Published in final edited form as:

Chembiochem. 2009 October 12; 10(15): 2455–2459. doi:10.1002/cbic.200900369.

Mechanistic Insights into Water Activation in SAM Hydroxide Adenosyltransferase (duf-62)

Dr. Hai Deng^{[a],[b]}, Dr. Stephen A. McMahon^[a], Dr. Alessandra S. Eustáquio^[c], Dr. Bradley S. Moore^{[c],[d]}[Prof.], Dr. James H. Naismith^[a][Prof.], and Dr. David O'Hagan^[a][Prof.]

David O'Hagan: do1@st-and.ac.uk

^[a] Centre for Biomolecular Sciences, School of Chemistry University of St. Andrews, St. Andrews, KY16 9ST (UK) Fax: (+44) 133-446-3808

^[c] Scripps Institution of Oceanography University of California in San Diego La Jolla, CA 92093-0204 (USA)

^[d] Skaggs School of Pharmacy and Pharmaceutical Sciences University of California at San Diego, La Jolla, CA 92093-0204 (USA) Fax: (+1)858-822-6652

Keywords

biosynthesis; catalysis; enzyme catalysis; enzymes

Enzymes that mediate S_N2 substitution reactions at carbon are relatively rare.[1] S-Adenosyl-L-methionine (SAM) is, however, amenable to substitution, due to its sulfonium moiety and the leaving group propensity of the resultant thioether.[2] Different enzymes have been identified that catalyse nucleophilic attack at all three carbons (C5', Me, C_γ) of SAM positioned adjacent to this sulfonium sulfur (Scheme 1). The methyltransferase enzyme class attacks the sulfonium Me utilising C, N, O, S nucleophiles to demethylate SAM and generate S-adenosyl-L-homocysteine (SAH).[3] Spermidine[4] and spermine[5] synthases catalyse the attack of aminopropyl nucleophiles at the C_γ methionyl arm of decarboxylated S-adenosyl-L-methionine (dcAdoM)—an important reaction during polyamine biosynthesis.[6] Two enzymes, the fluorinase from *Streptomyces cattleya*[7,8] and the chlorinase from *Salinispora tropica*[9] catalyse halide ion S_N2 substitution at the C5' of SAM to generate 5'-fluoro-5'-deoxyadenosine and 5'-chloro-5'-deoxyadenosine, respectively (Scheme 1).

Both the fluorinase and chlorinase enzymes are involved in the first step of halogen metabolite biosynthesis in their respective organisms. These enzymes share 30% homology and belong to a large microbial superfamily, which is annotated as domains of unknown function-62 (duf-62). We have recently demonstrated that the nonhalogenating members of this superfamily unusually catalyse nucleophilic attack of water (hydroxide) at the C5' carbon of SAM to generate adenosine and L-methionine.[10,11] Structures of four nonhalogenating duf-62 proteins in complex with adenosine have been deposited.[2,10] As expected they have the same fold as the fluorinase and chlorinase structures; this confirms their evolutionary relationship. However, the duf-62s have highly conserved active sites including a His–Arg–Asp triad of amino acids that are not found in the two halogenating (fluorinase and chlorinase) members of the superfamily, a striking difference between these

Correspondence to: Alessandra S. Eustáquio; David O'Hagan, do1@st-and.ac.uk.

^[b]Current address: Marine Biodiscovery Centre Department of Chemistry, University of Aberdeen Aberdeen, AB24 3UE (UK)

proteins. This triad is highly conserved in ~300 duf-62 amino acid sequences and its role in water activation for nucleophilic attack is intriguing and not obvious. The active-site region of the *Pyrococcus horikoshii* structure with adenosine (product) bound is shown in Figure 1A and the catalytic triad involving His127, Arg75 and Asp68 is highlighted, as is a bound water (red) molecule, which is hydrogen bonded between Asp68 and the adenosine hydroxyl group.

