Methyltransferase directed labeling of biomolecules and its applications

Neely, Robert; Hofkens, Johan; Janssen, Kris; Leen, Volker; Deen, Jochem; Vranken, Charlotte

DOI:
10.1002/anie.201608625

Document Version
Peer reviewed version

Citation for published version (Harvard):
Neely, R, Hofkens, J, Janssen, K, Leen, V, Deen, J & Vranken, C 2017, 'Methyltransferase directed labeling of biomolecules and its applications', Angewandte Chemie (International Edition) , vol. 56, no. 19, pp. 5182–5200. https://doi.org/10.1002/anie.201608625

A Journal of the Gesellschaft Deutscher Chemiker

Angewandte International Edition Chemie

GDCh

www.angewandte.org

Accepted Article

Title: Methyltransferase directed labeling of biomolecules and its applications

Authors: Jochem Deen, Charlotte Vranken, Volker Leen, Robert K Neely, Kris Pieter Frans Janssen, and Johan Hofkens

To be cited as: *Angew. Chem. Int. Ed.* 10.1002/anie.201608625
Angew. Chem. 10.1002/ange.201608625

Link to VoR: http://dx.doi.org/10.1002/anie.201608625
http://dx.doi.org/10.1002/ange.201608625

WILEY-VCH

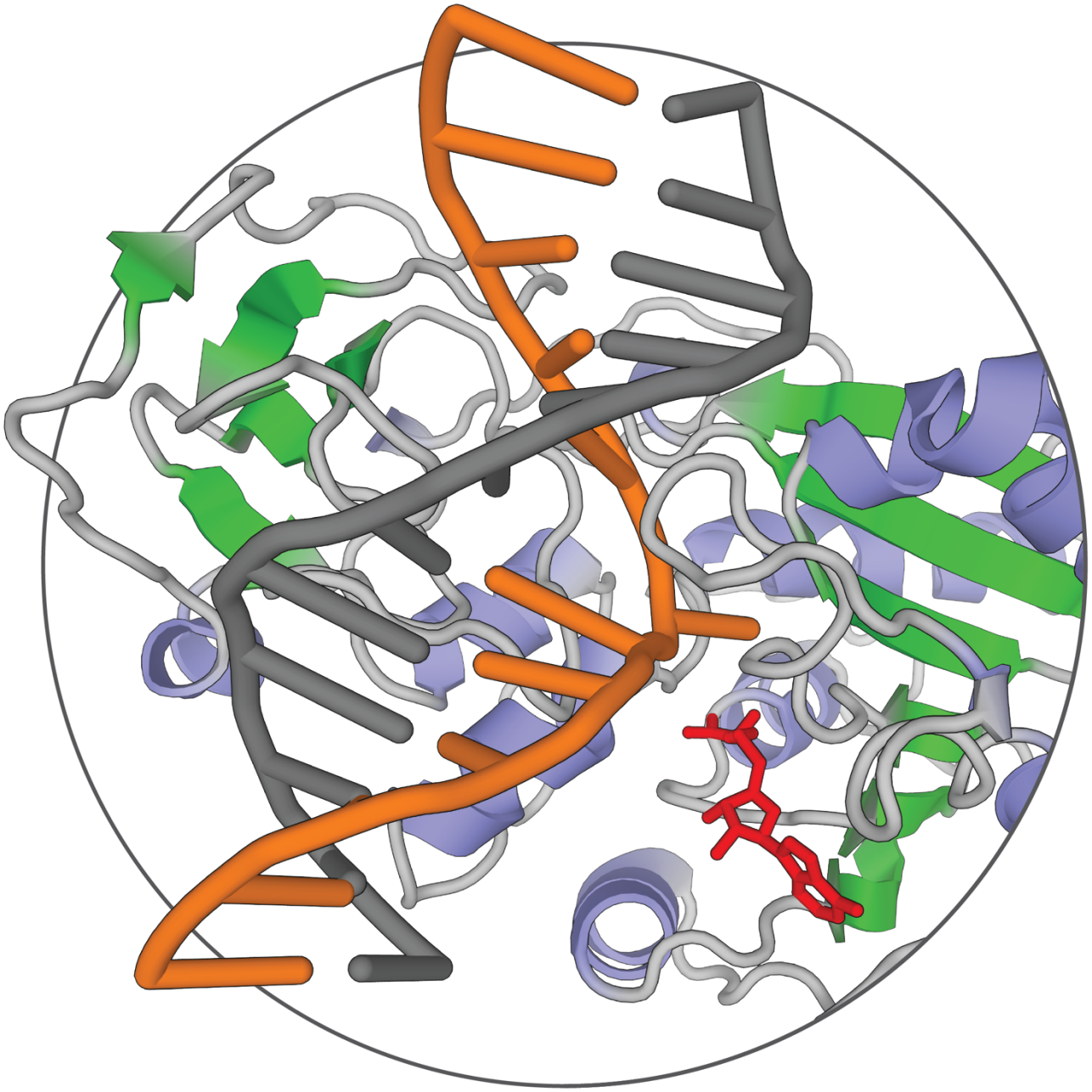
10.1002/anie.201608625

WILEY-VCH

REVIEW

Methyltransferase directed labeling of biomolecules and its applications

Jochem Deen†,[a] Charlotte Vranken†,[b] Volker Leen,[b] Robert K. Neely,[c] Kris P. F. Janssen*,[b] and Johan Hofkens*[b]



10.1002/anie.201608625

Abstract: Methyltransferases (MTases) compose a large family of enzymes which have the ability to methylate a diverse set of targets, ranging from the three major biopolymers DNA, RNA and protein to small molecules. Most of these MTases use the cofactor S-Adenosyl-L-Methionine (AdoMet) as their methyl source. Given the important biological role of methylation, e.g. in epigenetic regulation of gene activity and the vast potential of targeted functionalization in biology, diagnostics and nanotechnology, it should come as no surprise that recent years have seen significant efforts in the development of AdoMet analogues with the aim of transferring moieties other than simple methyl groups. Two major classes of AdoMet analogues currently exist- the doubly-activated and aziridine based molecules- each of which employs a different approach for transalkylation, as opposed to transmethylation, of the target molecule. In this review, we discuss the various strategies for labelling and functionalizing biomolecules using AdoMet-dependent MTases and AdoMet analogues. We cover the synthetic routes to AdoMet analogues, their stability in biological environments and their application in transalkylation reactions. Finally, some perspectives are presented for the potential use of AdoMet analogues in biology research, (epi)genetics and nanotechnology.

1. Introduction

Biological systems are complex and never merely the sum of their basic components. Instead, the intricate interplay between countless numbers of biomolecules gives rise to new, 'emergent' properties which could not have been predicted when considering each component individually.[1] In that respect, life itself is perhaps the most intriguing emergent property of all. Nonetheless, even a tiny change in even a single macromolecule can have profound effects on a living organism. Such is the case in epigenetic regulation of gene activity where small, highly specific modifications to DNA, RNA or interacting proteins will result in significant modification of biological function.

Chemists can utilize the inherent specificity of the enzymatic machinery involved in epigenetic regulation in order to deliver functional- or reporter groups to defined macromolecular targets. As such, this review will focus on a subset of the methyltransferase (MTase) enzymes, the *S*-adenosyl-L-methionine (AdoMet **1**, Scheme 1) dependent methyltransferases. In nature, these enzymes methylate their targets, which range from small molecules to proteins and oligonucleotides. They perform this function by binding the methyl-donor, *S*-adenosyl-L-methionine (AdoMet **1**) and catalyzing the covalent transfer of a methyl-group from AdoMet to their target. Moreover, many MTases have been shown to be sufficiently malleable such that they will perform a similar catalytic transfer of much more extended and functional moieties than the methyl-group. This has allowed them to become an essential tool in the functionalization of the three major biopolymers, RNA, DNA and proteins. The field has been the focus of increasing attention and over the past decade, several new approaches allowing the incorporation of functional chemical moieties onto biomolecules in a site-specific manner have been developed.

