From Natural Methylation to Versatile Alkylations Using Halide Methyltransferases

Qingyun Tang,[a] Ioannis V. Pavlidis,[b] Christoffel P. S. Badenhorst,*[a] and Uwe T. Bornscheuer*[a]

Halide methyltransferases (HMTs) enable the enzymatic synthesis of S-adenosyl-L-methionine (SAM) from S-adenosyl-L-homocysteine (SAH) and methyl iodide. Characterisation of a range of naturally occurring HMTs and subsequent protein engineering led to HMT variants capable of synthesising ethyl, propyl, and allyl analogues of SAM. Notably, HMTs do not depend on chemical synthesis of methionine analogues, as required by methionine adenosyltransferases (MATs). However, at the moment MATs have a much broader substrate scope than the HMTs. Herein we provide an overview of the discovery and engineering of promiscuous HMTs and how these strategies will pave the way towards a toolbox of HMT variants for versatile chemo- and regioselective biocatalytic alkylations.

Background

The “magic methyl” effect refers to the ability of well-placed methyl groups to dramatically alter the physical and biological properties of a molecule.[1] This makes methylation one of the most atom-efficient strategies for modulation of biological effects.[2] Going beyond methylation, structural diversification of natural products by alkyl-diversification would create libraries of compounds with profoundly enhanced properties. However, catalyst-controlled chemo- and regioselective alkylation of molecules bearing two or more heteroatoms with similar reactivities is still a major challenge in organic synthesis. This is particularly true for N-heterocyclic compounds like pyrazoles, triazoles, and pyridines, since tautomerisation can lead to different heteroatoms (e.g. nitrogen and oxygen) having similar reactivities (Scheme 1a).[3] Since drug molecules often contain amines, N-heterocycles, and unprotected polar groups, selective heteroatom alkylation is a reaction on many organic chemists’ wish list.[4]

Enzymes can address this challenge, due to their chemo-, regio-, and stereoselectivities.[5] For example, the S-adenosyl-L-methionine-dependent methyltransferases (MTs) catalyse the highly chemo- and regioselective methylation of a broad range of biomolecules at C, O, N, S, and P atoms.[6] Very recently, Hammer’s group demonstrated the incredible potential of protein engineering by designing a panel of methyltransferases for the regioselective methylation of pyrazoles, with regioisomeric ratios up to 99% (Scheme 1b).[7] Before, there were no chemical, or even biological, catalysts known capable of controlling pyrazole alkylation so precisely.[8]

Interestingly, like many naturally-occurring MTs, Hammer’s biocatalysts are promiscuous, in terms of the alkyl donor, and can transfer a range of alkyl groups, as long as the corresponding S-adenosyl-L-methionine (SAM) analogues are available.[6c,7] Unfortunately, SAM analogues are rare in nature and hard to synthesise. A prominent exception is carboxy-SAM (cxSAM, Scheme 1c), which was identified as a co-purified ligand in the crystal structure of CmoA, the enzyme that synthesises cxSAM from SAM and prephenate.[8] CxSAM is used by the carboxymethyl transferase CmoB for tRNA alkylation.[9] A carboxy-S-adenosyl-L-ethionine (cxSAE) analogue (Scheme 1c) has been synthesised from ethionine, using a human methionine adenosyltransferase (MAT) and CmoA.[8] The MAT uses ethionine, a methionine analogue, to produce SAE, the corresponding SAM analogue. MATs use ATP as source of the adenosyl moiety and are remarkably promiscuous with regards to methionine analogues, allowing a broad range of SAM analogues to be synthesised (Scheme 2a).[10] Halogenases have also been used to synthesise SAM analogues from methionine analogues. These enzymes use 5′-chloro-5′-deoxyadenosine or 5′-fluoro-5′-deoxyadenosine as donors of the adenosyl moiety (Scheme 2a).[10] MATs and halogenases are often coupled with MTs in one-pot cascade reactions to form alkylated products. Details about these enzymes and their applications in biocatalytic alkylation cascades have been described in recent reviews.[10a,11a,12]