Assay of the *P. horikoshii* SAM hydroxide adenosyltransferase in an H₂ ¹⁸O enriched buffer resulted in [5'-¹⁸O]-adenosine, but no oxygen-18 isotope was incorporated into the Asp68 carboxylate of the enzyme; this indicates direct attack of SAM by water rather than an initial attack by the carboxylate group of Asp68 followed by hydrolysis. On the basis of the structural (adenosine complex) and isotope labelling evidence we proposed the mechanism illustrated in Figure 1 B.[10] The structural water molecule completes a cyclic network by hydrogen bonding to the Asp68 carboxylate and the amide NH of His127. This working hypothesis implies that either upon or before SAM binding, the enzyme must recruit and activate a second water molecule that would act as the nucleophile. The water is predicted to be close to O5 of the adenosine in the complex structure. Two general processes could account for water activation. The first involves an electronic proton relay in which hydrogen bonding between His127 and Arg75 weakens the Arg75 to Asp68 interaction, and this in turn allows the Asp68 carboxylate to act as a base to assist deprotonation of a proximate water molecule, which either acts directly as a nucleophile or in turn it deprotonates a second water molecule, which then acts as a nucleophile. Alternatively, the hydrogen bonded triad is conserved through this family of proteins as a scaffold to help order/position the nucleophilic water. In the second hypothesis the binding energy and localisation of SAM pushes the C5' carbon into a proximate reaction distance with the water molecule and deprotonation would occur after nucleophilic attack.

In order to obtain additional insight into the nature of this catalysis we have carried out site-directed mutagenesis studies on the duf-62 enzyme from *S. arenicola*, and we have obtained a co-crystal of the *P. horikoshii* enzyme with the unreactive substrate analogue SAH.

SAH is often a potent inhibitor of SAM-dependent enzyme reactions,[12] and it proved to be an efficient inhibitor of the *P. horikoshii* SAM hydroxide adenosyltransferase ($K_i = 3.6 \mu\text{M}$). [10] Therefore, we determined the structure of the SAH complex as a model for the ternary complex (Figure 2). The enzyme is a trimer and the SAH ligand is bound at the three interfaces between subunits in a similar manner to that found in the fluorinase and chlorinase structures.[7,8] The protein structure is at the gross level unchanged from the adenosine co-crystal structure[10,11] A water molecule (green) is hydrogen bonded to the Asp68 residue, in a site previously observed for the adenosine structure (yellow water molecule). The new structure does locate the binding site of the methionyl moiety of SAM, which forms hydrogen bonds to the highly conserved side chain of Thr125 and Ser226 from the same monomer.

As hypothesised, the SAH complex reveals there is now a second outer water molecule (blue) hydrogen bonded to the inner water molecule (green) at the active site and located proximate to the electrophilic 5'-carbon atom of SAH. This outer (blue) water immediately emerges as a candidate for the putative nucleophilic water molecule. Since the enzyme does not catalyze SAH hydrolysis, it is insufficiently activated for nucleophilic attack of SAH and the nucleophilic water is reasonably secured within the structure. Given the very close correspondence between the adenosine structure and the SAH structure, and the conservation of the methionyl binding site we do feel that the SAH complex is a robust model for the ternary complex. Both the outer (blue) and inner (green) water molecules are potential candidates for nucleophilic attack at C5', as both are within 3.5 Å of C5', and the

true nucleophile cannot be unambiguously identified in this study. The inner (green) water actually has the more favourable trajectory for attack (140° vs. 98°) in this structure, although this water is found in both product and substrate structures and is judged more likely to be a structural water molecule. The exact trajectory of the outer (blue) water can expect to vary a little in going from this unproductive SAH structure to a catalytically competent SAM complex. In the adenosine complex, the Thr125 amide carbonyl, hydrogen bonds to the adenosine hydroxyl. In the SAH complex this carbonyl hydrogen bonds to the outer (blue) water. In doing so the main chain rotates slightly; this suggests that the hydrogen bond to Thr125 is preserved during the reaction course, as that water molecule becomes the C5'—O of adenosine.

