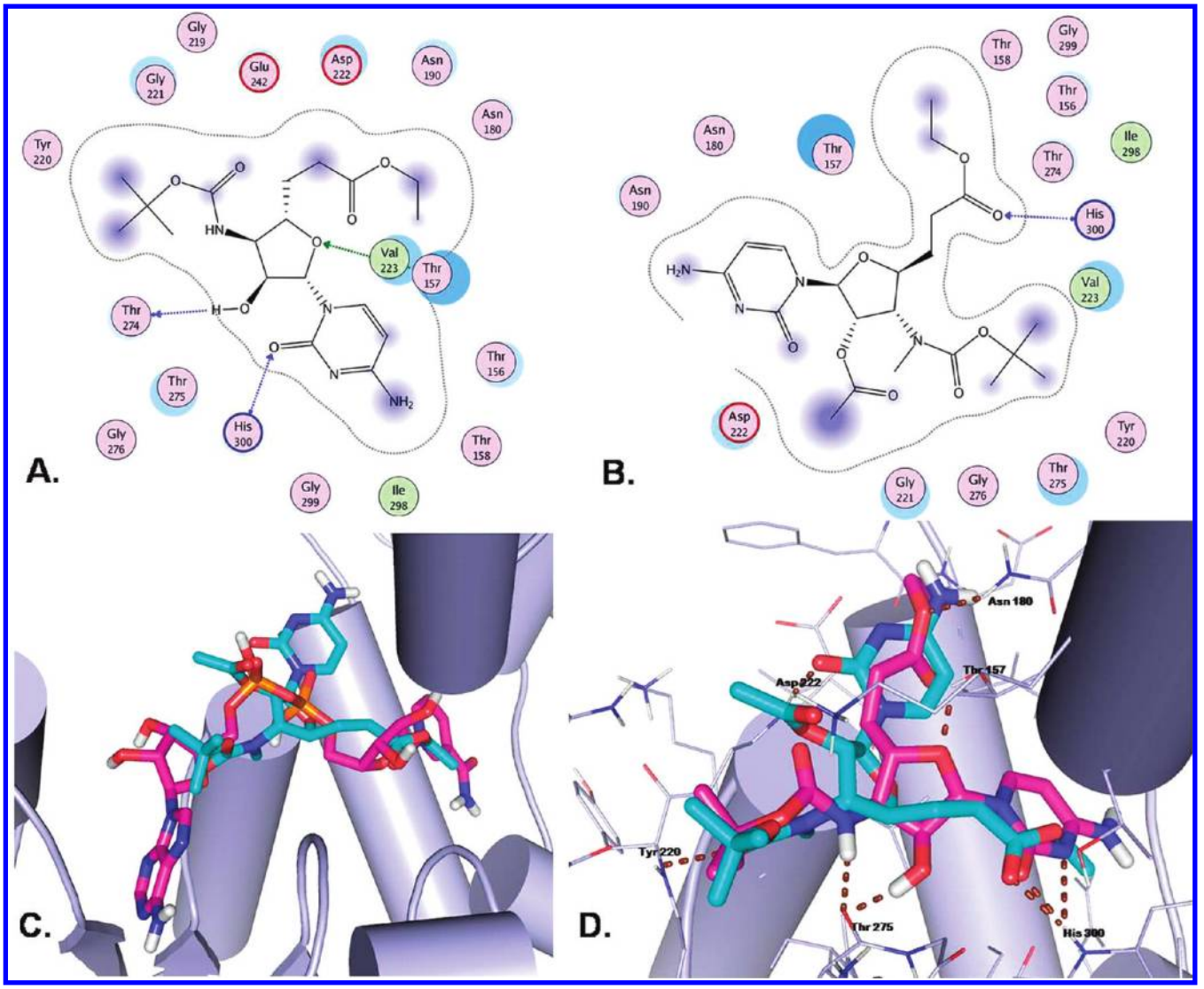


Figure 7. Binding pose analysis of compounds **1** and **2** at the NAD^+ binding site of *Ld* AdoHcyase. (A) 2D binding pose view of compound **1**. (B) 2D binding pose view of compound **2**. (C) Comparative 3D binding pose view of compound **2** (cyan color), along with NAD^+ (pink color). (D) Comparative 3D binding pose view of compound **2** (cyan color) along with compound **1** (pink color).

Table 2. Summary of the MM/GBSA Endpoint Free Energies for the Five Active Hits[a]

comp name	$G_{complex}$	G_{ligand}	ΔG_{bind}	ΔG_{Coul}	ΔG_{Cov}	ΔG_{vdW}	ΔG_{SolvSA}	ΔG_{SolvGB}
1	44645.69	−80.32	**-15.0**	−14.14	10.12	−49.08	2.59	35.51
2	44645.42	−63.95	**-31.64**	−27.68	9.07	−54.99	2.25	39.72
3	44651.05	−64.70	**-25.26**	−9.73	3.73	−50.79	6.64	24.89
4	44662.01	−45.41	**-33.59**	−21.22	1.00	−47.57	0.77	33.43
5	44637.22	−48.50	**-55.30**	−13.56	8.58	−76.78	−6.75	33.22
NAD^+	44578.14	−105.66	**-57.21**	−117.20	5.78	−71.63	0.85	124.99

[a]All reported energies are in kilocalories per mole. The total energy of the *Ld* AdoHcyase protein was 44 741.01 kcal/mol. $G_{complex}$ the total energy of the complex; G_{ligand} the energy of the unbound ligand; ΔG_{bind} free energy of protein−ligand binding; ΔG_{Coul} the columbic binding free energy; ΔG_{Cov} the covalent binding free energy; ΔG_{vdW} van der Waals binding free energy; ΔG_{SolvSA} solvation binding free energy of the surface area; ΔG_{SolvGB} generalized Born solvation binding free energy.

14.8%; **3**, 13.9%; **4**, 15.8%; and **5**, 19.1%. The corresponding overall yields of **2**−**5**, starting from azide **7**, were 29.7%, 27.8%, 31.6%, and 32.5%, respectively.

Enzyme Inhibition Assay of *Ld* AdoHcyase. The purified *Ld* AdoHcyase enzyme was subjected to the inhibition assay for the five synthesized hits gained from the SBVS. The assay of *Ld* AdoHcyase activity in the hydrolytic direction was performed spectroscopically by measuring the rate of the product (Hcy) formed by reaction with DTNB. The inhibition assay was carried out by preincubating 4.7 μg of *Ld* AdoHcyase with 50, 100, and 200 μM concentrations of each compounds (inhibitors) for 1 h at 37 °C. NADH was used as the positive control. The IC_{50} value for each test compound was determined from the probit analysis after carrying the enzyme inhibition

dx.doi.org/10.1021/ci2005862 | *J. Chem. Inf. Model.* 2012, 52, 777−791

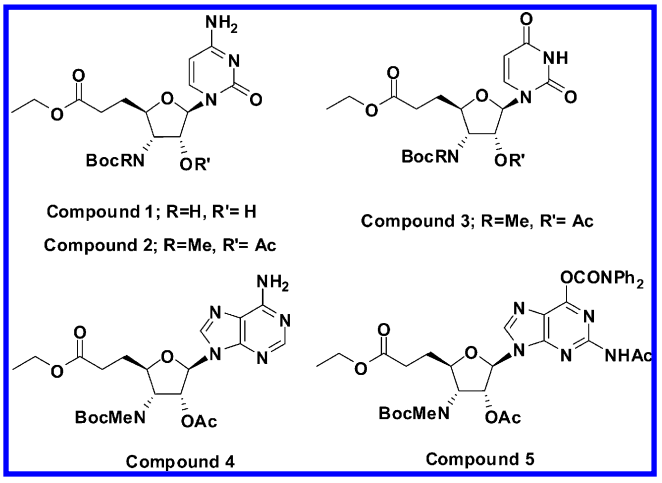


Figure 8. Structures of compounds **1**−**5**.

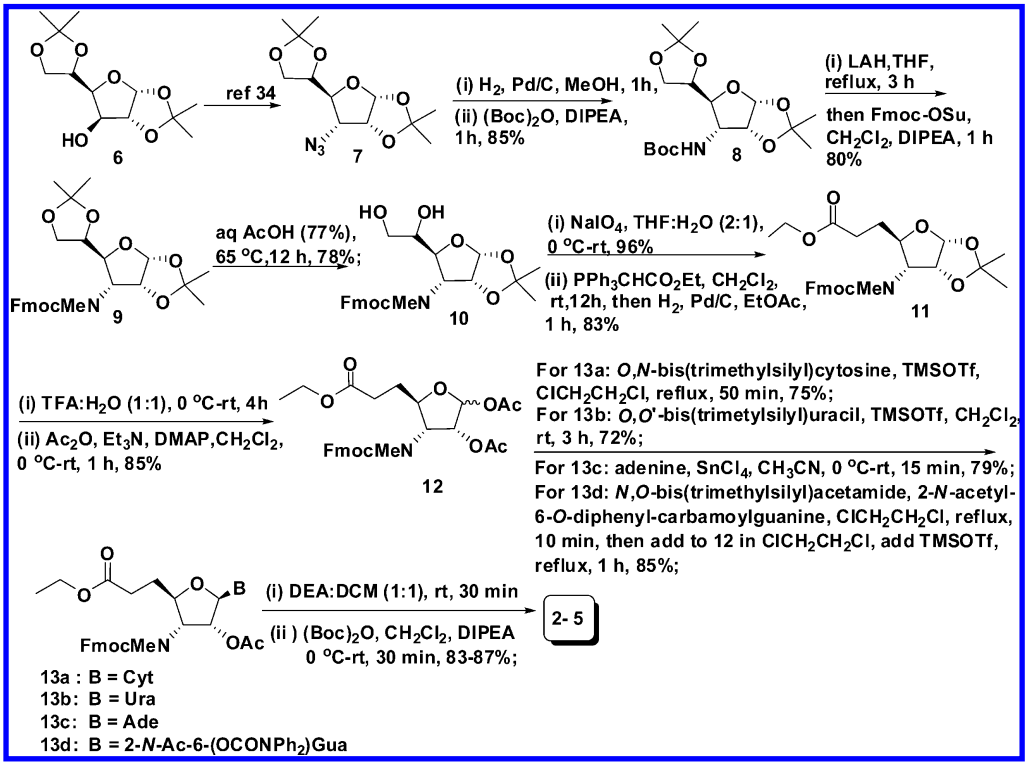
assays at three different concentrations 50, 100, and 200 μM. The dose response plot for each compound clearly demonstrated the maximum inhibitory activity at a concentration of 200 μM (Figure 9).