[a] Laboratory of Nanoscale Biology, School of Engineering, EPFL, STI IBI-STI LBEN BM 5134 (Bâtiment BM), Station 17, CH-1015 Lausanne, Switzerland

[b] Laboratory of Photochemistry and Spectroscopy, Department of Chemistry, KU Leuven, Celestijnenlaan 200F, B-3001 Heverlee, Belgium. Fax: +32 16 327990; Tel: +32 16 327804; E-mail: johan.hofkens@chem.kuleuven.be

[c] School of Chemistry, University of Birmingham, Edgbaston, Birmingham, B15 2TT, United Kingdom

† Authors contributed equally

Johan Hofkens received his MSc and PhD degrees in Chemistry from the KU Leuven. After postdoctoral research with Prof. Masuhara at Osaka University and Prof. Barbara at the University of Minneapolis, he rejoined the KU Leuven where he started the Single Molecule Unit. In 2005 he was appointed Research Professor at the KU Leuven and in 2008 he was promoted to full professor. His research interests are single molecule spectroscopy, fluorescence and non-linear microscopy and the application of these techniques in timely topics including materials science and biosciences. With respect to these topics he received an ERC advanced grant in 2012.

Kris Janssen, born in 1981, obtained an MSc degree in Bioscience Engineering and Catalysis in 2005 from KU Leuven. After fulfilling positions in private research and industry he returned to KU Leuven in 2008 to pursue a PhD degree in the lab of Prof. Jeroen Lammertyn, focusing on the development of DNA based biosensors. After graduating in 2013 he successfully applied for a position as an FWO postdoctoral fellow and is now working with the group of Prof. Johan Hofkens. His research interests lie in the application of super-resolution fluorescence and electron microscopy to study the interactions of DNA–RNA and proteins with inorganic nanomaterials.

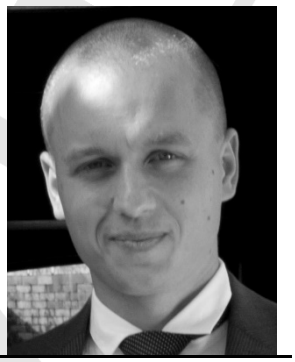
10.1002/anie.201608625

Jochem Deen, born in 1984, received his BSc in engineering from Fontys Eindhoven in 2008 and a MSc in biophysics in 2010 from KU Leuven. He obtained his PhD in the lab of Prof. Johan Hofkens in 2016 co-supervised by Prof. Rob Neely. During his PhD, he focused on the development of DNA mapping using DNA methyltransferases and super-resolution fluorescence microscopes.

Charlotte Vranken, born in 1988, obtained an MSc degree in Chemistry at the KU Leuven in 2011. She obtained her PhD under the supervision of Prof. Johan Hofkens and Wim Dehaen in 2016. Her research was focused on the synthesis of AdoMet analogues for sequence-specific DNA modification.

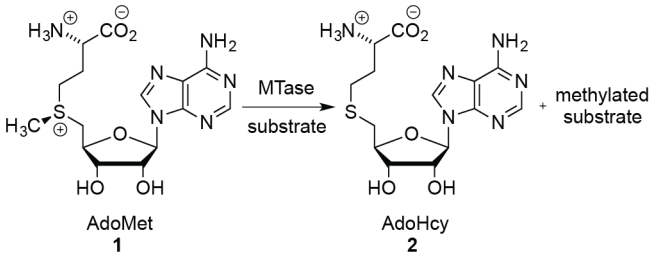
Rob Neely is currently a Senior Lecturer and EPSRC Healthcare Technologies Fellow at the University of Birmingham. He obtained his PhD in 2005, with Prof Anita Jones at the University of Edinburgh. He was awarded an EPSRC postdoctoral fellowship (Edinburgh) and spent a glorious year with Sir Richard Roberts at New England Biolabs, MA working on the DNA methyltransferases. He moved to Prof. Hofkens' group in Leuven, Belgium in 2009 as a Marie Curie Fellow and was appointed as a Lecturer in Birmingham in 2014.

Volker Leen obtained MSc and PhD degrees in Chemistry (Organic Synthesis) from the KU Leuven, having worked on pyrrole and dye synthesis. After postdoctoral studies on antiviral compounds, he joined the research Group of Johan Hofkens. His current research interest is biotechnology development from an organic synthesis point of view.



2. Labeling strategies using AdoMet dependent MTases

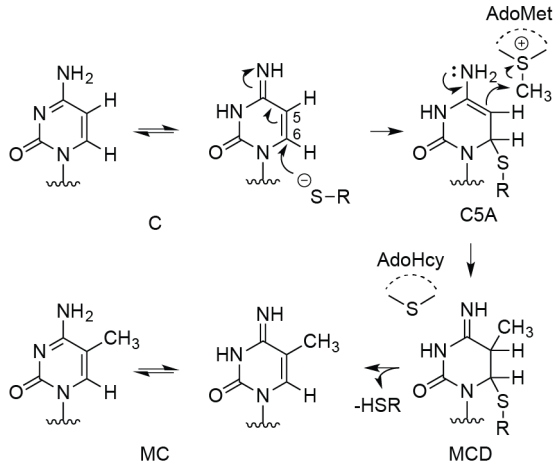
AdoMet dependent MTases catalyze the transfer of a methyl group from their ubiquitous cofactor to a tremendously diverse range of biomolecules (Scheme 1). This makes them fundamental to a wide variety of biological pathways, ranging from small-molecule biosynthesis to protein repair, signal transduction, chromatin regulation and gene silencing.[2]



Scheme 1. Scheme showing the transfer of methyl group from the cofactor AdoMet (1) to the substrate by the MTase.

In order to catalyze methylation, the MTases form ternary complexes with their target molecule and AdoMet. These complexes involve a network of intermolecular interactions but some general principles/chemistries are common to all systems. The transferable methyl group of AdoMet is bound to a sulfonium center. The molecule is inherently unstable towards nucleophilic attack and the methyl group readily participates in substitution reactions.[3] This results in a half-life of 17 hours for AdoMet in the M.*Hha*I reaction buffer at pH 7.4, 37°C[4] and the susceptibility to nucleophilic attack is instrumental to the function of MTases. A mechanism for C5 cytosine methylation by M.*Hha*I, resulting in 5-methyl cytosine (5mC) was first described by Santi *et al*.[5] The authors proposed a two-step concerted mechanism which is shown in Scheme 2, an enzymatically catalyzed version of the Morita-Baylis-Hillman reaction.[5a] In the first step of the reaction, cytosine undergoes nucleophilic attack at C6 by the thiol of a cysteine residue in the catalytic pocket of M.*Hha*I. This results in the 6-Cys-cytosine compound (C5A). Irreversible nucleophilic attack at the transferable methyl group by C5A results in the stable intermediate 5-methyl-6-Cys-5,6-dihydrocytosine (MCD) and simultaneously converts AdoMet to AdoHcy.[5b] Deprotonation and subsequent release of the enzyme of MCD restores the aromatic conjugation and results in the DNA 5-methylcytosine (5mC).[5a] In N6-adenine DNA MTases, the 6-amino group of adenine forms two hydrogen bonds, which increases the electron density of N6 and contributes to its activation for nucleophilic attack.