Methionine analogues need to be chemically synthesised as they are not readily available (Scheme 2a).[10a,11c,13] The harsh chemical synthesis conditions are not compatible with biotransformations and therefore have to be performed separately, limiting their application. SAH can be directly alkylated using alkyl triflates or alkyl bromides, but this results in a mixture of (S,S)- and (R,S)-epimers, the latter being potent MT inhibitors (Scheme 2b).[14] Furthermore, residual SAH is also a potent MT inhibitor.[11] Consequently, chemically synthesised...
SAM analogues have to be purified by – non-trivial – chromatography, limiting the application of this approach.\(^{[7a,16]}\)

### Regeneration of SAM Analogues

SAM has to be used in stoichiometric amounts, producing an equivalent amount of SAH, which is a potent MT inhibitor.\(^{[15]}\) A fundamental problem is that SAM-dependent enzymes transfer a 15 Da methyl group using a 384 Da handle that has to be degraded because it inhibits MT activity.\(^{[17]}\) Poor atom economy has therefore always been a significant drawback of SAM-dependent methylations. A landmark publication by the Andexer group showed that SAM could be regenerated using a five-enzyme cascade, starting with the hydrolysis of SAH to adenosine and homocysteine, followed by three phosphorylations of adenosine to form ATP, which is used by an MAT to convert methionine to SAM. Unfortunately, this cascade fails after 11 cycles, produces homocysteine as waste, and depends on methionine and its analogues as alkyl donors.\(^{[18]}\)

Recently, Liao and Seebeck described a one-enzyme system for SAM recycling (Scheme 3a).\(^{[17]}\) Halide methyltransferases (HMTs) use SAM to methylate halide ions, annually releasing megatons of methyl halides into the atmosphere.\(^{[19]}\) Liao and Seebeck realized that methyl halide synthesis is driven by the evaporation of volatile methyl halides and the metabolic degradation of SAH. Therefore, they could employ the reaction in reverse, which is thermodynamically favoured, enabling the production of SAM from SAH and methyl iodide, with iodide as the sole by-product (Scheme 3a). A stoichiometric amount of methyl iodide is provided as methyl group donor and a catalytic amount of SAH is recycled up to 580 times.\(^{[17]}\) The HMT-synthesized SAM was utilized by different O\(^{\mathrm{N}}\)-, N\(^{\mathrm{N}}\)- and C-MTs to produce several regioselectively methylated products (Scheme 3b).\(^{[17,20]}\)

### Promiscuous HMTs

Recently, the Bornscheuer\(^{[21]}\) and Hammer\(^{[26]}\) groups demonstrated that HMTs can also be used for the synthesis and regeneration of a range of SAM analogues. These HMTs, combined with alkyl-promiscuous MTs, produced the regioselectively alkylated products 4-‘O-ethylflavonol, 4-allyloxy-3-hydroxybenzaldehyde, 1-ethyl-5-cyclopropylpyrazole, and 1-propyl-5-cyclopropylpyrazole (Scheme 3c). In our work, more than 40 regeneration cycles were achieved for the preparative scale alkylation of flavonoid and phenolic compounds, using 100 \(\mu\)M of SAH.\(^{[21]}\) Importantly, these cycle numbers may well be underestimated, as additional SAH added to the reactions may not have been necessary. The traces of SAH and/or SAM bound to purified proteins are sufficient for bioalkylation cascades.\(^{[21]}\)

A key realisation was that, just like methyltransferases in general, HMTs should be alkyl-promiscuous, enabling them to accept donors other than methyl iodide. While the *Chloracido bacterium thermophilum* HMT used by Liao and Seebeck did not have significant ethyltransferase activity, the enzymes from *Arabidopsis thaliana*, *Raphanus sativus*, *Aspergillus clavatus*, *Batis maritima*, and *Synechococcus elongatus* did.\(^{[21b,21]}\) As we predicted, even more promiscuous HMTs can be found in nature, and Bengel et al. reported that the *Aspergillus clavatus* HMT had significant propyltransferase activity and could even produce a small amount of the butyl analogue of SAM (Table 1).\(^{[4b]}\) This enzyme was identified by screening only seven of the HMTs described in the literature, suggesting that even more useful HMTs could be found in nature. Therefore, we suggest further investigation of the 89 HMTs that have already been functionally characterised in the past.\(^{[22]}\) However, it is very unlikely that an HMT would be so promiscuous that it could be used for any alkyl group of interest. Therefore, HMT engineering will probably play an important role in the future of bioalkylation.