The details of enzyme inhibition assay have been described in the Experimental Section. The *Ld* AdoHcyase inhibitory activity of the five synthesized nucleoside analogues along with their Glide and GOLD docking scores is listed in Table 3, where compound **5** has shown a very good enzyme inhibitory effect (87.58%) with IC_{50} of 43 μM compared to the other nucleoside analogues against *Ld* AdoHcyase. The higher *Ld* AdoHcyase inhibitory activity of compound **5** was due to its higher mimicking tendency with NAD^+ (Figure 6C and D) which is further substantiated by its highest GOLD fitness score as well as higher Glide XP score as described in the SBVS section. It has also been observed that NADH showed >98% inhibitory activity over *Ld* AdoHcyase enzyme. This inhibition is probably due to competitive binding of NADH to the NAD^+ binding site on the enzyme as the loose binding NAD^+ is unique for the *Ld* AdoHcyase.

CONCLUSION

In the present study, we have successfully demonstrated the application of computational approaches for targeting the NAD^+ binding site of *Ld* AdoHcyase for the identification of novel *Ld* AdoHcyase inhibitors. This approach is significantly important as the NAD^+ is loosely bind in case of *Ld* AdoHcyase and can be easily targeted. These ligands will not affect the activity of *Human* AdoHcyase because the human enzyme binds NAD^+ tightly. Initially, the development and validation of the 3D structure of *Ld* AdoHcyase using the *Lm* AdoHcyase as a template was carried out. At the same time, cloning of the *Ld* AdoHcyase gene from clinical strains, its overexpression and purification were performed. In the next step, the model was used in combined docking and molecular dynamics studies to validate the binding site of NAD in *Ld*. The unique loose binding of NAD^+ at the NAD^+ binding site for the *Ld* AdoHcyase has been targeted for the discovery of novel *Ld* AdoHcyase inhibitors, and SBVS of a small focused library of 8000 nucleoside analogues at the NAD^+ binding site of *Ld* AdoHcyase was performed. The five active hits gained from the SBVS paradigm were further subjected for the molecular dynamics studies to ensure their stable binding. These hits were synthesized and evaluated for their *Ld* AdoHcyase inhibitory activity, which resulted in the identification and understanding of the mode of action of the five nucleoside analogues as *Ld* AdoHcyase inhibitors. The nonspecific binding of five synthesized hits was observed through docking studies. However, the hydrophobic cleft at the binding site surrounded

Scheme 1. Synthesis of Ribonucleoside Amino Acids 2−**5**



dx.doi.org/10.1021/ci2005862 | *J. Chem. Inf. Model.* 2012, 52, 777−791

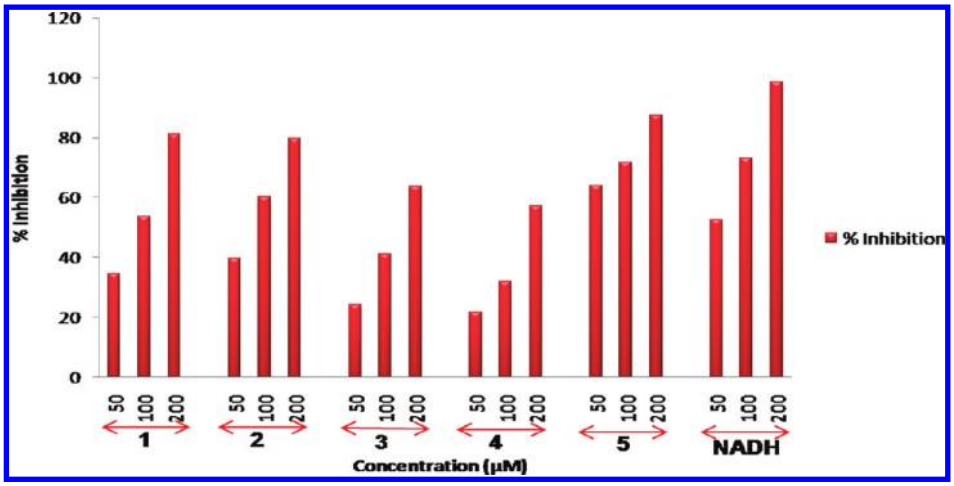


Figure 9. Dose–response graph for the five hits at three different concentrations.

Table 3. Inhibitory Activity of the Top Five Scoring Hits on *Ld* AdoHcyase

SN	compound name	% inhibitory effect[a] ± SD	IC_{50} (μM)	Glide XP score	GOLD fitness score
1	1	81.13 ± 3.69	95.4	−3.68	54.11
2	2	79.87 ± 3.50	79.6	−4.76	30.19
3	3	63.62 ± 4.51	144	−4.72	43.02
4	4	57.21 ± 3.49	171	−3.51	61.33
5	**5**	**87.58 ± 2.15**	**43**	**−4.39**	**65.96**
6	$NADH^+$	98.71 ± 3.46	35.5	−16.64	85.96

[a]% *Ld* AdoHcyase inhibitory effect at 200 μM. Data are averages of three measurements.

by the residues Val-243, Gly-219, and Glu-242 seemed to be important for ligand binding. It has also been observed that the interaction with any of the three consecutive threonine residues 156, 157, and 158 also play an important role in the ligand binding at the NAD^+ binding site. The reported *Ld* AdoHcyase homology model may be useful for the design and synthesis of novel *Ld* AdoHcyase inhibitors as well as for mutagenesis studies. The SBVS paradigm may serve as an alternative to experimental high-throughput screening in drug discovery, and *Ld* AdoHcyase could be considered as an important target for developing novel antileishmanial drugs. The structural overlap among the three protozoan AdoHcyase (*L.donovani*, *T.brucei*, and *T.cruzi*) and difference from that of the human enzyme suggest the possibility of selective drug design. These *Ld* AdoHcyase inhibitors may be further subjected for the ex vivo study on the amastigotes and promastigote stages of *L. donovani* for evaluation of their antiparasital effects. In this context the most potent inhibitor, **5**, may serve as a lead for developing more potent *Ld* AdoHcyase inhibitors as potential antileishmanial agents.

EXPERIMENTAL SECTION

Computational Tools. All calculations were performed on a Linux workstation equipped with four parallel Intel Xeon X5460 processors (2.8 GHZ) with 12 GB total RAM.

Sequence Alignment and Model Building. Our earlier submitted was retrieved from the UniProt database (accession code: D3K115). The NAD^+ bound crystal structures of *Lm* AdoHcyase protein (PDB ID: 3G1U) were taken into account as templates for homology modeling of *Ld* AdoHcyase. Multiple sequence alignment and model building was carried out using the Prime[18] module of the Schrodinger software package.[19]

Model Quality Assessments (MQA). These QA methods used in the present study were Procheck,[20] ProSA-web,[21a,b] QMEAN,[22] and Errat.[23] These methods have been used for the quality assessment and relevant model selection in the present study. Procheck is employed for geometric evaluations, whereas ProSA-web is used to evaluate the quality of consistency between the native fold and the sequence and to examine the energy of residue–residue interactions. The QMEAN value represents the model reliability in terms of *geometrical features whereas* the Errat value signifies the overall protein quality factor.

Ligand and Protein Preparation. Protein was prepared using the Protein Preparation Wizard implemented in the Schrodinger package using default options: bond orders were assigned, hydrogens were added, metals were treated, and water molecules 5 Å beyond heterogroups were deleted. Hydrogens were then optimized using the exhaustive sampling option, and the protein was minimized to an rmsd limit from the starting structure of 0.3 Å using the Impref module of Impact with the OPLS_2005 force field. The Ligands from the in-house library of nucleoside analogues were prepared using LigPrep with Epik to expand protonation and tautomeric states at 7.0 ± 2.0 pH units. Conformational search for each ligand was performed using MacroModel with the OPLS_2005 force field, water as solvent, and default parameters.