10.1002/anie.201608625



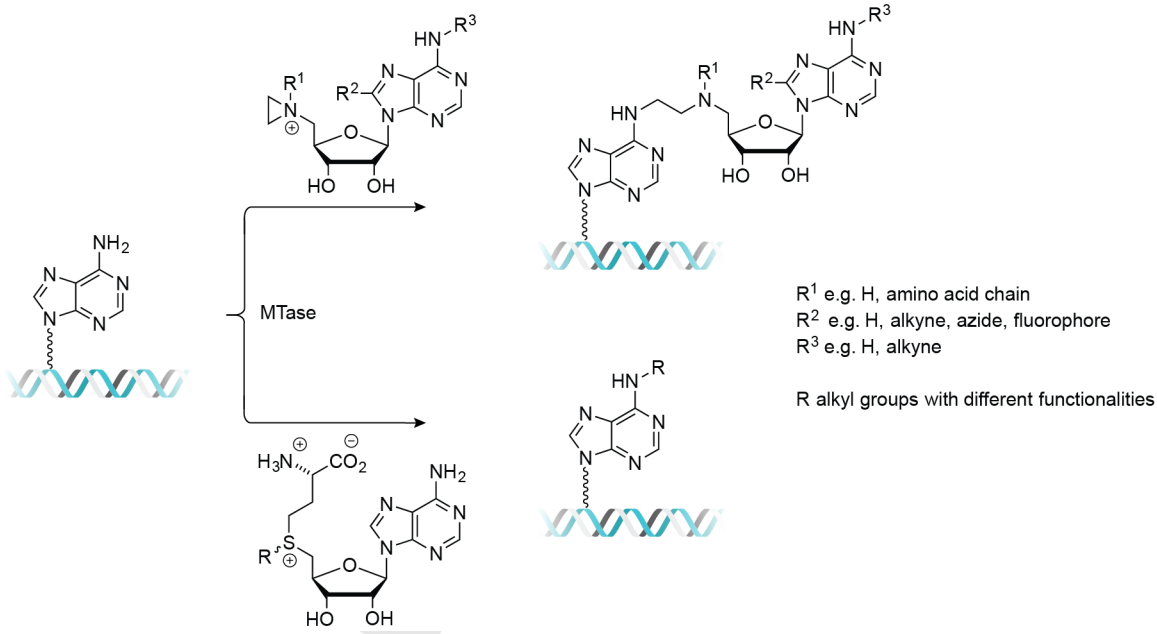
Scheme 2. Reaction mechanism of 5-C methylation by M.*Hha*I (=R).[5a]

This mechanism is not specific to M.*Hha*I and was shown to apply to many other MTases.[3, 6] An S_N2 mechanism is proposed at the sp^3 carbon center, with SAH acting as a good and stable leaving group. This has proven instrumental in allowing the development of a large variety of artificial AdoMet analogues.

2.1 AdoMet analogues

AdoMet is a structural hybrid of methionine and adenosine for which both the *R*- and the *S*-diastereoisomers bind similarly to MTases. However, only the *S*-diastereoisomer has the correct geometry at the sulfonium center to allow for proper catalytic function.[7] In nature, the active *S*-diastereoisomer is formed upon stereospecific reaction of L-methionine (L-Met) and adenosine triphosphate (ATP), catalyzed by the methionine adenosyltransferase enzymes (MAT).[8]

The vast majority of artificial AdoMet analogues can be categorized into two major groups: aziridinoadenosines and doubly-activated AdoMet analogues (Scheme 3). To date, the aziridinoadenosines have only been prepared synthetically. However, the doubly-activated AdoMet analogues can be prepared via either total synthesis or enzymatically, using a synthetic methionine analogue and a MAT enzyme.[9]



Scheme 3. DNA labeling using aziridine based or doubly activated AdoMet analogues.

Aziridine based AdoMet analogues

Aziridine-based AdoMet analogues (Scheme 4) constitute some of the earliest examples of AdoMet analogues and were first developed by the group of Weinhold.[10] These AdoMet analogues differ from natural AdoMet as their 5'-sulfonium is replaced by an aziridine ring (**11**, Scheme 8). This aziridine group ring-opens as a result of a nucleophilic attack from the target compound. Since the 'leaving group' is conjugated to the transferable moiety in this case, following the S_N2 reaction mechanism to completion sees the transfer of the entire cofactor analogue to the target molecule (DNA in Scheme 3).[7]

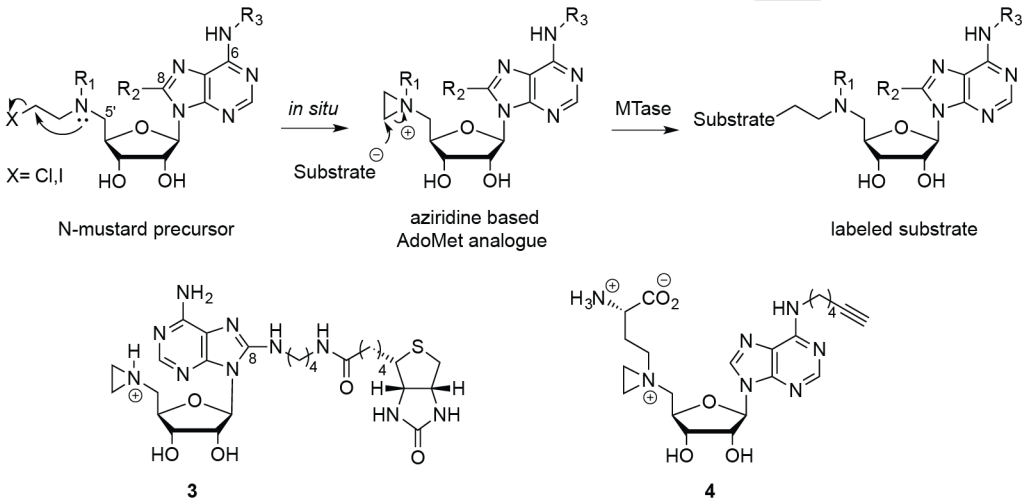
Several DNA MTases were found to be capable of transferring these compounds to DNA. The process of introducing reporter groups (e.g. biotin, **3**, Scheme 4) to a variety of DNA sequences has been termed 'Sequence-specific Methyltransferase Induced Labeling' (SMILing). Many different reporters such as e.g. fluorescent dyes have been attached to the adenine moiety of this aziridine AdoMet analogue for application in DNA labeling.[10b, 11] Modelling of the AdoMet-binding pocket from crystallographic data showed that steric interactions between the cofactor analogues and the enzyme are likely a key factor determining compatibility of AdoMet analogues with specific

10.1002/anie.201608625

methyltransferases. As an example, modelling of the M.*Taq*I AdoMet binding pocket showed that the 8-position of the adenine moiety of a bound cofactor is accessible to the solvent. This implies that the enzyme is therefore likely to tolerate cofactor analogues that incorporate bulky modifications at this site. Successful examples of such bulky modifications to the cofactor analogue include an azide group (**12**, Table 1, SI) and a range of fluorophores.[11-12] Pljevaljčić *et. al.* identified similar opportunities for modification of the adenine moiety (6-, 7- or 8- position) for a range of MTases from their crystallographic structures.