---

**Scheme 1.** a) Due to tautomerisation of heteroatoms, chemical alkylation of \(N\)-heterocyclic compounds like pyrazoles, triazole, and pyridone usually leads to mixtures of alkylated products. b) Regioselective methylation of pyrazoles catalysed by SAM-dependent MTs. By using different MT variants, either 1-methylpyrazoles or 2-methylpyrazoles can be synthesised with regiosomeric ratios up to 99%. c) CxSAM is a rare example of a naturally occurring SAM analogue. CxSAM is synthesised from prephenate and SAM by CmoA. Starting from 5-adenosyl-l-ethionine (SAE), the cxSAE analogue can be synthesised.

**Table 1.** Comparison of methyltransferase activity of the HMTs described in this work with key enzymes in SAH and SAM metabolism.
Scheme 2. a) Chemoenzymatic synthesis of SAM analogues. Methionine analogues can be synthesized by reductive alkylation of homocysteine. SAM analogues can be synthesized by MATs from methionine analogues and ATP, or by halogenases from methionine analogues and 5'-chloro-5'-deoxyadenosine or 5'-fluoro-5'-deoxyadenosine. The box shows alkyl groups of SAM analogues produced by MATs and halogenases. b) Chemical synthesis of SAM analogues. SAH can be directly alkylated using alkyl triflates or alkyl bromides, resulting in a mixture of epimers.

Table 1. Conversion [%] of SAH and alkyl iodides to SAM analogues by promiscuous HMTs.

| Enzyme                        | Methyl iodide | Ethyl iodide | Ethyl bromide | Propyl iodide | Allyl iodide | Butyl iodide |
|-------------------------------|---------------|--------------|---------------|---------------|--------------|--------------|
| Arabidopsis thaliana HMT (WT) | 100           | 44           | 1             | 0             | –            | 0            |
| Arabidopsis thaliana HMT (V140T) | 90            | –            | –             | 50            | 71           | –            |
| Aspergillus clavatus HMT     | 100           | 78           | 99            | 31            | 1            | –            |
| Synechococcus elongatus HMT  | 99            | 81           | 100           | 1             | –            | 3            |

[a] The Raphanus sativus HMT was also tested by Tang et al.[21] but is not mentioned because activity was only determined in U/mg and conversion similar to other data in this table was not determined. [b] Reactions were performed with crude E. coli cell lysate, 1 mM SAH, and 40–80 equivalents of alkyl halides at 25 °C for 20 h, as described by Bengel et al.[4b] [c] Reactions were performed on preparative scale with purified enzyme, 15 mg SAH and 8 equivalents of alkyl iodides at 25 °C for 24 h, as described by Tang et al.[21]
Utilizing a very sensitive iodide assay, we identified a variant (V140T) of the Arabidopsis thaliana HMT that can synthesise the ethyl, propyl, and allyl analogues of SAM (Table 1). We used ethyl iodide as substrate for screening our site-directed mutagenesis libraries, as the wild-type Arabidopsis thaliana HMT already had activity towards ethyl iodide and therefore the chances of finding improved variants were good. We believe that screening larger libraries using bulkier substrates like butyl- or benzyl iodide would lead to the identification of variants capable of efficiently producing the corresponding SAM analogues. The results we obtained by screening with ethyl iodide led to two important conclusions. First, screening using a slightly larger donor (e.g., ethyl vs. methyl iodide) can result in the identification of variants active on even larger substrates (e.g., propyl and allyl iodide). Second, the V140T variant was generally the best catalyst for transferring all the alkyl groups we tested. This suggests that a single substitution can make an HMT promiscuous and that engineering robust general alkyl transfer catalysts should be possible.