Parasites. The *L. donovani* clinical isolate 2001 procured from a patient admitted to the Kala-azar Medical Research Centre, Muzzafarpur (affiliated to Institute of Medical Sciences, BHU, Varanasi), India. The strains have also been maintained in hamsters through serial passage, i.e., from amastigote to amastigote. For bulk cultivation, promastigotes were grown in L-15 medium (Sigma-Aldrich) with L-glutamine, supplemented with 10% tryptose phosphate broth (Himedia), 0.1% gentamicin, and 10% FBS (Life Technologies).[39] Parasites were harvested after 3 days of culture.

Cloning and Overexpression and Purification of *Ld* AdoHcy Hydrolase (1314bp). *Ld* strain 2001 genomic DNA was isolated from 10^8 promastigotes, washed, and suspended in NET buffer (10 mM Tris-HCl (pH 7.5), 100 mM NaCl, and 1 mM EDTA) and incubated with proteinase K (1 mg/mL; Invitrogen Life Technologies) and 0.5% SDS at 50 °C for 4 h. Nucleic acidswere extracted by phenol:chloroform:isoamyl

dx.doi.org/10.1021/ci2005862 | J. Chem. Inf. Model. 2012, 52, 777–791

alcohol extraction and ethanol precipitation. Genomic DNA was spooled and subjected to RNase (100 ug/mL) treatment. The *Ld* AdoHcyase gene was amplified using Taq Polymerase (Sigma-Aldrich) lacking a 3′−5′ exonuclease activity. PCR was performed using *Ld* AdoHcyase specific primers (based on the *Lm* AdoHcyase gene sequence) forward, 5′ CATATGATGGCGGACTACAAGGTAAAGGACATC 3′ and reverse, 5′ AAGCTTGTAGCGGTAGTGGTCCGGCTTGAACG 3′ (NdeI and *Hin*dIII site underlined) in a Thermocycler (Bio-Rad) under conditions at one cycle of 96 °C for 4 min, 30 cycles of 94 °C for 1 min, 55 °C for 1 min, and 72 °C for 1 min 30 s, and finally one cycle of 72 °C for 10 min. Amplified PCR product was electrophoresed in agarose gel and eluted from the gel by GenElute Columns (Qiagen). Eluted product was cloned in pTZ57R/T (T/A) cloning vector (Fermentas) and transformed into competent DH5 cells. The transformants were screened for the presence of recombinant plasmids with the ADHT insert by gene-specific PCR under similar conditions as previously mentioned. Isolated positive clones were sequenced from Delhi University (New Delhi) and submitted to the NCBI (http://www.ncbi.nlm.nih. gov/nuccore/GU353334.1; accession no. GU353334). *Ld* AdoHcyase was further subcloned at the NdeI and *Hin*dIII site in bacterial pET28a (Novagen) The expression of *Ld* AdoHcyase checked in bacterial cells by transforming the *Ld* AdoHcyase +28a construct in *E. coli* Rosetta strain. The transformed cells were inoculated into 5 mL test tube culture medium (Luria−Bertani) and allowed to grow at 37 °C in a shaker at 220 rpm. Cultures in logarithmic phase (at OD600 of 0.5− 0.6) were induced for 14 h 1.0 mM isopropyl-D-thiogalactopyranoside (IPTG) at 26 °C After induction, cells were lysed in 5 sample buffer (0.313 M Tris-HCl (pH 6.8), 50% glycerol, 10% SDS, and 0.05% bromphenol blue, with 100 mM DTT) and whole cell lysate (WCL) were analyzed by 12% SDS-PAGE.[40] Uninduced control culture was analyzed in parallel. The separated proteins from the polyacrylamide gel were transferred onto a nitrocellulose membrane in a semidry blot apparatus (Amersham) as described earlier.[41] Membrane was incubated for 1 h in blocking buffer followed by a 2 h incubation at room temperature with mouse anti-His Ab (Novagen) as primary Ab (1/2500 dilution) and then incubated with goat antimouse HRP conjugate Ab (1/10 000; Bangalore Genie) for 1 h at room temperature. The blot was developed using DAB (sigma). *Ld* AdoHcyase was overexpressed in *E. coli* strain Rosetta (DE3) harboring the plasmid 28a+ *Ld* AdoHcyase. Cells were grown in Luria−Bertani (LB) broth with 40 μg/mL kanamycin and induced with 1 mM isopropyl β-thiogalactopyranoside (IPTG). Cultures were then grown at 26 °C for 20 h before harvest. Cells were harvested by centrifugation at 8000*g* for 10 min at 4 °C. Protein was purified using Ni-nitrilotriacetic acid agarose affinity chromatography.

Structure Based Virtual Screening (SBVS). A small focused virtual library of 8000 nucleoside analogues has been generated through CombiGlide[25] and screened through docking and scoring function of Glide and GOLD software for their competitive binding at NAD^+ binding site of *Ld* AdoHcyase. Default parameters (except non-hydrogen atoms 2−35, carbon 1−30, peroxides 0, and N-OS 0−9) were used for the SMART pattern filter algorithm implemented in the Discovery Studio 2.1 software package.[28] The LogP value was calculated in Discovery Studio 2.1. The MDS was run for a period of 100 ps, during which 10 samples of the complex were collected for the analyses. The MDS was run using settings as OPLS-2005 as force field, water as solvent model, constant dielectric as electrostatic treatment, Polak-Ribiere Conjugate Gradient (PRCG) as minimization method, maximum iterations of 500, molecular dynamics as dynamics method, simulation temperature of 300 K, time step of 1.5 fs, equilibration time of 10 ps, and simulations time of 100 ps.

The key interactions between these inhibitors and the binding site residues of *Ld* AdoHcyase are presented in the two-dimensional (2D) form using the MOE software package,[42] while the 3D molecular graphics were produced using the PyMol program.[43]

Chemistry. All the chemicals and reagents were obtained from Aldrich (USA) or Spectrochem (India) and were used without purification. Oxygen- or moisture-sensitive reactions were carried out in flame- or oven-dried glassware sealed with rubber septa under a dry nitrogen atmosphere. Similarly, sensitive liquids and solutions were transferred using a gastight syringe or cannula. Reactions were monitored by thin layer chromatography (TLC) carried out on 0.25 mm silica gel plates with UV light, I2, 7% ethanolic phosphomolybdic acid heat, and 2.5% ethanolic anisaldehyde (with 1% AcOH and 3.3% conc H_2SO_4) heat as developing agents. Silica gels of 60−120 and 200−300 mesh were used for chromatographic separation. Infrared spectra were recorded as neat or in KBr on a Fourier transform spectrometer (FTIR). ^{1}H NMR spectra are reported in ppm using tetramethylsilane (0.00 ppm) or solvent ($CDCl_3$: 7.26 ppm) as an internal standard in 200, 300, 400, and 500 MHz instruments at 30 °C. Data are reported as s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, b = broad, and coupling constant(s) are in hertz. Proton-decoupled ^{13}C NMR spectra were recorded at 75, 100, and 150 MHz instruments as indicated beneath of the spectra and are reported in parts per million using the solvent as an internal standard ($CDCl_3$: 77.00 ppm). Unless noted otherwise on the spectra, NMR spectra were recorded in $CDCl_3$. Mass spectra were obtained under electron spray ionization (ESI) and high resolution mass spectrometer (HRMS-ESI) techniques. Optical rotations were measured using Autopol III manufactured by Rudolph using a sodium (546 nm, D line) lamp and are reported as follows: $[\alpha]_\lambda^{\mathrm{ToC}}$ (c = g/100 mL, solvent). Melting points were recorded using a Fisher-Johns melting point apparatus and are uncorrected. Analytical HPLC analyses were performed on a Merck Hitachi HPLC system equipped with a 5 μ Thermo C-8 column (250 × 4.6 mm) and a UV/vis detector setting of λ = 254 nm to ensure the purity of the tested compounds. All compounds were eluted with isocratic elution (water:ACN; 20:80) to verify the purity of active virtual hits. HPLC analysis confirmed >95% purity for the tested compounds.

Synthesis of ***8****.* To a solution of the azide **7** (15 g, 52.63 mmol) in ethanol (105 mL), 10% Pd−C (1.5 g) was added and the mixture was hydrogenated at room temperature (rt) using an H_2-filled balloon for 1 h. The solution was then filtered through a short pad of Celite, and the filter cake was washed with ethanol (3 × 20 mL). The filtrate and washings were combined and concentrated in vacuo. The reaction mixture was then cooled to 0 °C and to the same reaction mixture $(Boc)_2O$ (15 mL, 63.15 mmol) was added followed by DIPEA (10.1 mL, 63.15 mmol) in CH_2Cl_2. The reaction mixture was brought to room temperature and stirred at the same temperature for 30 min. It was quenched with saturated aqueous NH_4Cl solution, extracted with EtOAc, washed with brine, dried (Na_2SO_4), and concentrated in vacuo. Purification by column chromatography (SiO_2, 6−8% EtOAc in petroleum ether eluant) afforded **8**

(16.1 g, 85%) as a colorless oil. R_f = 0.45 (SiO_2, 15% EtOAc in petroleum ether). Specific rotation $[\alpha]_D^{25}$ = +18.2 (*c* 0.61, $CHCl_3$). IR (neat): ν_{max} 3758, 3414, 2925, 1704, 1454, 1376, 1166, 1057, 753 cm^{-1}. ^1H NMR (300 MHz, $CDCl_3$): δ 5.79 (d, *J* = 3.5 Hz, 1H), 4.96 (d, *J* = 8.8 Hz, 1H), 4.58 (t, *J* = 4.0 Hz, 1H), 4.24 (m, 1H), 4.12−3.77 (m, 4H), 1.60−1.25 (m, 21H). ^{13}C NMR (75 MHz, $CDCl_3$): δ154.96, 112.39, 109.54, 103.93, 79.11, 78.70, 77.27, 75.67, 65.35, 54.63, 28.26, 25.41, 26.36, 26.55. MS (ESIMS): *m*/*z* (%): 360 (100) $[M + H]^+$.