Further work on the aziridine-type AdoMet analogues has focused on substitutions at the 5'-N to circumvent the production of the rather unstable aziridine-based AdoMet analogues.[13] Indeed, introduction of an amino acid side chain at the 5'-N position (**13**, Table 1, SI) enhances the MTase catalyzed DNA alkylation and generation of small molecule derivatives.[14] These analogues can easily be synthesized *in situ* from the N-mustard precursor, which undergoes an intramolecular cyclization to form the aziridine-based cofactor (Scheme 4). Similarly, inclusion of a functional group such as an azide (**14, 15**, Table 1, SI) or a terminal alkyne (**4**, Scheme 4, **16** and **17**, Table 1, SI) enables the use of these N-mustard cofactor analogues in bio-orthogonal ligations.[12-13]

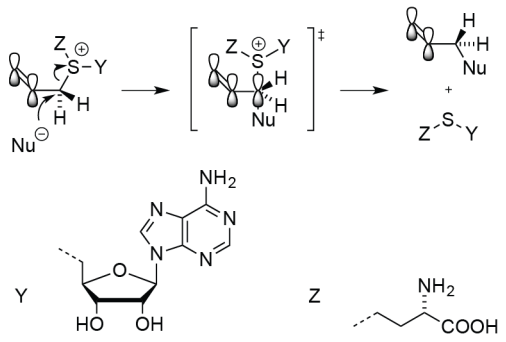
The use of the aziridine-based cofactor analogues for transalkylation reactions suffers from several drawbacks. Most notably, the transalkylation is not catalytic and stoichiometric amounts of the MTase must be used in the reaction. This was first observed by Osborne *et al.* who demonstrated the protein methyltransferase 1-catalysed (PRMT1) transalkylation of a peptide using a N-mustard-based cofactor. The product of this reaction, essentially a peptide with a covalently-bound cofactor analogue, is a potent inhibitor for the MTase.[15] Hence, the reaction is self-limiting. Furthermore, the aziridine-based analogues are very reactive. This makes them prone to rather rapid degradation and also results in an increased propensity to effect non-specific alkylation, even in the absence of MTases.[7]



Scheme 4. Aziridoadenosine based AdoMet analogues.

Doubly activated AdoMet analogues

MTases have an inherent ability to catalyze the transfer of alkyl groups larger than methyl. However, for DNA MTases, the transfer rate decreases rapidly with increasing size of the transferable moiety (methyl > ethyl > propyl). The reaction occurs via an S_N2 mechanism with inversion of the configuration at the α-carbon next to the sulfur atom.[16] As such, incorporation of an unsaturated bond at the β-position can potentially stabilize the p-orbital of the α-carbon formed at the intermediate stage of this substitution reaction. Weinhold and Klimašauskas were able to demonstrate that an allylic or a propargylic carbon-carbon bond at the β-position relative to the sulfonium center can restore the reaction rate by conjugative stabilization of the S_N2 transition state (Scheme 5).[17]



Scheme 5. Stabilization of the S_N2 transition state.[17b]

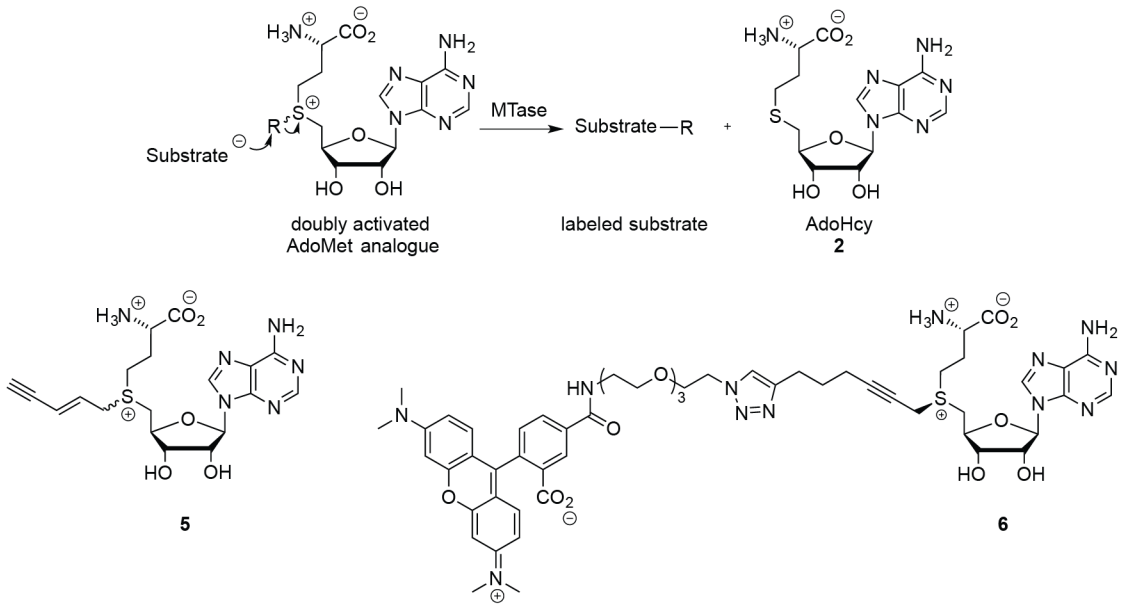
10.1002/anie.201608625

This approach to labeling, using so-called 'doubly-activated' cofactor analogues has been termed 'methyltransferase-directed Transfer of Activated Groups' (mTAG).[18] Unlike aziridine based labeling, mTAG is a truly catalytic process and does not require stoichiometric amounts of the MTase.[17b] Even though multiple turnovers per minute can be achieved, the process is still typically an order of magnitude slower than the equivalent methyl-transfer.[17b]

AdoMet analogues with alkyl, alkenyl (**19**, see Table 2, SI) and alkynyl (**20**, see Table 2, SI) side chains have been prepared and successfully applied in mTAG.[17]

The approach offers a simple way to label targets with a functional group that can be used for further ligation of more complex moieties. For example, incorporation of different transferrable amine moieties (**21**[4, 18], **22**[4, 19], **23**[4, 19b] Table 2, SI), can be followed by a -coupling reaction with the NHS-ester of an amine-reactive probe.[18] In addition, several examples exist on the application of AdoMet analogues with a transferrable terminal alkyne (**5**[20], Scheme 6, **20**[20g, 20h], **24**[20b, 20c, 20g], and **25**[4, 19b], Table 2, SI) or azide (**26**[20g, 21], **27**[4]). Transferred alkynes and azides can be further functionalized using the bio-compatible and highly-efficient azide-alkyne cycloaddition reaction (one of the 'click' series of reactions). However, some of these analogues, such as compound **20** (Table 2, SI), featuring a terminal alkyne, are highly unstable.[4, 20g, 20h] Other AdoMet analogues, such as the AdoEnyYn compound (**5**, Scheme 6), are more stable, making them more suitable for use in bio-orthogonal ligation.[20b-h]

Ketones are yet another interesting transferrable moiety as they can be used to react with hydroxylamines and hydrazides. This was demonstrated with AdoMet analogue **28** (Table 2, SI), which was successfully transferred to DNA and subsequently used for coupling with a hydroxylamine coupled to a fluorophore.[22] Finally, AdoMet analogues directly incorporating a fluorescent dye allow for single-step, direct labeling reactions. One such example in which a TAMRA dye was coupled to the AdoMet analogue (**6**, Scheme 6) and subsequently used for labeling reactions was first described by Grunwald *et al*.[23]