In order to shed more light on the role of the His–Arg–Asp amino acid triad, mutagenesis studies were carried out, this time on the *Salinispora arenicola* enzyme. The *S. arenicola* duf-62 gene was identified after a full genome sequence analysis of this oceanic actinomycete bacterium. It has a 34% identity to *P. horikoshii* enzyme. The mutagenesis studies focused on the amino acid residues that constitute the His–Arg–Asp triad. These residues were sequentially replaced by alanine. Replacement of Asp72 (Asp68 in *P. horikoshii* enzyme numbering) resulted in complete loss of catalytic activity and similarly mutation of Arg79 (Arg75 in *P. horikoshii* enzyme numbering) to alanine gave an almost inactive protein. This indicates that these residues are very important in catalysis. However, the histidine (His125 in *P. horikoshii* enzyme numbering) to alanine (His130-Ala) mutation generated a protein that retained a significant level of catalytic efficiency (25 %). A more radical replacement of His130 with tyrosine (as found in fluorinase and chlorinase) reduced the catalytic efficiency further to 5% of wild-type levels.

Blast sequence analysis of ~300 duf-62 proteins available in GenBank shows that the rate of substitution of the Asp–Arg–His triad is 3% for Asp (Tyr in fluorinase and chlorinase, and Glu or Asn in four proteins each), 10% for Arg (including Thr in fluorinase and chlorinase, Asn, Gln and Ser) and 13% for His130 (including Tyr in fluorinase and chlorinase, Pro, Arg, Gln and Thr). The low substitution rate found for Asp72 (*S. arenicola* numbering) is consistent with the essential role of this residue in catalysis. Furthermore, the fact that the Asp–Arg–His triad is highly but not absolutely conserved throughout the family might suggest some plasticity of the triad while still maintaining activity. This seems particularly true for the His residue as shown by the mutagenesis studies. Alternatively, the possibility that some family members carry out slightly different biochemistry, for example, by using diverse nucleophiles as in the case of halogenases, should be considered.

The two active site water molecules in the unproductive SAH structure emerge as candidate nucleophiles. Both lie close to the C5' of SAH, which is expected to approximate the location of the electrophilic carbon of SAM. The outer (blue) water is perhaps the better candidate for the nucleophile. It hydrogen bonds to Thr125 and the movement of Thr125 residue in going from the substrate model (SAH) complex to the product (adenosine) complex suggests that the backbone of the protein moves with the nucleophilic water molecule as it carries out the chemistry. Considering the structure, mutation of the Asp residue will have two obvious effects. Firstly, it will perturb the *positioning* of the inner water (green) molecule, as the Asp carboxylate is no longer available as a hydrogen bonding acceptor. Secondly, if the Asp acts as a base initiating a proton-transfer process to induce nucleophilicity, then that will also be compromised. Mutation of the Arg79 residue in the *S. arenicola* enzyme resulted also in almost complete loss of catalytic activity. This mutation would significantly perturb both the hydrogen bonding and the positioning of the Asp residue, and subsequent ordering/positioning of the two water molecules. The observation that His130 of the *S. arenicola* enzyme can be mutated to Ala with substantial retention of catalytic activity (~25%) is surprising as this is a conserved residue in 260 of 300 duf-62

proteins. In the structure the His hydrogen bonds to the Arg of the triad, and as the pK_a of His and Arg differ significantly, proton transfer from the arginine to the histidine should not be significant, however, a key structural role for His is implied. Nonetheless, our data suggest that water activation/positioning depends principally on the Asp–Arg ion pair.

In summary, SAH was co-crystallised with SAM hydroxide adenosyltransferase from *P. horikoshii*. A second water molecule was located in the active site and is implicated as the nucleophile in the substitution reaction. The amino acid triad and particularly the Asp–Arg ion pair appears to play an important role in holding the water for nucleophilic attack as SAM binds and becomes proximate. Further studies exploring the role of each water molecule and the relative importance of charge transfer versus a structural/positioning of water by hydrogen bonding, the role of the His residue of the triad, and the Thr125 carbonyl (*P. horikoshii* numbering) are being explored by QM/MM theory calculations.

Experimental Section

Chemicals

SAM-*p*-toluenesulfonate was purchased from Sigma and adenosine from Acros Organics.