Synthesis of ***9***. To a solution of **8** (15 g, 41.78 mmol) in THF (125 mL), LAH (3.45 g, 83.56 mmol) was added at 0 °C, and the reaction mixture was refluxed for 3 h. The reaction was carefully quenched by addition of saturated Na_2SO_4 at 0 °C. The resulting suspension was stirred vigorously for 30 min, filtered through Celite, dried (Na_2SO_4), and concentrated in vacuo. The product was directly used for the next reaction.

To a stirred solution of amine in CH_2Cl_2 (125 mL) at 0 °C under nitrogen atmosphere was added DIPEA (8.62 mL, 50.13 mmol). After 10 min, Fmoc-OSu (21.14 g, 62.67 mmol) was added at 0 °C. The reaction mixture was warmed to room temperature and stirred for 1 h at the same temperature. It was quenched with saturated aqueous NH_4Cl solution, extracted with EtOAc, washed with brine, dried (Na_2SO_4) and concentrated in vacuo. Purification by column chromatography (SiO_2, 6−8% EtOAc in petroleum ether eluant) afforded **9** (16.54 g, 80%). The product contains two rotamers. R_f = 0.3 (SiO_2, 10% EtOAc in petroleum ether). Specific rotation $[\alpha]_D^{25}$ = +23.8 (*c* 0.58, $CHCl_3$). IR (neat): ν_{max} 3438, 2928, 2375, 2132, 1704, 1450, 1376, 1165, 1034, 757 cm^{-1}. ^1H NMR (300 MHz, $CDCl_3$): δ 7.82−7.50 (m, 4H), 7.49−7.29 (m, 4H), 5.74 (d, *J* = 3.4 Hz, 0.6H), 5.47 (d, *J* = 3.4 Hz, 0.4H), 4.77−4.21 (m, 5H), 4.19−3.66 (m, 4H), 3.08 (s, 1.7H), 2.96 (s, 1.3H), 1.65−1.15 (m, 12H). ^{13}C NMR (75 MHz, $CDCl_3$): δ156.81, 155.68, 144.13, 144.01, 143.83, 141.44, 141.24, 127.63, 127.52, 127.15, 126.90, 125.01, 124.58, 119.92, 119.77, 112.65, 112.44, 109.63, 109.58, 103.77, 103.43, 80.25, 76.64, 76.31, 74.87, 74.78, 67.71, 66.85, 66.69, 66.08, 59.42, 59.85, 47.20, 32.42, 26.58, 26.44, 26.27, 26.08, 25.09, 25.04. MS (ESIMS): *m*/*z* (%): 496 (50) $[M + H]^+$.

Synthesis of ***10***. A solution of **9** (15 g, 30.3 mmol) in 77% aq AcOH (21 mL, 0.7 mL/mmol) was stirred at 65 °C for 3 h. The reaction mixture was then azeotroped with CH_2Cl_2 (3 × 20 mL). The reaction mixture then cooled to 0 °C and slowly quenched with solid $NaHCO_3$ until the effervescence ceased. The reaction mixture was diluted with water (1 × 10 mL followed by extraction with CH_2Cl_2 (3 × 100 mL), washed with brine, dried (Na_2SO_4), and concentrated in vacuo. Purification by column chromatography (SiO_2, 45 to 50% EtOAc in petroleum ether eluant) afforded **10** (9.65 g, 70%). The product contains two rotamers of diol. R_f = 0.3 (SiO_2, 50% EtOAc in petroleum ether). Specific rotation $[\alpha]_D^{25}$ = +17.1 (*c* 0.56, $CHCl_3$). IR (neat): ν_{max} 3758, 3422, 2922, 1685, 1450, 1326, 1162, 1018, 749 cm^{-1}. ^1H NMR (300 MHz, $CDCl_3$): δ 7.74 (d, *J* = 7.4 Hz, 2H), 7.58 (d, *J* = 6.9 Hz, 2H), 7.43−7.22 (m, 4H), 5.72 (m, 0.65H), 5.48 (m, 0.35H), 4.74−4.10 (m, 5H), 3.88 (m, 1H), 3.74−3.33 (m, 3H), 3.19 (m, 1H), 3.03 (s, 1.9H), 2.91 (s, 1.1H), 2.61 (m, 1H), 1.62−1.39 (m, 3H), 1.37 (m, 3H). ^{13}C NMR (75 MHz, $CDCl_3$): δ157.24, 155.99, 144.15, 143.99, 143.85, 141.36, 127.76, 127.13, 125.13, 125.04, 124.86, 124.72, 120.01, 119.90, 112.81, 112.67, 103.73, 103.39, 80.44, 75.49, 75.33, 72.51, 72.07, 67.99, 67.15, 63.24, 62.81, 58.03, 47.27, 32.56, 30.93, 26.70, 26.55, 26.22. MS (ESIMS): *m*/*z* (%): 456 (60) $[M + H]^+$.

Synthesis of ***11***. $NaIO_4$ (11.27 g, 52.74 mmol) was added portionwise to a cooled solution of **10** (8.0 g, 17.58 mmol) in THF:H_2O (52 mL, 2:1). The mixture was stirred for 30 min. The inorganic salts were filtered off, and washed several times with THF (3 × 25 mL). The combined organic phases were concentrated under vacuo to yield the aldehyde (7.13 g, 96%) as a light yellow syrup. The crude aldehyde was used without further purification for the next step.

Aldehyde (7.13 g, 16.85 mmol) was dissolved in dichloromethane (50 mL), and (carboethoxymethylene)-triphenylphosphorane (11.79 g, 33.71 mmol) was added portionwise under a nitrogen atmosphere at rt. The reaction mixture was stirred at rt for 12 h, and dichloromethane was evaporated under reduced pressure. Purification by column chromatography (SiO_2, 9−10% EtOAc in petroleum ether eluant) afforded ester (7.89 g, 95%) which was directly subjected to hydrogenation. To a solution of the ester (7.0 g, 14.19 mmol) in EtOAc (42 mL), 10% Pd/C (700 mg) was added and the mixture was hydrogenated at rt using a H_2-filled balloon for 4−6 h. The solution was then filtered through a short pad of Celite, and the filter cake was washed with EtOAc (3 × 20 mL). The filtrate and washings were combined and concentrated in vacuo. Purification by coloumn chromatography (SiO_2, 10% EtOAc in petroleum ether eluant) afforded pure compound **11** (5.83 g, 83%). The product contains two rotamers of ester. R_f = 0.4 (SiO_2, 20% EtOAc in petroleum ether). Specific rotation $[\alpha]_D^{25}$ = +14.2 (*c* 0.53, $CHCl_3$). IR (neat): ν_{max} 3412, 2922, 2375, 2140, 1649, 1440, 1326, 1166, 1026, 742 cm^{-1}. ^1H NMR (300 MHz, $CDCl_3$): δ7.76 (d, *J* = 6.7 Hz, 2H), 7.66−7.48 (m, 2H), 7.46−7.23 (m, 4H), 5.74 (m, 0.6H), 5.45 (m, 0.4H), 4.68 (m, 1H), 4.43 (t, *J* = 7.0 Hz, 1.3H), 4.35−4.21 (m, 1.7H), 4.19−4.03 (m, 3H), 3.80 (m, 1H), 3.05 (s, 1.7H), 2.91 (s, 1.3H), 2.59−2.23 (m, 2H), 2.21−1.65 (m, 2H), 1.61−1.14 (m, 9H). ^{13}C NMR (75 MHz, $CDCl_3$): δ173.31, 173.07, 157.05, 156.12, 144.38, 144.09, 141.48, 127.88, 127.78, 127.42, 127.22, 125.21, 125.14, 124.78, 124.61, 120.15, 120.01, 119.92, 112.46, 112.25, 103.86, 103.54, 79.95, 73.61, 73.70, 68.02, 67.07, 61.18, 60.57, 47.44, 32.05, 31.87, 30.24, 30.04, 29.83, 27.39, 27.25, 26.60, 26.45, 26.25, 26.17, 14.36. MS (ESIMS): *m*/*z* (%): 495 (80) $[M^+]^+$.

Synthesis of ***12***. An ice-cold solution of **8** (5.0 g, 10.39 mmol) in TFA:H_2O (26 mL, 1:1) was stirred for 15 min at 0 °C and at rt for 4 h. The reaction was then cooled to 0 °C and slowly quenched with solid $NaHCO_3$ until the effervescence ceased. The reaction mixture was then diluted with water and extracted with CH_2Cl_2. The organic layer was washed with brine, dried (Na_2SO_4), and concentrated in vacuo to give lactol as a thick liquid. The crude product was directly subjected to acetylation.