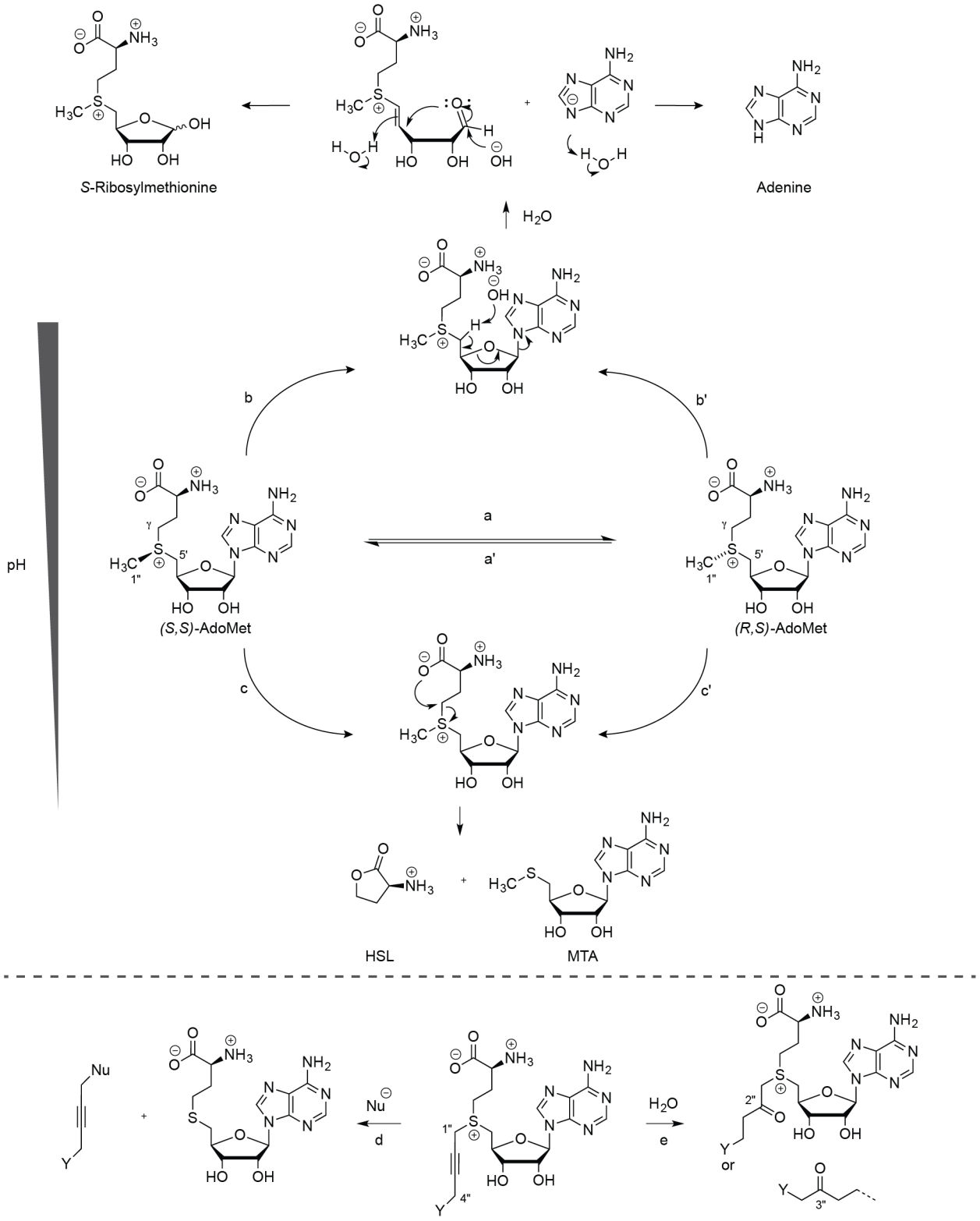


Scheme 6. Doubly activated AdoMet analogues.

2.2 Stability of AdoMet analogues

AdoMet analogues are prone to spontaneous decomposition in aqueous environments through a range of different pathways (Scheme 7).[4, 24] Reversible racemization (route a) to the *(R)*-diastereomer of AdoMet is also common and can result in the formation of an inactive AdoMet analogue. The presence of the sulfonium center activates the adjacent carbon atoms towards decomposition reactions. In alkaline conditions, deprotonation at the 5'-C occurs (route b), which results in formation of adenine and *S*-ribosylmethionine. Under more acidic conditions, intramolecular attack of the α-carboxylate group onto the γ-C of the methionine group (route c) yields methylthioadenosine (MTA) and homoserine lactone (HSL). In the case of AdoMet, no nucleophilic attack is observed at the methyl group, but this changes when considering AdoMet analogues with extended chains containing a triple bond in β-position, relative to the sulfonium center. The partial positive charge at the 1"-C increases making it more susceptible to nucleophilic attack (route d). When electronegative groups (e.g. X= NH_2) are in close proximity to the triple bond, this can lead to a higher electron deficiency at both 4"-C and 1"-C, which enables the addition of water (route e) at both and results in the hydrated compound showing almost no reactivity towards most DNA MTases and protein MTases.[4, 24]

10.1002/anie.201608625



Scheme 7. Decomposition pathways of (S,S)-AdoMet (analogues). a: Inversion at the sulfonium center of *(S,S)*-AdoMet results in *(R,S)*-AdoMet. b: deprotonation at the C-5' and subsequent elimination of adenine base to give *S*-ribosylmethionine. c: nucleophilic attack of the α-carboxylate onto the γ-carbon of methionine delivers HSL and MTA. d: nucleophilic addition at 1"-C. e: addition of water on the 2"-C or 4"-C position.

10.1002/anie.201608625

For a pH of 7.5 and a temperature of 37 °C, the rate constants for racemization, cleavage to HSL/MTA and hydrolysis to adenine /S-ribosylmethionine were reported to be 1.8 x 10^{-6} s^{-1}, 4.6 x 10^{-6} s^{-1} and 3 x 10^{-6} s^{-1} respectively. The hydrolysis rate shows a significant decrease at lower pH values.[25]

Recently, Huber *et al.* developed novel AdoMet analogues lacking the N7 of the adenine moiety and further omitting the carboxylic acid of the amino acid chain in favor of a tetrazole ring. This makes them less likely to decompose through both pathways b and c and restricts epimerization at the sulfonium center.[26]

To improve the stability and reactivity of the AdoMet analogues, several molecules have been studied where sulfur is replaced with another chalcogen, such as selenium (SeAdoMet, **7**, Figure 1) and tellurium (TeAdoMet, **8**, Figure 1). For AdoMet, three decomposition pathways are common: racemization to the inactive diastereomer, deprotonation of the 5'-C to yield S-ribosylmethionine and adenine or intramolecular nucleophilic attack of the carboxylate group to result in HSL and MTA. SeAdoMet (**7**, Figure 1) was found to decompose via two of these reactions whereas the tellurium analogue TeAM (**8**, Figure 1) was inert to all three pathways.[24] It could be shown that the trend in electrophilicity of these AdoMet analogues followed SeAdoMet > AdoMet > TeAdoMet and that the 5'-C acidity follows the series AdoMet > SeAdoMet> TeAdoMet.[27] Selenium analogues are therefore both more reactive towards nucleophilic attack, making them better suited as cofactors for transalkylation and more stable towards deprotonation of the 5'-H (pathway b, Scheme 7). As a result, they have been subject of extensive investigation. Weinhold *et al.* have developed SeAdoYn (**9**, Figure 1) as a cofactor for protein methyltransferase directed fluorescent labeling of proteins. This analogue showed improved reactivity and higher stability toward hydrolysis compared to the similar sulfur-containing AdoEnYn and AdoYn analogues.[20h] Comparative decay studies of propargylic SeAdoYn (**9**, Figure 1) and ProSAM (**20**, Table 2, SI) analogues showed that decomposition of the ProSAM compound results firstly in the hydrated product (keto byproduct, **28**, Table 2, SI), which further decomposes to the thioether. However, the SeAdoYn follows a different mechanism and directly forms the selenoether.[20h, 28] SeAdoYn was used as a substrate for transalkylation by a large variety of wild-type MTases, which is in contrast with many of the AdoMet analogues, with larger transferable groups, that were only active with mutated MTases.[9b, 19a, 20h, 28-29]