Purification of SAM hydroxide adenosyltransferase from *P. horikoshii*

The plasmid pET-PH0463 was constructed in the previous report.[10] The resultant plasmid was transferred into competent *E. coli* Rosetta cells (Novagen) and grown in Luria broth containing kanamycin (100 mg mL^{-1}) at 37°C until an absorbance of 0.6 at 600 nm was reached. Over-expression was induced by the addition of isopropylthiogalactoside (IPTG; final concentration 1 mM) and incubation was continued at 16°C for 24 h. Cells were collected and lysed by sonication, and the apoenzyme was purified to homogeneity as described previously. The yield was about 30 mg L^{-1} media.

Purification of SAM hydroxide adenosyltransferase from *S. arenicola*

Plasmid pJH1 containing the parent gene in pHIS8 has been previously described.[11] Point mutants were generated by using the “QuickChange® Site-Directed Mutagenesis Kit” (Stratagene) and pJH1 template DNA (100 ng), with DMSO (5%) in the PCR reaction. One® Shot TOP10 chemically competent cells (Invitrogen) were used instead of XL1-Blue. Complementary primer pairs were designed for each mutation (mutation site underlined) as follows (only the sense strand sequence is shown). D72A (GTC GCC GTG GTC GCC CCC GGC GTG GGC), R79A (GTG GGC ACG GCC GCC CGG GCG ATC G), H130A (GGA TGT CCG CGA CCT TTG CCG GCC GGG ACG TC) and H130Y (GGA TGT CCG CGA CCT TTT ACG GCC GGG ACG TC). Expression in *E. coli* BL21 (DE3) and purification of the His₈-tagged mutant proteins were carried out as previously described except that growth temperature was 28°C throughout and imidazole (40 mM) was added to lysis and wash buffers.[9] The yield of soluble protein was approximately 1 mg per litre , and it was stored in phosphate buffer (50 mM , pH 7.9) and 50% (v/v) glycerol at -20°C as described for the parent protein.[13]

Activity assay

Kinetic constants of parent and mutant proteins were determined by incubating the respective enzyme (100 nM) with SAM (1 to $200 \text{ }\mu\text{M}$) in phosphate buffer (50 mM , pH 7.9) at 37°C for 3 (wild type), 5 (His130Ala) or 10 (His130Tyr) minutes as previously described. [11] Relative activity of the Arg72Ala mutant in comparison to the wild type was based on K_{cat} only, which was investigated by incubating the enzyme ($1 \text{ }\mu\text{M}$) with SAM (0.2 mM) for 10 min.

Crystallography

Crystal growth conditions were identified by using the sitting-drop vapour diffusion technique and a selection of in-house stochastic crystallisation screens. Protein was concentrated to 17 mg mL⁻¹ and SAH was added to a 5_M excess concentration. The protein/SAH mixture was screened at 293 K with total drop sizes of 0.3 µL (containing 0.15 µL of protein and 0.15 µL precipitant) and 0.45 µL (containing 0.3 µL of protein and 0.15 µL precipitant) prepared by using a Cartesian Honeybee 963 nanodrop crystallisation robot. Thin rod-shaped crystals (30–50×30–50×100–200 µm) were observed with at least five different precipitating agents. Each of these crystals was cryoprotected by addition of PEG 400 (25%, v/v) to the crystallising mother liquor and tested in-house for protein diffraction. Three of these crystals diffracted in-house to higher than 3 Å resolution and were therefore stored at 100 K for transport to the synchrotron. No optimisation of these crystals was performed.