To a stirred solution of the lactol in CH_2Cl_2 (32 mL) at 0 °C under a nitrogen atmosphere was added Et_3N (4.35 mL, 31.18 mmol). After 10 min, Ac_2O (2.35 mL, 24.94 mmol) followed by DMAP (254 mg, 2.078 mmol) was added at 0 °C. The reaction mixture was warmed to room temperature and stirred for 1 h at same temperature. It was quenched with saturated aqueous NH_4Cl solution, extracted with EtOAc, washed with brine, dried (Na_2SO_4), and concentrated in vacuo. Purification by column chromatography (SiO_2, 15−20% EtOAc in petroleum ether eluant) afforded an anomeric mixture **12** (4.59 g, 82%). The product contains two rotamers of diacetate. R_f = 0.4 (SiO_2, 30% EtOAc in petroleum ether). Specific rotation $[\alpha]_D^{25}$ = +16.8 (*c* 0.6, $CHCl_3$). IR (neat): ν_{max} 3780, 3462, 2985, 1742, 1447, 1374, 1242, 1048, 936, 783 cm^{-1}. ^1H

dx.doi.org/10.1021/ci2005862 | J. Chem. Inf. Model. 2012, 52, 777−791

NMR (300 MHz, $CDCl_3$): δ7.76 (d, J = 7.5 Hz, 2H), 7.64−7.49 (m, 2H), 7.44−7.21 (m, 4H), 6.37 (d, J = 5.0 Hz, 0.26H), 6.22 (d, J = 5.0 Hz, 0.24H), 6.03 (m, 0.5H), 5.16 (m, 1H), 4.89−4.20 (m, 5H), 4.12 (q, J = 7.0 Hz, 2H), 2.99−2.85 (m, 3H), 2.53−2.23 (m, 2H), 2.15−1.59 (m, 8H), 1.25 (t, J = 7.0 Hz, 3H). ^{13}C NMR (75 MHz, $CDCl_3$): δ172.86, 169.32, 169.04, 156.92, 156.19, 144.09, 143.88, 141.44, 127.82, 127.15, 124.92, 124.67, 120.08, 98.29, 93.91, 93.69, 77.97, 77.38, 76.47, 70.95, 70.74, 67.79, 67.43, 60.59, 59.16, 58.17, 57.86, 47.33, 31.72, 31.57, 30.10, 29.84, 29.69, 28.78, 28.50, 21.20, 20.86, 20,42, 14.26. MS (ESIMS): m/z (%): 562 (100) $[M + Na]^+$.

*General Procedure **A** (Coupling of Cytosine and Uracil with **12**).* A suspension of nucleobase and $(NH_4)_2SO_4$ (trace) in HMDS (2 mL/mmol) was stirred at reflux under N_2 atmosphere until a clear solution was formed. Volatiles were evaporated under reduced pressure, and toluene was added and coevaporated several times. The residue was dried in vacuum, and a solution of the carbohydrate derivative in 1,2-dichloroethane (under N_2) was added. TMSOTf (1.2 equiv) was added at 0 °C, and the solution was brought to rt and stirred for 2 h at the same temperature for **13b**, refluxed for 1 h at 80 °C in case of **13a**. The reaction mixture was cooled to 0 °C, and CH_2Cl_2 (10 mL) and saturated aqueous $NaHCO_3$ (5 mL) were added. The organic layer was separated, and the aqueous layer was extracted with CH_2Cl_2 (3 × 100 mL). The combined organic layers were washed with brine and dried (Na_2SO_4). The solvent was removed under reduced pressure, and the residue was purified by silica gel chromatography.

*Synthesis of **13a**.* Procedure **A**, with **12** (200 mg, 0.37 mmol), cytosine (95 mg, 0.85 mmol), trimethylsilyltrifluoromethanesulfonate (0.13 mL, 0.74 mmol), and $ClCH_2CH_2Cl$ (5 mL), chromatography (SiO_2, 5−6% MeOH in $CHCl_3$ eluant) afforded **13a** (164 mg, 75%). The product contains two rotamers. R_f = 0.40 (SiO_2, 10% MeOH in $CHCl_3$). Specific rotation $[\alpha]_D^{25}$ = +18.0 (c 0.54, $CHCl_3$). IR (neat): ν_{max} 3342, 3018, 2367, 1653, 1491, 1375, 1220, 1065, 763 cm^{-1}. ^{1}H NMR (300 MHz, $CDCl_3$): δ7.82−7.67 (m, 2H), 7.57 (d, J = 7.0 Hz, 2H), 7.45−7.26 (m, 4H), 7.19 (d, J = 7.4 Hz, 0.6H), 6.97 (d, J = 7.4 Hz, 0.4H), 5.94 (d, J = 7.4 Hz, 0.4H), 5.87 (d, J = 7.4 Hz, 0.59H), 5.68 (m, 1H), 5.43 (m, 0.56H), 5.06 (m, 0.43H), 4.66−4.20 (m, 4H), 4.19−3.95 (m, 3H), 4.12 (q, J = 7.0 Hz, 2H), 2.98−2.83 (m, 3H), 2.55−2.36 (m, 2H), 2.20−1.76 (m, 5H), 1.24 (t, J = 7.0 Hz, 3H). ^{13}C NMR (75 MHz, $CDCl_3$): δ173.17, 172.91, 169.47, 169.30, 166.28, 166.11, 156.74, 156.26, 155.63, 144.02, 143.82, 141.45, 141.26, 127.86, 127.37, 127.20, 124.94, 124.85, 120.10, 96.25, 92.01, 90.74, 77.36, 75.78, 67.91, 67.20, 60.66, 59.74, 47.33, 32.50, 32.01, 30.03, 29.77, 29.43, 28.40, 20.90, 14.29. MS (ESIMS): m/z (%): 613 (70) $[M + Na]^+$.

*Synthesis of **13b**.* Procedure **A**, **12** (200 mg, 0.37 mmol), uracil (124 mg, 1.11 mmol), trimethylsilyltrifluoromethanesulfonate (0.13 mL, 0.74 mmol), and $ClCH_2CH_2Cl$ (5 mL), chromatography (SiO_2, 2−3% MeOH in $CHCl_3$ eluant) afforded **13b** (158 mg, 72%). The product contains two rotamers. R_f = 0.4 (SiO_2, 5% MeOH in $CHCl_3$). Specific rotation $[\alpha]_D^{25}$ = +8.33 (c 0.5, $CHCl_3$). IR (neat): ν_{max} 3602, 3020, 2364, 1695, 1452, 1217, 1109, 936, 766 cm^{-1}. ^{1}H NMR (300 MHz, $CDCl_3$): δ7.81−7.68 (m, 2H), 7.57 (d, J = 7.4 Hz, 2H), 7.44−7.20 (m, 5H), 6.82 (d, J = 8.0 Hz, 0.42H), 5.89 (d, J = 3.5 Hz, 0.6H), 5.78 (d, J = 8.0 Hz, 1H), 5.49 (m, 0.4H), 5.31 (m, 0.6H), 4.74 (m, 1H), 4.57−3.77 (m, 6H), 4.13 (q, J = 6.9 Hz, 2H), 2.96−2.80 (m, 3H), 2.55−2.27 (m, 2H), 2.13−1.76 (m, 5H), 1.25 (t, J = 6.9 Hz, 3H). ^{13}C NMR (75 MHz, $CDCl_3$): δ172.90, 172.77, 169.53, 163.18, 163.06, 156.63, 156.10, 150.17, 149.84, 143.83, 141.46, 140.45, 140.12, 127.83, 127.18, 124.92, 124.72, 120.10, 103.44, 103.28, 90.24, 89.19, 75.52, 74.96, 67.91, 66.59, 60.76, 60.25, 47.64, 47.32, 33.01, 30.61, 29.83, 28.61, 28.13, 20.72, 14.29. MS (ESIMS): m/z (%): 592 (60) $[M + H]^+$.

*Synthesis of **13c**.* Adenine (55.0 mg, 0.41 mmol) was added to a solution of **12** (200 mg, 0.37 mmol) in dry CH_3CN (2 mL) under a nitrogen atmosphere. The reaction mixture was cooled to 0 °C, and $SnCl_4$ (0.10 mL, 0.74 mmol) in dry CH_3CN (1 mL) was added. The reaction mixture was warmed to room temperature and stirred for 15 min, and the cloudy solution turned clear. The reaction mixture was cooled to 0 °C, CH_2Cl_2 (10 mL) and saturated aqueous $NaHCO_3$ (5 mL) were added. The organic layer was separated, and the aqueous layer was extracted with CH_2Cl_2 (3 × 100 mL). The combined organic layers were washed with brine, dried (Na_2SO_4), concentrated under reduced pressure, and the residue was purified by silica gel chromatography (SiO_2, 2−3% MeOH in $CHCl_3$ eluant) which afforded **13c** (180 mg, 79%). The product contains two rotamers. R_f = 0.45 (SiO_2, 5% MeOH in $CHCl_3$). Specific rotation $[\alpha]_D^{25}$ = −0.6 (c 0.56, $CHCl_3$). IR (neat): ν_{max} 3478, 3109, 2371, 1704, 1639, 1474, 1218, 1065, 763, 669 cm^{-1}. ^{1}H NMR (300 MHz, $CDCl_3$): δ8.36 (s, 0.5H), 8.12 (s, 0.5H), 7.92 (s, 0.4H), 7.84−7.48 (m, 4.6H), 7.47−7.15 (m, 5H), 6.17−5.92 (m, 2H), 5.75 (m, 1H), 4.82−4.20 (m, 5H), 4.11 (q, J = 7.0 Hz, 2H), 3.04−2.85 (m, 3H), 2.54−2.35 (m, 2H), 2.16−1.74 (m, 5H), 1.23 (t, J = 7.0 Hz, 3H). ^{13}C NMR (75 MHz, $CDCl_3$): δ172.96, 172.83, 169.56, 156.48, 156.28, 155.45, 152.90, 143.77, 143.61, 141.38, 141.30, 139.29, 127.75, 127.09, 124.83, 124.53, 120.03, 88.24, 87.69, 78.05, 75.58, 67.73, 67.20, 60.84, 60.57, 47.38, 47.27, 33.75, 29.88, 29.68, 28.80, 28.05, 20.75, 20.65, 14.17. MS (ESIMS): m/z (%): 615 (50) $[M + H]^+$.