The group of Luo has also demonstrated that selenium substitution can result in increased reactivity of the sulfur-centered AdoMet analogues when applied as substrates for the protein MTases. Their study shows that the decomposition rate of SeAdoMet is 10-fold higher than that of AdoMet, but the protein MTase-catalyzed transalkylation reaction is only 3-5 fold faster for the SeAdoMet analogue. Hence, the authors suggest that the protein MTase-catalyzed reaction rates are not determined solely by the strength of the chalcogen-carbon bond, but are likely caused by other factors.[28, 30] Interestingly, the study also suggests that the β-sp^2 carbon, which is essential for activity with *S*-alkyl AdoMet analogues, is likely not required for some protein MTases when Se-alkyl AdoMet analogues are used.[30]

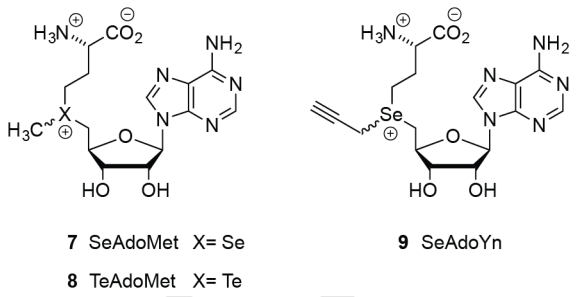
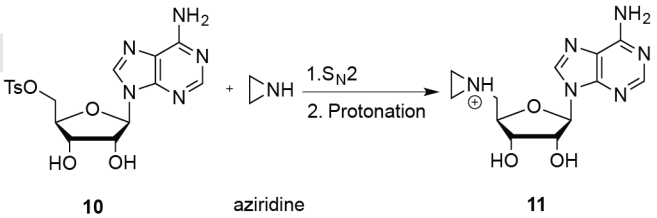


Figure 1. Chalcogen containing AdoMet analogues.

2.3 Synthesis of AdoMet analogues

The first example of the synthesis of aziridine-based AdoMet analogues was described by Pignot *et al.* in 1998 who focused on the development of aziridine-based cofactor analogues.[10a, 31] N-Adenosylaziridine was synthesized using a nucleophilic substitution of the tosylate group in 5'-deoxy-5'-tosyladenosine with aziridine and subsequently activated as an alkylating agent by protonation of the nitrogen atom in the aziridine ring (Scheme 8).[10a]



Scheme 8. Synthesis of *N*-Adenosylaziridine AdoMet analogue. Step 1: Nucleophilic substitution of the tosylate group in 5'-deoxy-5'-tosyladenosine with aziridine.

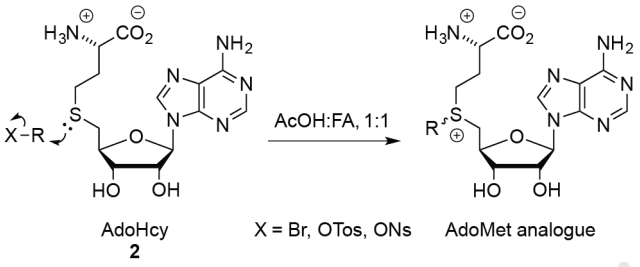
The ring-strain in this three membered heterocycle makes it susceptible to nucleophilic ring-opening, which is further facilitated by nitrogen quaternization. This aziridine based AdoMet analogue can be used for the introduction of different functional or reporter groups through modification of the adenine moiety. One such example shows the attachment of a dansyl fluorophore at the adenine 8-position.[11] The synthesis starts from 8-bromo-2',3'-*O*-isoporpylideneadenosine, which was treated with diaminobutane to introduce a flexible linker and was in a few steps (substitution and deprotection) converted to the *N*-adenosyl-aziridine derivative. A similar route has been used to prepare the N-mustard precursors of the aziridine-based AdoMet analogues (Scheme 4).[32]

These aziridine-based AdoMet analogues suffer from a long- and low yielding-synthetic route.[10, 12b, 13-14] Townsend *et al.* have reported improved synthetic routes for the synthesis of N-chloromustard-substituted adenosines using reductive amination as the key step and have also demonstrated the synthesis of a photo caged derivative, which benefits from increased stability

10.1002/anie.201608625

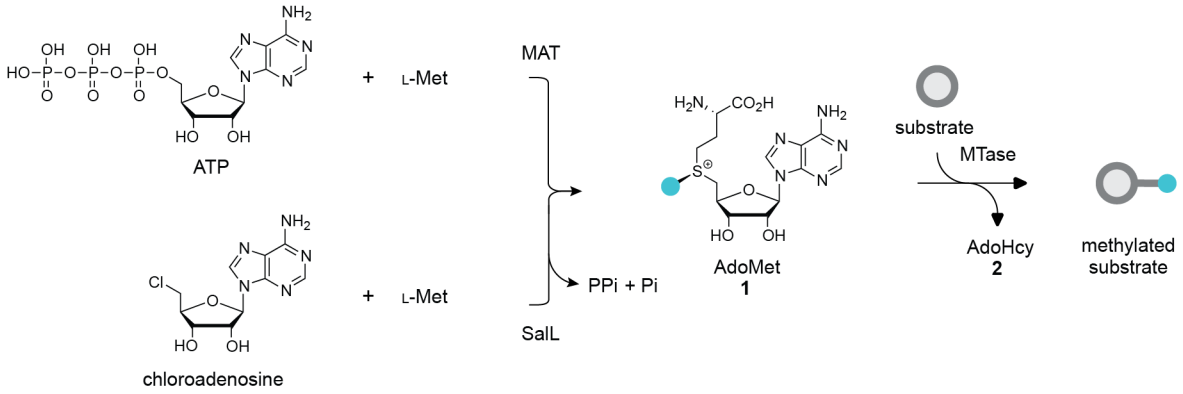
compared with standard aziridine-based cofactors and is readily activated as a cofactor using UV-irradiation.[33]

Doubly-activated AdoMet analogues are typically synthesized by combining *S*-adenosyl-L-homocysteine (AdoHcy, **2**, Scheme 9) with an excess of strong electrophiles, e.g. alkyltriflates or alkylbromides, in acidic conditions.[34] The first chemical synthesis of the diastereomers of AdoMet from AdoHcy was published in 1959.[35] In acidic environment, the propensity for nucleophilic attack by e.g. AdoHcy amines, hydroxyls or the carboxyl acid is strongly reduced, leaving the thioether as the sole reactant.[17a] Both the *(R)*- and the *(S)*-epimer are formed during this S_N2 reaction. However, as only the active *(S)*-epimer can be used in transmethylation reactions, a separation of the diastereomers is preferable. Reverse-phase HPLC (RP-HPLC) is commonly used in efforts to separate the diastereomers. This, however, remains challenging.



Scheme 9. Synthesis of doubly activated AdoMet analogues

The synthetic routes to AdoMet analogues are not trivial. In addition, these compounds often display limited stability. As a result, there have been several attempts have to prepare both AdoMet and its analogues enzymatically.