HMT Engineering

We originally tested the HMTs from Chloracidobacterium thermophilum (the enzyme used by Liao and Seebeck), Raphanus sativus (highest activity towards methyl iodide), and Arabidopsis thaliana. We found that the Arabidopsis thaliana HMT was engineered because it had the highest activity towards ethyl iodide between the three candidates and because its 3D structure was available, which facilitated semi-rational protein engineering. Currently, this is still the only solved HMT structure. Further HMT engineering would benefit from detailed structural and mechanistic understanding, suggesting that solving more HMT structures will be important.

While molecular docking suggested that the Arabidopsis thaliana HMT can accommodate even butyl iodide in its spacious active site, substrates larger than ethyl iodide are not converted (Figure 1). Currently, it is believed that SAM-dependent MTs catalyse alkyl transfer by binding the substrates (SAM analogue and acceptor) in reactive conformations, pre-aligning electron orbitals for efficient $S_{N}2$ displacement. This is likely also the case for HMTs, but we are not aware of detailed mechanistic studies. The lower activity with larger alkyl halides
Figure 1. Crystal structure of Arabidopsis thaliana HMT with the ligand SAH and docked methyl, ethyl, propyl, and butyl iodide in the active site. V140, SAH, and alkyl iodides are shown as sticks. Other sites mutated by Tang et al. are shown as lines.

probably results from less efficient binding of the larger substrates in the same reactive conformation as methyl iodide.

A better understanding of the natural functions of HMTs would also facilitate application and engineering of these enzymes. Suggestions that the function of HMTs is to regulate intracellular halide concentrations seem highly unlikely; some studies have suggested that the formation of methyl halides is a "metabolic accident." Generally, HMT activities depend on substrate in the order SCN⁻ > HS⁻ > I⁻ > Br⁻ > Cl⁻, which corresponds to the nucleophilicity of the anions.

A few studies have suggested that the intended substrates of HMTs are toxic thiols resulting from the hydrolysis of, for example, glucosinolates. The catalytic efficiencies of Arabidopsis thaliana HMT against SCN⁻ and (NH₄)₂S are even five orders of magnitude higher than against Cl⁻, which indicates that methylation of halides might be an unspecific side-reaction. This is a significant consideration because it suggests that halide-methyl transfer is a broadly distributed promiscuous activity of methyltransferases and testing a much larger number of MTs for promiscuous HMT activity is advisable. HMT and thiol methyltransferase (TMT) protein sequences share larger number of MTs for promiscuous HMT activity is advisable. promiscuous activity of methyltransferases and testing a much longer alkyl chains. HMT activity on less electrophilic alkyl donors like alkyl bromides and chlorides would also enable more selective alkylation. Interestingly, ethyl bromide has been shown to be well accepted as alkyl donor by the Aspergillus clavatus, Synechococcus elongatus, and Batis maritima HMTs (Table 1). It would be very interesting to see whether HMTs can be engineered for higher activity towards these substrates, and our previously described ultrasensitive bromide assay would facilitate such engineering.

Another exciting avenue for further research is in situ generation of alkyl halides, decreasing the dependence on alkyl iodides. An emerging possibility is the use of non-haem iron halogenases for the stereo- and regioselective halogenation of non-activated sp²-carbon atoms. A potential limitation is the preference of these enzymes for chlorination. A simple general approach bridging the chloride-forming halogenases with the iodide-prefering HMTs would be enzymatic transhalogenation, which uses "transhalogenases" to convert alkyl chlorides to alkyl iodides. In this light, biocatalytic alkylation is expected to move beyond mere alkyl-diversification towards a generalised ability to selectively link two molecular scaffolds.

Future Directions

The inherent chemical instability of SAM under physiological conditions (half-life of about 16 h at pH 8) is a frequently discussed concern, complicating for example the purification of SAM analogues. Intramolecular cyclization results in the formation of 5'-deoxy-5'-methylthioadenosine and homoserine lactone, while depurination results in degradation to adenine and 5'-ribosylmethylthionine (Scheme 4a). Now that efficient systems for SAM analogue regeneration from SAH are available, stability becomes a more significant consideration than the degradation of the coproduct SAH to avoid inhibition of methyltransferases. Huber et al. used an MAT to produce the more stable SAM isosteres S⁴⁰⁴AM (half-life of 21 h), SA'M (half-life of 83 h), and S⁴⁰⁴A'M (no degradation at pH 8) (Scheme 4b). These isosteres could be used by a prototypical Class I model MT but it is currently not known whether HMTs can accept these isosteres in HMT-MT cascade reactions to achieve more regeneration cycles (Scheme 4c). Furthermore, it is exciting to think about applying "bump and hole" strategies for making SAH analogues that are selectively activated by mutant HMTs and used exclusively by mutant MTs, thus creating ideal bioorthogonal alkylation cascades.