*Synthesis of **13d**.* Bis(trimethylsilyl)acetamide (BSA) (0.22 mL, 0.89 mmol) was added to a solution of **2**-*N*-acetyl-6-*O*-diphenylcarbamoylguanine (173 mg, 0.44 mmol) in $ClCH_2CH_2Cl$ (2.4 mL). The mixture was refluxed for 10 min, cooled to rt, and added to a solution of **13** (200 mg, 0.37 mmol) in $ClCH_2CH_2Cl$ (1.2 mL) at 0 °C. Trimethylsilyl trifluoromethanesulfonate (0.08 mL, 0.44 mmol) was added dropwise at 0 °C. The brown solution was refluxed for 1 h. The reaction mixture was cooled to 0 °C, and CH_2Cl_2 (10 mL) and saturated aqueous $NaHCO_3$ (5 mL) were added. The organic layer was separated and the aqueous layer was extracted with CH_2Cl_2 (3 × 50 mL). The combined organic layers were washed with brine and dried (Na_2SO_4). Solvent was removed under reduced pressure, and the residue was purified by silica gel chromatography (SiO_2, 2−3% MeOH in $CHCl_3$ eluant) afforded **13d** (259 mg, 80%). The product contains two rotamers. R_f = 0.5 (SiO_2, 5% MeOH in $CHCl_3$). Specific rotation $[\alpha]_D^{25}$ = −5.5 (c 0.37, $CHCl_3$); IR (neat): ν_{max} 3748, 3424, 2987, 2370, 2112, 1686, 1377, 1218, 1028, 769 cm^{-1}. ^{1}H NMR (300 MHz, $CDCl_3$): δ8.53 (m, 0.6H), 8.12 (s, 0.4H), 7.98 (s, 0.5H), 7.85−7.67 (m, 2H), 7.66−7.16 (m, 16H), 5.91 (m, 1H), 5.21−4.37 (m, 4H), 4.34−4.01 (m, 3H), 3.05−2.85 (m, 3H), 2.58−2.22 (m, 6H), 2.15−1.74 (m, 5H), 1.23 (t, J = 7.0 Hz, 3H). ^{13}C NMR (75 MHz, $CDCl_3$): δ172.97, 169.62, 169.56, 169.48, 156.55, 152.31, 150.31, 143.82, 143.76, 143.01, 141.80, 141.77, 141.40, 129.21, 129.07, 127.78, 127.51, 127.09, 126.98, 126.49, 124.91, 124.83, 121.35, 120.61, 120.01, 88.09, 75.62, 67.75, 60.63, 47.30, 47.27, 29.68, 28.64, 25.12, 24.91, 20.67, 14.17; MS (ESIMS): m/z (%): 868 (100) $[M + H]^+$.

dx.doi.org/10.1021/ci2005862 | J. Chem. Inf. Model. 2012, 52, 777−791

General Procedure B. Dichloromethane and diethylamine was added to the carbohydrate derivative (1 equiv) in a 1:1 ratio under a nitrogen atmosphere at 0 °C. The reaction mixture was brought to room temperature and stirred at the same temperature for 30 min. The solvent was removed under reduced pressure. The reaction mixture was then azeotroped with CH_2Cl_2 (3 × 10 mL). The reaction mixture then cooled to 0 °C, and to the same reaction mixture, $(Boc)_2O$ (1.2 equiv) was added followed by DIPEA (1.2 equiv) and DMAP (0.1 equiv) in CH_2Cl_2. The reaction mixture was brought to room temperature and stirred at the same temperature for 30 min. Solvent was removed under reduced pressure, and the residue was purified by silica gel chromatography.

Synthesis of 2. Procedure **B**, **13a** (100 mg, 0.17 mmol), DCM:DEA (3 mL, 1.5:1.5), $(Boc)_2O$ (0.05 mL, 0.20 mmol), DIPEA (0.035 mL, 0.20 mmol), DMAP (2.0 mg, 0.017 mmol) in CH_2Cl_2 (3 mL), chromatography (SiO_2, 5−6% MeOH in $CHCl_3$ eluant) afforded **2** (68.2 mg, 86%). The product contains two rotamers. R_f = 0.4 (SiO_2, 10% MeOH in $CHCl_3$). Specific rotation $[\alpha]_D^{25}$ = +3.94 (*c* 0.29, $CHCl_3$). IR (neat): ν_{max} 3436, 2930, 2369, 1728, 1649, 1398, 1217, 1129, 770 cm^{-1}. ^1H NMR (400 MHz, $CDCl_3$): δ7.32 (m, 1H), 6.24 (m, 1H), 5.84 (m, 1H), 5.56 (m, 0.4H), 5.25 (m, 0.6H), 4.75 (m, 1H), 4.32 (m, 1H), 4.19 (q, *J* = 6.9 Hz, 2H), 3.05−2.95 (m, 3H), 2.59−2.41 (m, 2H), 2.23−1.75 (m, 5H), 1.51−1.47 (m. 9H), 1.24 (t, *J* = 6.9 Hz, 3H). ^{13}C NMR (75 MHz, $CDCl_3$): δ172.42, 170.62, 170.46, 166.12, 160.81, 152.50, 147.01, 142.04, 133.25, 105.51, 105.20, 101.21, 93.24, 91.36, 91.25, 80.87, 80.35, 79.74, 79.32, 78.90, 78.30, 77.18, 76.33, 67.90, 61.60, 61.41, 56.01, 53.73, 53.44, 47.20, 42.01, 31.96, 29.64, 29.36, 27.61, 27.36, 12.01, 11.72. MS (ESIMS): *m*/*z* (%): 491 (100) $[M + Na]^+$. HRMS (ESIMS): calcd for $C_{21}H_{32}N_4O_8Na$ $[M + Na]^+$ 491.2118, found 491.2098.

Synthesis of 3. Procedure **B**, **13b** (100 mg, 0.17 mmol), DCM:DEA (3 mL, 1.5:1.5), $(Boc)_2O$ (0.05 mL, 0.20 mmol), DIPEA (0.035 mL, 0.20 mmol), DMAP (2.0 mg, 0.017 mmol) in CH_2Cl_2 (3 mL), chromatography (SiO_2, 2−3% MeOH in $CHCl_3$ eluant) afforded **3** (67.1 mg, 84%). The product contains two rotamers. R_f = 0.3 (SiO_2, 5% MeOH in $CHCl_3$). Specific rotation $[\alpha]_D^{25}$ = +12.2 (*c* 0.66, $CHCl_3$). IR (neat): ν_{max} 3020, 2367, 1725, 1683, 1448, 1217, 1151, 765 cm^{-1}. ^1H NMR (300 MHz, $CDCl_3$): δ7.21 (m, 1H), 5.99−5.59 (m, 2H), 5.44 (m, 0.4H), 5.18 (m, 0.6H), 4.84−4.23 (m, 2H), 4.13 (q, *J* = 6.9 Hz, 2H), 3.02 (s, 2.4H), 2.95 (s, 0.6H), 2.52−2.45 (m, 2H), 2.25−1.75 (m, 5H), 1.49−1.46 (m. 9H), 1.25 (t, *J* = 6.9 Hz, 3H). ^{13}C NMR (75 MHz, $CDCl_3$): δ172.76, 172.56, 172.01, 171.90, 160.08, 160.02, 152.20, 151.72, 148.14, 147.94, 147.39, 141.88, 139.17, 104.24, 102.99, 102.64, 93.03, 89.69, 89.17, 87.17, 86.84, 84.17, 83.74, 76.93, 61.33, 60.66, 57.61, 34.22, 34.05, 31.87, 30.55, 29.87, 29.79, 29.64, 29.60, 28.30, 28.20, 27.57, 27.40, 22.64, 22.06, 21.71, 14.16, 14.07. MS (ESIMS): *m*/*z* (%): 492 (100) $[M + Na]^+$. HRMS (ESIMS): calcd for $C_{21}H_{31}N_3O_9$ $[M + Na]^+$ 492.1958, found 492.1939.