A 1.77 Å dataset was collected on ID14-4 at the European Synchrotron Radiation Facility, Grenoble France at 100 K from a single crystal previously tested in-house. This crystal was grown from PEG 6000 (19.83%), KCI (0.26_M) and MOPS (0.1_M, pH 7) and cryoprotected by addition of PEG 400 (25%, v/v). Data were collected as 0.5 s 0.5° oscillations by using 90% beam attenuation on a Q315r ADSC CCD detector with a wavelength of 0.979 Å. The crystal to detector distance was 237 mm. All data were integrated and reduced in space group R3 by using HKL2000,[14] with unit-cell parameters of $a = b = 65.9$, $c = 142.8$ Å, $\alpha = \beta = 90^\circ$, $\gamma = 120^\circ$; full details are given in Table 1. The structure was solved by molecular replacement by using PHASER[15,19] as implemented in CCP4 (CCP4, 1994) with the monomer from project PH0463 from *Pyrococcus horikoshii* OT3 (PDB ID: 1WU8) as the search model. One protein molecule was found in the asymmetric unit consistent with a Matthews coefficient of 2.11 and solvent content 41.2%. TLS parameters were defined and employed throughout refinement with REFMAC5.[17,18] Manual rebuilding was with COOT.[19] The model was refined to $R = 22$ and $R_{\text{free}} = 25.7\%$ were upon clear density could be observed for SAH. SAH was modelled in this density and refinement was duly completed. The structure was examined throughout for correctness with MOLPROBITY.[20] Both the structure and structure factors have been deposited with the PDB (Table 2).

Acknowledgments

We thank the BBSRC for grants (BB/F007426/1 and BBS/B/14426), the NIH (CA127622 to B.S.M.) and the Scottish Funding Council (SULA and SSPF). A.S.E. is a Tularik postdoctoral fellow of the Life Sciences Research Foundation. Kenneth Johnson is acknowledged for his assistance in data collection.

References

1. Dewar MJS. Enzyme. 1986; 36:8–20. [PubMed: 3792297]
2. Deng H, O'Hagan D. Curr Opin Chem Biol. 2008; 12:582–592. [PubMed: 18675376]
3. Komoto J, Yamada T, Takata Y, Konishi K, Ogawa H, Gomi T, Fujioka M, Takusagawa F. Biochemistry. 2004; 43:14385–14394. [PubMed: 15533043]
4. Wu H, Min J, Ikeguchi Y, Zeng H, Dong A, Loppnau P, Pegg AE, Plotnikov AN. Biochemistry. 2007; 46:8331–8339. [PubMed: 17585781]
5. Wu H, Min J, Zeng H, McCloskey DE, Ikeguchi Y, Loppnau P, Michael AJ, Pegg AE, Plotnikov AM. J Biol Chem. 2008; 283:16135–16146. [PubMed: 18367445]
6. Golding BT, Nassereddin IK. J Am Chem Soc. 1982; 104:5815–5817.
7. O'Hagan D, Schaffrath C, Cobb SL, Hamilton JTG, Murphy CD. Nature. 2002; 416:279. [PubMed: 11907567]
8. Dong C, Huang FL, Deng H, Schaffrath C, Spencer JB, O'Hagan D, Naismith JH. Nature. 2004; 427:561–565. [PubMed: 14765200]

9. Eustáquio A, Pojer F, Noel JP, Moore BS. *Nat Chem Biol.* 2008; 4:69–74. [PubMed: 18059261]
10. Deng H, Botting CH, Hamilton JTG, Russell RJM, O'Hagan D. *Angew Chem.* 2008; 120:5437–5441.
11. Eustáquio AS, Härle J, Noel JP, Moore BS. *ChemBioChem.* 2008; 9:2215–2219. [PubMed: 18720493]
12. Ghosh AK, Wang Y. *J Chem Soc Perkin Trans 1.* 1999; 24:3597–3601.
13. Jez JM, Ferrer JL, Bowman ME, Dixon RA, Noel JP. *Biochemistry.* 2000; 39:890–902. [PubMed: 10653632]
14. Otwinowski Z, Minor W. *Methods Enzymol.* 1997; 276:307–326.
15. McCoy AJ, Grosse-Kunstleve RW, Storoni LC, Read RJ. *Acta Crystallogr Sect D.* 2005; 61:458–464. [PubMed: 15805601]
16. Storoni LC, McCoy AJ, Read RJ. *Acta Crystallogr Sect D.* 2004; 60:432–438. [PubMed: 14993666]
17. Murshudov GN, Vagin AA, Dodson EJ. *Acta Crystallogr Sect D.* 1997; 53:240–255. [PubMed: 15299926]
18. Murshudov GN, Vagin AA, Lebedev A, Wilson KS, Dodson EJ. *Acta Crystallogr Sect D.* 1999; 55:247–255. [PubMed: 10089417]
19. Emsley P, Cowtan K. *Acta Crystallogr Sect D.* 2004; 60:2126–2132. [PubMed: 15572765]
20. Davis W, Leaver-Fay A, Chen VB, Block JN, Kapral GJ, Wang X, Murray LW, Arenall WB, Snoeyink J, Richardson JS, Richardson DC. *Nucleic Acids Res.* 2007; 35(Suppl 2):W375–W383. [PubMed: 17452350]

Figure 1.