Scheme10. Enzymatic AdoMet synthesis using SalL and MAT

AdoMet itself can be obtained by isolation from yeast, grown in media supplemented with L-methionine.[8b] Small-scale[8a, 36] (μmol) enzymatic syntheses of AdoMet from ATP and L-methionine (L-Met) as well as larger scale[8b] (mmol) synthesis have been reported. The AdoMet formation is catalyzed by L-methionine *S*-adenosyl transferase enzyme (MAT or AdoMet synthetase) in a two-step fashion where the complete tripolyphosphate (TPP) chain of ATP is cleaved and subsequently hydrolyzed to pyrophosphate (PPi) and phosphate (Pi). The enzymatic synthesis results in high yields of the preferred *(-)*-epimer (Scheme 10, top).[36a, 37]

Recently, several groups have demonstrated enzymatic synthesis of AdoMet analogues as well.[9, 38] The group of Thorson tested a range MAT enzymes for their ability to catalyze the formation of several AdoMet analogues. They tested a library of forty-four non-native S/Se l-Met analogues with five different MATs. Of these, human MAT II was the most permissive. Furthermore, the synthesized analogues could be successfully used in the alkylation of small molecules (indolocarbazoles).[9b] The group of Burkart demonstrated the chemoenzymatic synthesis of several AdoMet analogues using either a fluorinase (FDAS) from *Streptomyces cattleya* or a chlorinase (SalL) from marine bacterium *Salinaspora tropica*.[38-39] Both halogenases are known to catalyze the breakdown of AdoMet, resulting in L-Met and fluoro-5'-deoxyadenosine (FDA) or 5'-chloro-5'-deoxyadenosine (CIDA). However, the reaction can be reversed at low chloride/fluoride and high L-Met concentrations (Scheme 10, bottom). This enzymatic pathway can also be exploited using various L-Met derivatives for the enzymatic synthesis of AdoMet analogues.[9a, 38, 40] but it should be noted that decreased activity of the L-Met analogues with increasing size of the L-Met analogue was observed.[40]

2.4 Methyltransferases

Methyltransferases are typically categorized in 5 distinct families (classes I-V) based on structural similarities.[2a] Of these, Class I MTases, are by far the largest group. They catalyze the majority of methylation reactions and include all DNA MTases, some protein MTases and RNA MTases.[2a] The largest group of protein MTases are the protein lysine MTases, and together with the protein arginine MTases, they play an important role in histone modification and thus ultimately in gene transcription. Most of the work in the past decade using modified AdoMet analogues has focused on class I MTases and class V MTases, which are also the focus of this review. Other, less common, classes include the MetH reactivation domain (class II), precorrin-4 MTases (class III) and the SPOUT family of RNA MTases (class IV).[2]

While the MTases share little sequence similarity, they do share a highly conserved structural fold. The core element of this conserved fold are seven stranded β-sheets with three helices on each side, a structural feature commonly referred to as the Rossman fold.[2b, 41] This core is shared among MTases that act on all different substrates, ranging from small molecules to DNA and proteins.[3, 42] This structural similarity suggests that the labeling reactions with AdoMet analogues are universal and applicable across the entire range of MTases.

In labeling reactions, wildtype MTases are most frequently used in combination with an AdoMet analogue. However, in some cases this is not possible because the native fold of the MTases is incompatible with the AdoMet analogue. This could be due to bulky groups on the AdoMet analogues that prevent a good steric fit or disruption of the needed binding interaction to the AdoMet binding pocket. In these scenarios it could prove useful to engineer the AdoMet binding pocket for a more favorable interaction with the AdoMet analogues, for example by directed evolution of the MTase.[43]

In other cases, it might be preferable to use an MTase that interacts more readily with the AdoMet analogue rather than the natural AdoMet. For example, when using the labeling reaction in live cells in the presence of AdoMet. In this scenario it would be useful to engineer the AdoMet binding pocket such that the enzyme prefers the AdoMet analogue over the natural AdoMet. A typical approach in such a case is to aim for a 'bump-and-hole' strategy.[44]

DNA transalkylation

In bacteria, the DNA MTases play a role in defense of the bacterium against viral invasion by enabling the distinction between their host genome and invading viral DNA to be made.[6, 45] In eukaryotic cells the MTase's main role is in the regulation of genes.[6] DNA methyltransferases can be subdivided depending on their target for modification (cytosine C5, cytosine N4 or adenine N6). Of these groups the cytosine C5 and adenine N6 MTases are found in many species of fungi, bacteria and protozoa, while the cytosine N4 MTases occur only in bacteria.[6, 46]

The DNA MTases typically recognize short, palindromic DNA sequences (2-8 bases long). Ordinarily, the target base for methylation lies buried within the DNA helix and as a result, the DNA MTases extrude this base from the DNA duplex, flipping it into the enzyme's catalytic pocket, where transmetylation occurs (Figure 2).[47] The byproduct of this catalysis, *S*-adenosyl-L-homocysteine (AdoHcy, **2**, Scheme 9), is subsequently released from the AdoMet binding pocket and, in cells, digested by a AdoHcy hydrolase enzyme.[48]

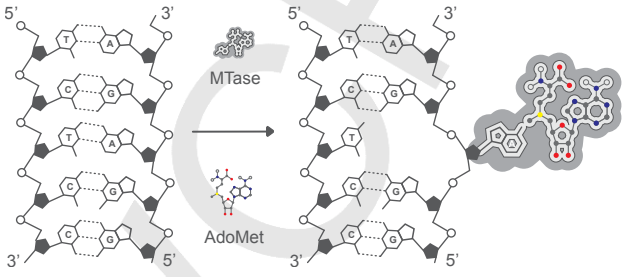


Figure 2. Base flipping mechanism. Most of the work on DNA MTases with AdoMet analogues is based on M.*Taq*I, a DNA MTase from a thermophilic bacteria.

DNA MTases were the first MTases that were shown to catalyze transalkylation reactions using the AdoMet analogues. Part of the reason for this is likely the ease with which the progress of the labeling reaction can be followed. Every bacterial MTase has a partner restriction endonuclease that cleaves DNA at the same site targeted by the MTase. However, if that site is methylated (or alkylated) cleavage is prevented. This can be exploited to readily confirm the activity of the methyltransferase with the AdoMet analogue using gel electrophoresis.

The first example of a transalkyation reaction using an AdoMet analogue and a DNA MTase was the aziridine transfer with the M.*Taq*I DNA MTase in 1998.[10a]. Many aziridine analogues have subsequently been successfully transferred to DNA by, for example, adenine DNA MTases M.*Eco*RI[12b] and M*Bsec*I[49] and cytosine DNA MTases such as M.*Hha*I[12a] and *M.Sss*I[14a]. In these cases functional moieties (including fluorophores[11]) were placed on various locations of the aziridine cofactor. The aforementioned study by Pljevaljcic *et. al.*, which models cofactor binding pockets of several DNA MTases together with the cofactor analogues, concludes that 4 of the 6 screened enzymes (M.*Taq*I, M.*Rsr*I, M.*Dpn*M, M.*Pvu*II) are likely to tolerate cofactors carrying bulky modifications at the 8-position of the cofactor's adenine moiety. For M.HhaI and M.MboII the 7 or 6 positions of that same adenine moiety appear to be feasible sites for the attachment of bulky modifications.[10b] However, to the best of our knowledge, only the M.*Taq*I and M.*Hha*I enzymes have been shown to transfer an aziridine AdoMet analogue to DNA. Unpublished work in our lab failed to show any transfer with the M.*Pvu*II enzyme and aziridine analogues carrying modifications at the adenine 6-, 7-, and 8-positions.