Future Directions

The inherently chemical instability of SAM under physiological conditions (half-life of about 16 h at pH 8) is a frequently discussed concern, complicating for example the purification of SAM analogues. Intramolecular cyclization results in the formation of 5′-deoxy-5′-methylthioadenosine and homoserine lactone, while depurination results in degradation to adenine and 5′-ribosylmethionine (Scheme 4a). Now that efficient systems for SAM analogue regeneration from SAH are available, stability becomes a more significant consideration than the degradation of the coproduct SAH to avoid inhibition of methyltransferases. Huber et al. used an MAT to produce the more stable SAM isosteres S⁴⁰⁴AM (half-life of 21 h), SA'M (half-life of 83 h), and S⁴⁰⁴A'M (no degradation at pH 8) (Scheme 4b). These isosteres could be used by a prototypical Class I model MT but it is currently not known whether HMTs can accept these isosteres in HMT-MT cascade reactions to achieve more regeneration cycles (Scheme 4c). Furthermore, it is exciting to think about applying “bump and hole” strategies for making SAH analogues that are selectively activated by mutant HMTs and used exclusively by mutant MTs, thus creating ideal bioorthogonal alkylation cascades.

Future Directions

The inherently chemical instability of SAM under physiological conditions (half-life of about 16 h at pH 8) is a frequently discussed concern, complicating for example the purification of SAM analogues. Intramolecular cyclization results in the formation of 5′-deoxy-5′-methylthioadenosine and homoserine lactone, while depurination results in degradation to adenine and 5′-ribosylmethionine (Scheme 4a). Now that efficient systems for SAM analogue regeneration from SAH are available, stability becomes a more significant consideration than the degradation of the coproduct SAH to avoid inhibition of methyltransferases. Huber et al. used an MAT to produce the more stable SAM isosteres S⁴⁰⁴AM (half-life of 21 h), SA'M (half-life of 83 h), and S⁴⁰⁴A'M (no degradation at pH 8) (Scheme 4b). These isosteres could be used by a prototypical Class I model MT but it is currently not known whether HMTs can accept these isosteres in HMT-MT cascade reactions to achieve more regeneration cycles (Scheme 4c). Furthermore, it is exciting to think about applying “bump and hole” strategies for making SAH analogues that are selectively activated by mutant HMTs and used exclusively by mutant MTs, thus creating ideal bioorthogonal alkylation cascades.

Conclusions

The recently described promiscuous halide MTs have expanded the bioalkylation toolkit and enable selective ethylation, propylation, and alkylation of target compounds. Furthermore, our group has demonstrated that the substrate scope of HMTs
can be expanded, and we believe that a combination of high-throughput screening, facilitated by our halide assays,\textsuperscript{[21,34]} with rational library design, using computational tools like FuncLib,\textsuperscript{[4b,37]} will rapidly expand the range of easily accessible SAM analogues. However, detailed mechanistic understanding of how HMTs function is currently lacking. Determining the crystal structures and catalytic mechanism of more HMTs would enable more rational design, perhaps allowing SAM analogues of currently unimaginable complexity to be enzymatically synthesised from alkyl halides. We believe that HMTs might soon displace MATs from their prominent position in biocatalytic alkylation cascades.

Acknowledgements

Q. T. thanks the China Scholarship Council for financial support of her PhD thesis project (File No.: 201606150073). Open access funding enabled and organized by Projekt DEAL.

Conflict of Interest

The authors declare no conflict of interest.

Keywords: alkylation · alkyl iodide · halide methyltransferase · methylation · SAM analogue