A) The active-site region of SAM hydroxide adenosyltransferase from *P. horikoshii* with the product adenosine (yellow) bound. The His127–Arg75–Asp68 hydrogen bonded triad is highlighted, as is an ordered water molecule (red). B) A putative mechanism is illustrated for nucleophilic substitution based on isotope labelling and structural evidence.[10,11]

Figure 2. SAM hydroxide adenosyltransferase from *P. horikoshii* is represented with two views of the X-ray structure of the active site region of Asp-Arg-His triad with SAH (substrate analogue) bound. The structure reveals the location of the methionyl residue and two ordered water molecules (outer: blue; inner: green). The structure is overlaid with that of the previous adenosine (product) structure,[10] which had a single water molecule (yellow) hydrogen bonded to Asp68. It is noteworthy that the amide carbonyl of Thr125 moves/rotates between the substrate and product structures; this indicates a hydrogen bond that might track nucleophile movement.

Scheme 1.

Nucleophilic reactions of SAM-dependent enzymes including the recently identified SAM hydroxide transferase (duf-62).

Table 1

Mutagenesis studies on the duf-62 enzyme from *Salinispora arenicola*. Apparent kinetic constants of wild-type and mutant proteins.

<i>S. arenicola</i> duf-62	K_M [μM]	k_{cat} [min^{-1}]	K_{cat}/K_M [$\mu\text{M}^{-1} \text{min}^{-1}$]	Relative catalytic efficiency k_{cat}/K_M [%]
wild-type	1 ± 0.15	0.45 ± 0.07	0.45	100
Asp72Ala	–	–	–	–
Arg79Ala	n.d.	n.d.	n.d.	0.15*
His130Ala	1.4 ± 0.15	0.16 ± 0.01	0.114	25
His130Tyr	1.7 ± 0.3	0.04 ± 0.008	0.023	5

n.d.: not determined; the low activity of the Asp72Ala mutant makes accurate determination of kinetic constants difficult;

* based on relative turnover (K_{cat}) only by using 200 μM SAM and 1 μM enzyme.

Table 2

Crystallography data and data collection statistics. Values in parentheses refer to the highest resolution shell.

ESRF ID14-4	
wavelength [Å]	0.979
resolution [Å]	50.0–1.77(1.83–1.77)
space group	R3
temperature [K]	100
detector	Q315r ADSC CCD
detector distance	237
unit-cell parameters [Å]	$a = b = 65.9$, $c = 142.8$ Å, $\alpha = \beta = 90^\circ$, $\gamma = 120^\circ$
solvent [%]	41.2
unique reflections	22346
$I/\sigma(I)$	11.4 (3.6)
average redundancy	4.6 (3.7)
data completeness [%]	99.7 (100)
R_{merge} [%] ^[a]	4.5 (50.1)
<i>Refinement</i>	
R_{factor} [%] <i>b</i>	16.8
R_{free} [%] <i>C</i>	21.7
rmsd bond distance [Å]	0.017
rmsd bond angle [°]	1.622
average B factors [Å ²]	
main chain	40.67
side chain	42.7
ligand	32.7
solvent	45.35
<i>In Ramachandran</i> [%]	
core	96.1
disallowed	0
no. of protein atoms	2038
no. of ligand atoms	26
no. of solvent atoms	117
PDB ID	2wr8

[a] $R_{\text{merge}} = \frac{\sum \sum I(h)i - (I(h)) \sum I(h)i}{\sum I(h)i}$, where $I(h)$ is the measured diffraction intensity and the summation includes all observations.