Synthesis of 4. Procedure **B**, **13c** (100 mg, 0.16 mmol), DCM:DEA (3 mL, 1.5:1.5), $(Boc)_2O$ (0.05 mL, 0.19 mmol), DIPEA (0.033 mL, 0.19 mmol), DMAP (2.0 mg, 0.016 mmol) in CH_2Cl_2 (3 mL), chromatography (SiO_2, 2−3% MeOH in $CHCl_3$ eluant) afforded **4** (70.4 mg, 87%). The product contains two rotamers. R_f = 0.45 (SiO_2, 5% MeOH in $CHCl_3$). Specific rotation $[\alpha]_D^{25}$ = −2.8 (*c* 0.82, $CHCl_3$). IR (neat): ν_{max} 3018, 2370, 1744, 1653, 1458, 1217, 935, 766 cm^{-1}. ^1H NMR (300 MHz, $CDCl_3$): δ8.86 (m, 1H), 8.15 (m, 1H), 6.22 (d, *J* = 3.4 Hz, 0.6H), 5.99 (d, *J* = 2.6 Hz, 0.4H), 5.83 (m, 0.42H), 5.67 (m, 0.58H), 5.09 (m, 1H), 4.52 (m, 1H), 4.11 (q, *J* = 7.0 Hz, 2H), 3.08 (s, 2.2H), 2.98 (s, 0.8H), 2.55−2.38 (m, 2H), 2.24−1.82 (m, 5H), 1.51−1.47 (m. 9H), 1.23 (t, *J* = 7.0 Hz, 3H). ^{13}C NMR (75 MHz, $CDCl_3$): δ172.86, 172.52, 171.97, 171.91, 152.70, 152.30, 152.08, 151.94, 150.90, 150.47, 150.37, 144.42, 143.45, 129.70, 129.30, 88.33, 88.01, 84.30, 84.01, 83.76, 83.64, 78.01, 77.95, 77.85, 77.32, 61.52, 60.58, 60.53, 58.65, 35.03, 30.51, 29.71, 29.61, 28.63, 27.76, 27.74, 27.53, 22.23, 21.86, 14.12. MS (ESIMS): *m*/*z* (%): 493 (80) $[M + H]^+$. HRMS (ESIMS): calcd for $C_{22}H_{33}N_6O_7$ $[M + H]^+$ 493.2410, found 493.2390.

Synthesis of 5. Procedure **B**, **13d** (100 mg, 0.11 mmol), DCM:DEA (3 mL, 1.5:1.5), $(Boc)_2O$ (0.04 mL, 0.14 mmol), DIPEA (0.024 mL, 0.19 mmol), DMAP (1.4 mg, 0.011 mmol) in CH_2Cl_2 (3 mL), chromatography (SiO_2, 2−3% MeOH in $CHCl_3$ eluant) afforded **5** (71.2 mg, 83%). The product contains two rotamers. R_f = 0.4 (SiO_2, 5% MeOH in $CHCl_3$). Specific rotation $[\alpha]_D^{25}$ = −4.87 (*c* 0.24, $CHCl_3$). IR (neat): ν_{max} 3755, 3432, 2368, 1630, 1368, 1219, 1156, 770 cm^{-1}. ^1H NMR (400 MHz, $CDCl_3$): δ8.61 (brs, 1H), 8.21 (m, 1H), 7.45−7.11 (m, 10H), 5.62−5.59 (m, 2H), 4.42 (m, 1H), 4.21−4.03 (m, 3H), 4.13 (q, *J* = 7.0 Hz, 2H), 3.11 (s, 2.4H), 3.01 (s, 0.6H), 2.53−2.39 (m, 2H), 2.35−1.75 (m, 8H), 1.51−1.44 (m, 9H), 1.24 (t, *J* = 7.0 Hz, 3H). ^{13}C NMR (75 MHz, $CDCl_3$): δ173.21, 173.17, 172.48, 155.77, 152.86, 148.11, 147.51, 141.51, 141.43, 129.42, 129.38, 127.62, 126.50, 126.46, 90.05, 88.50, 84.01, 83.76, 78.65, 77.22, 76.41, 60.63, 57.10, 56.90, 34.01, 31.90, 29.67, 29.47, 27.99, 27.86, 27.74, 27.61, 22.66, 22.61, 22.35, 14.18, 14.09. MS (ESIMS): *m*/*z* (%): 746 (80) $[M + H]^+$. HRMS (ESIMS): calcd for $C_{37}H_{44}N_7O_{10}$ $[M + H]^+$ 746.3144, found 746.3168.

Assay for Enzyme Activities. The assay of *Ld* AdoHcyase activity in the hydrolytic direction was performed spectroscopically by measuring the rate of the product (Hcy) formed by reaction with DTNB.[17] Briefly the reaction volume was 1 mL containing 4.7 μg of adenosylhomocysteinase and 4 units of Ado deaminase, 100 μM NAD^+, 1 mM EDTA, adenosylhomocystein (500 mM), and 250 mM DTNB in 50 mM potassium phosphate buffer, pH 7.2, in a final volume of 1 mL. The reaction mixture was maintained at 37 °C for 2 min, and it was monitored at 412 nm continuously using an UV spectrophotometer (Spectra max plus, Molecular device, USA). The initial rate was obtained by fitting the data to zero-order.[13,44]

Enzyme Inhibition Assay. All the test compounds were dissolved in DMSO at 0.4 ng/mL concentration. Dilution with triple distilled water (TDW) was carried out to make their final concentrations of 50, 100, and 200 μM for each test compound. The inhibition assay was carried out by preincubating 4.7 μg of *Ld* AdoHcyase with 50, 100, and 200 μM concentrations of each compounds (inhibitors) for 1 h at 37 °C. The reaction started with the addition of the enzyme reaction buffer into the *Ld* AdoHcyase−inhibitor mixture (500 mM S-adenosylhomocysteine, 4 U Adodeaminase, 100 μM NAD^+, and 250 mM DNTB in 50 mM potassium phosphate buffer with 1 mM EDTA, pH 7.2) to make the final volume of 1 mL. NADH was used as the positive control. The positive control wells contain all components of the reaction mixture with the positive control (NADH) in place of the test compound. The blank wells contain all components of the reaction mixture including DMSO but not any test compound to check any detergent effect. All the assays were replicated thrice, and the means of the replicates were calculated. NAD^+ competition assays were

dx.doi.org/10.1021/ci2005862 | *J. Chem. Inf. Model.* 2012, 52, 777−791

also carried out with an initial 1 h preincubation of test compounds and protein. It was found that NAD^+ could attenuate inhibition in a concentration dependent manner, suggesting reversible NAD^+ competitive inhibitors.

■ ASSOCIATED CONTENT

Supporting Information

Validation of homology model for the *Ld* AdoHcyase through Procheck, ProSA-web, and Errat servers, graphical plot analysis of Cα backbone rmsd, potential and solvation energy of protein ligand complex of 10 sampled structures for each ligand. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Mailing address: Medicinal and Process Chemistry Division, C.S.I.R.-Central Drug Research Institute, Lucknow, 226001, India. Tel: +9152226124112/ex 4386. Fax: +9152226123405. E-mail: chakraborty@cdri.res.in (T.K.C.), anuradha_dube@hotmail.com (A.D.), anilsak@gmail.com (A.K.S.).

Author Contributions

#These authors have contributed equally to this work.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors are thankful for financial assistance in the form of fellowships by Indian Council of Medical Research (ICMR) (A.K.G.), Council of Scientific and Industrial Research (C.S.I.R.) (P.K. and K.Y.S.), University Grant Commission (U.G.C.) (P.K.G., A.K.J., and S.D.), and to Mr. A. S. Kushwaha for technical assistance. The authors are also thankful to Sophisticated Analytical Instruments Facility, C.D.R.I., Lucknow, for providing spectroscopic data. The CDRI communication number allotted to this manuscript is 8207.

■ REFERENCES

(1) WHO executive summary for NTD. http://www.who.int/neglected_diseases/en/ (accessed November 15, 2011).

(2) *WHO Expert Committee report on the control of Leishmaniases*; Geneva, March 22−26, 2010.

(3) Pink, R.; Hudson, A.; Mouries, M. A.; Bendig, M. Opportunities and challenges in antiparasitic drug discovery. *Nat. Rev. Drug Discov.* **2005**, *4*, 727−740.

(4) (a) Carter, C. K.; Henriquez, L. F.; Campbell, A. S.; C.W. Roberts, W. C.; Nok, A.; Mullen, B. A.; McFarlane, E. DNA vaccination against the parasite enzyme gamma-glutamylcysteine synthetase confers protection against Leishmania donovani infection. *Vaccine* **2007**, *25*, 4502−4509. (b) Handman, E. Leishmaniasis: current status of vaccine development. *Clin. Microbiol. Rev.* **2001**, *14*, 229−243.

(5) Jean-Moreno, V.; Rojas, R.; Goyeneche, D.; Coombs, G. H.; Walker, J. *Leishmania donovani*: differential activities of classical topoisomerase inhibitors and antileishmanials against parasite and host cells at the level of DNA topoisomerase I and in cytotoxicity assays. *Exp. Parasitol.* **2006**, *112*, 21−30.

(6) Palmer, J. L.; Abeles, R. H. The mechanism of action of S-adenosylhomocysteinase. *J. Biol. Chem.* **1979**, *254*, 1217−1226.

(7) Ueland, P. M. *Pharmacol. Rev.* **1982**, *34*, 223−253.

(8) Yin, D., Yang, X., Yuan, C. S.; and Borchardt, R. T. Applying Chemical Principles to the Understanding and Treatment of Diseases. In *Biomedical Chemistry*; Torrence, P. F., Ed.; Wiley: New York, 2000; pp 41−71.

(9) (a) Yuan, C. S.; Saso, Y.; Lazarides, E.; Borchardt, R. T.; Robins, M. J. Recent advances in S-adenosyl-L-homocysteine hydrolase inhibitors and their potential clinical applications. *Exp. Opin. Therap. Pat.* **1999**, *9*, 1197−1206. (b) Yuan, C. S.; Liu, S.; Wnuk, S. F.; Robins, M. J.; Borchardt, R. T. Design and synthesis of S-adenosylhomocysteine hydrolase inhibitors as broad-spectrum antiviral agents. *Adv. Antiviral Drug Design.* **1996**, *2*, 41−88. (c) Chiang, P. K. biological effects of inhibitors of S-adenosylhomocysteine hydrolase. *Pharmacol. Ther.* **1998**, *77*, 115−134. (d) Daelemans, D.; Este, J. A.; Witvrouw, M.; Pannecouque, C.; Jonckheere, H.; Aquaro, S.; Perno, C. F.; Clercq, E. D.; Vandamme, A. M. S-adenosylhomocysteine hydrolase inhibitors interfere with the replication of human immunodeficiency virus type 1 through inhibition of the ltr transactivation. *Mol. Pharmacol.* **1997**, *52*, 1157−1163.

(10) Henderson, D. M.; Hanson, S.; Allen, T.; Wilson, K.; Coulter-Karis, D. E.; Greenberg, M. L.; Hershfield, M. S.; Ullman, B. Cloning of the gene encoding Leishmania donovaniS-adenosylhomocysteine hydrolase, a potential target for antiparasitic chemotherapy. *Mol. Biochem. Parasitol.* **1992**, *53*, 169−183.

(11) Seley, K. L.; Schneller, S. W.; Rattendi, D.; Bacchi, C. J. (+)-7-Deaza-5′-noraristeromycin as an anti-trypanosomal agent. *J. Med. Chem.* **1997**, *40*, 622−624.

(12) Creedon, K. A.; Rathod, P. K.; Wellems, T. E. Plasmodium falciparum S-adenosylhomocysteine hydrolase. cDNA identification, predicted protein sequence, and expression in Escherichia coli. *J. Biol. Chem.* **1994**, *269*, 16364−16370.

(13) Yuan, C. S.; Yeh, J.; Liu, S.; Borchardt, R. T. Mechanism of inactivation of S-adenosylhomocysteine hydrolase by (Z)-4′,5′-didehydro-5′-deoxy-5′-fluoroadenosine. *J. Biol. Chem.* **1993**, *268*, 17030−17037.

(14) Minotto, L.; Ko, G. A.; Edwards, M. R.; Bagnara, A. S. Trichomonas vaginalis: Expression and Characterisation of RecombinantS-Adenosylhomocystein. *Exp. Parasitol.* **1998**, *90*, 175−180.

(15) Bagnara, A. S.; Tucker, V. E.; Minotto, L.; Howes, E. R.; Ko, G. A.; Edwards, M. R.; Dawes, I. W. Molecular characterisation of adenosylhomocysteinase from Trichomonas vaginalisase. *Mol. Biochem. Parasitol.* **1996**, *81*, 1−11.

(16) Xiaoda, Y. Ronald, T. B. Overexpression, purification, and characterization of S-adenosylhomocysteine hydrolase from Leishmania donovani. *Arch. Biochem. Biophys.* **2000**, *383*, 272−280.

(17) Xing, L.; McDonald, J. J.; Kolodziej, S. A.; Kurumbail, R. G.; Williams, J. M.; Warren, C. J.; O'Neal, J. M.; Skepner, J. E.; Roberds, S. L. Discovery of potent inhibitors of soluble epoxide hydrolase by combinatorial library design and structure-based virtual screening. *J. Med. Chem.* **2011**, *54*, 1211−1222.

(18) *Prime*, version 2.2; Schrödinger, LLC: New York, 2010.

(19) *Schrödinger*, version 9.1; Schrodinger, LLC: New York, 2005.

(20) Laskowski, R. A.; MacArthur, M. W.; Moss, D.; Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **1993**, *26*, 283−291.

(21) (a) Wiederstein; Sippl, M. J. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res.* **2007**, *35*, 407−410. (b) Sippl, M. J. Recognition of Errors in Three-Dimensional Structures of Proteins. *Proteins* **1993**, *17*, 355−362.

(22) Benkert, P.; Schwede, T.; Tosatto, S. C. E. QMEANclust: Estimation of protein model quality by combining a composite scoring function with structural density information. *BMC Struct. Biol.* **2009**, *9*, 35.

(23) Colovos., C.; Yeates, T. O. Verification of protein structures: patterns of nonbonded atomic interactions. *Protein Sci.* **1993**, *2*, 1511−1519.

(24) Gupta, A. K.; Chakroborty, S.; Srivastava, K.; Puri, S. K.; Saxena, A. K. Pharmacophore modeling of substituted 1,2,4-trioxanes for quantitative prediction of their antimalarial activity. *J. Chem. Inf. Model.* **2010**, *50*, 1510−1520.

(25) *CombiGlide*, version 2.5; Schrödinger, LLC: New York, 2009.

(26) *MacroModel*, version 9.8; Schrödinger, LLC: New York, 2010.

(27) *Glide*, version 5.6; Schrödinger, LLC: New York, 2010.

(28) *Discovery Studio*, version 2.1; Accelrys Software Inc.: San Diego, CA, 2008.
(29) *GOLD*, version 3.1; Cambridge Crystallographic Data Centre: Cambridge, UK, 2005.
(30) Chakraborty, T. K.; Gajula, P. K. Studies directed toward the development of amide-linked RNA mimics: synthesis of the monomeric building blocks. *J. Org. Chem.* **2008**, *73*, 6916−6919.
(31) Stevens, J. D. *Methods in carbohydrate chemistry*; Whistler, R. L., BeMiller, J. N., Eds.; Academic Press: Orlando, FL, 1972; Vol. *6*, pp 1230−1282.
(32) Yoshimura, Y.; Endo, M.; Miura, S.; Sakata, S. An alternative synthesis of the antineoplastic nucleoside 4′-thiofac and its application to the synthesis of 4′-thiofag and 4′-thiocytarazid. *J. Org. Chem.* **1999**, *64*, 7912−7920.
(33) Veeresa, G.; Datta, A. Stereoselective synthesis of antifungal antibiotic (+)-preussin. *Tetrahdron* **1998**, *54*, 15673−15678.
(34) Rozners, E.; Xu, Q. Total synthesis of 3′,5′-c-branched nucleosides. *Org. Lett.* **2003**, *5*, 3999−4001.
(35) Rozners, E.; Liu, L. Monomers for preparation of amide-linked rna: asymmetric synthesis of all four nucleoside 5′-azido 3′-carboxylic acids. *J. Org. Chem.* **2005**, *70*, 9841−9848.
(36) (a) Xu, Q.; Katkevica, D.; Rozners, E. Toward Amide-Modified RNA: Synthesis of 3′-Aminomethyl-5′-carboxy-3′,5′-dideoxy Nucleosides. *J. Org. Chem.* **2006**, *71*, 5906−5913. (b) Robins, M. J.; Doboszewski, B.; Timoshchuk, V. A.; Peterson, M. A. Glucose-Derived 3′-(Carboxymethyl)-3′-deoxyribonucleosides and 2′,3′-Lactones as Synthetic Precursors for Amide-Linked Oligonucleotide Analogues. *J. Org. Chem.* **2000**, *65*, 2939−2945.
(37) Zou, R.; Robins, M. J. High-yield regioselective synthesis of 9-glycosyl guanine nucleosides and analogues via coupling with 2-*N*-acetyl-6-*O*-diphenylcarbamoylguanine. *Can. J. Chem.* **1987**, *65*, 1436−1437.
(38) Tantry, S. J.; Venkataramanarao, R.; Chennakrishnareddy, G.; Sureshbabu, V. V. Total Synthesis of Cyclosporin O by Convergent Approach Employing Fmoc-Amino Acid Chlorides Mediated by Zinc Dust. *J. Org. Chem.* **2007**, *72*, 9360−9363.
(39) Dube, A. P.; Sharma, J. K.; Srivastava, A.; Misra, S. N.; Katiyar, J. C. Vaccination of langur monkeys (*Presbytis entellus*) against *Leishmania donovani* with autoclaved *L. major* plus BCG. *Parasitology* **1998**, *116*, 219−221.
(40) Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **1970**, *227*, 680−685.
(41) Towbin, H.; Staehelin, T.; Gordon, J. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl. Acad. Sci.* **1979**, *76*, 4350−4354.
(42) *MOE (Molecular Operating Environment), version 2009. 10;* ; Chemical Computing Group: Montreal, Quebec, Canada, 2009.
(43) *The PyMOL Molecular Graphics System*, version 1.3r1; Schrodinger, LLC: Portland, OR, 2010.
(44) Yuan, C. S.; Ault-Riche, D. B.; Borchardt, R. T. Chemical modification and site-directed mutagenesis of cysteine residues in human placental s-adenosylhomocysteine hydrolase. *J. Biol. Chem.* **1996**, *271*, 28009−28016.