As previously mentioned, the aziridine AdoMet analogues suffer from some disadvantages, which has inspired Weinhold and Klimašauskas to develop the doubly-activated AdoMet analogues. These cofactors have been employed in successful transalkylation reactions with several DNA MTases, with either adenine or cytosine as a target base. A comprehensive overview

10.1002/anie.201608625

can be found in Table 1. The attachment of a functional or fluorescent group to the DNA can be achieved using a second chemical coupling reaction, such as amine to NHS ester[19c] or click-chemistry-based conjugation.[20f] Alternatively, like the aziridine cofactors, the fluorescent group can be synthetically coupled to the cofactor, thus allowing one-step transfer of the fluorescent groups to the DNA substrate. This approach was first demonstrated with the DNA MTase M.*Taq*I.[23]

Rates of the transalkyation reaction with the doubly-activated AdoMet analogues have been dramatically improved by engineering of the cofactor binding pocket of the cytosine C5 MTases.[19d] Three amino acid side chains were selected based on their potential steric interaction with the cofactor, and replaced with shorter residues (two residues were replaced with alanine, while one residue was replaced with the smaller polar residue serine). The mutants showed a significant improvement in both binding affinity and the transfer rate of the AdoMet analogues, relative to the wild-type M.*Hha*I. In general, this effect was more marked for the longer cofactor analogues with transferrable groups with longer alkyl-chains than those with shorter chains. Furthermore, the mutant enzymes show a significant reduction in the methylation rate with the natural cofactor AdoMet. The binding constant of the enzyme with the natural cofactors AdoMet and AdoHcy was reduced as well. The weaker binding of AdoHcy may contribute to the increased catalytic efficiency of the AdoMet analogues.

Two of three mutations were located within the highly conserved sequence motifs of M.*Hha*I and were readily mapped to locations on other C5- DNA MTases. Indeed, the same mutations were successfully applied to cytosine-5 MTases M.*Hpa*II and M2.*Eco*31I, and later to the CpG-specific MTase M.*Sss*I.[50] However the same mutations did not result in an improved labeling with M.*BsaH*I.[20f]

Table 1. DNA MTases (on the right) shown to be compatible with various AdoMet analogues and the functionalities implemented on the AdoMet analogue

		MTAG					Aziridine				
	Target	Alkyl	Alkyne	Amine	Azide	Fluo	None	Alkyne	Azide	Biotin	Fluo
M.TaqI	TCGA	[17]	[17b, 20f],	[18]		[23]	[10-11, 12b, 14a]	[13, 14b]	[12b, 14b]	[12b, 49a]	[10b, 11, 51]
M.HhaI[a]	GCGC	[17b]		[18, 19c, 19d]	[4]		[12b]	[14b]	[12b, 14b]	[52]	
M.SssI[a]	CG			[50]	[50]						
M.BseCI	ATCGAT									[49]	
M.BcnIB[a]	CCSGG	[17b]	[17b]								
M2.Eco31I[a]	GGTCTC		[19d]	[19d]							
M.EcoRI	GAATTC						[12b, 14a]	[13]	[12b]		
M.HpaII	CCGG		[19d]	[19d]			[12b, 14a]				
M.XbaI	TCTAGA		[20f]								
MFokI	GGATG and CATCC		[20f]								

[a] Mutant enzyme with improved activity in the presence of a AdoMet analogue

10.1002/anie.201608625

RNA transalkylation

In this section we present a brief overview of the application of the of the RNA methyltransferase enzymes, which have been shown to catalyze transalkylation, using AdoMet analogues, of RNA. Early studies show some promise in this area but all have been realized in vitro using synthetically-prepared or in-vitro transcribed RNA substrates. Whether the methyltransferases can replace existing antibody-based approaches for targeted labelling or capture of RNA-based substrates remains an area for further investigation.

The first example of an RNA methyltransferase-directed transalkylation was with the tRNA methyltransferase, Trm1. [20d] This enzyme has been shown to catalyze the transalkylation of the N2 of guanosine 26 in $tRNA^{Phe}$ using a doubly-activated SAM analogue with an alkyne functionality. The modified tRNA was subsequently fluorescently labelled using the copper catalyzed azide-alkyne cycloaddition reaction (one of the 'click' reactions), in order to generate fluorescently labeled tRNA.

Using in vitro reconstitution of an archaeal box C/D small ribonucleoprotein RNA 2'-*O*-methyltransferase (C/D RNP) Tomkuvienė *et al*. were able to demonstrate transalkylation of in vitro transcripts of tRNA and pre-mRNA molecules. The C/D RNP complex includes a guide RNA molecule that is used to direct the specificity of the C/D RNP to these non-natural substrates. Again, in a second reaction, the alkylated RNA molecules were fluorescently labeled using the copper catalyzed azide-alkyne cycloaddition reaction.[19a]

Functionalization of both miRNA and small interfering RNA (siRNA) has been described by Plotnikova. *et. al*.[19b] Here, the HEN1 2'-*O*-methyltransferase from *Arabidopsis thaliana* was used to direct transalkylation to the 3'-terminal nucleotides of small double stranded RNA molecules. In fact, this study demonstrates the direct transfer of biotin to the 3'-ends of small RNA duplexes and their subsequent isolation using streptavidin-coated beads.[19b]

Protein transalkylation

The two most common AdoMet dependent protein MTases are protein arginine MTases and protein lysine MTases.[53] However, there are some protein MTases which target other sites such as other peptidyl chains, N- or C-termini.[54] The main targets of arginine and lysine protein MTases are histones, where methylation of the histone is associated with repression of transcription.[55] Besides histones, numerous other proteins have also been targets of methylation, where they play a role in many physiological paths such as signal transduction and protein translocation.[56] Because of the close role of histone methylation with repression of transcription, protein MTases have been implicated in cancer, neurodegenerative and other diseases.[57]

One notably early study in 2001 demonstrated that a mutant of the yeast MTase *Rmt*1 was selectively inhibited using N^6-substituted AdoMet analogues. This was the first successful demonstration of the bump-hole technique as a means to develop selective combinations of mutant methyltransferase enzymes and tailored AdoMet analogues.[44] The aziridine-based cofactor analogues and the PRMT1 protein methyltransferase have been successfully employed for protein transalkylation.[15, 58] The doubly-activated (mTAG) cofactors were first employed in protein transalkylation reactions by Peeters *et. al*.[20e] This, and a later study, describe the transfer of an alkyne group to the target of the two histone H3K9 protein MTases Dim-5 and SETDB1.[59]

The group of Luo, have developed combinations of several AdoMet analogues and engineered protein MTases to enable bio-orthogonal targeting of transalkylation reactions in cells. Their work is covered in detail in a recently published review[7] but we provide a brief overview here. The focus of their work has been the human protein MTases G9a (Also known as EuHMT2), GLP1 (also known as EuHMT1) and PRMT1 (for a complete overview see table 2). Enzyme-directed transalkylation reactions using azido-AdoMet with engineered G9a and GLP1[21, 60] and an alkyne-AdoMet with engineered G9a[20c] and PRMT1 have been demonstrated.[20g] Additionally, a Selenium-based SAM analogue with an alkyne linker was effectively employed in transalkyation reactions directed and catalyzed by the native protein MTases GLP1, G9a and SUV39H2.[28, 60-61]

An extensive study with 8 native enzymes showed limited transalkylation by the 3 wild type protein MTases G9a, GLP1 and SUV39H2 using an allyl-AdoMet analogue, and a complete absence of activity with the bulkier AdoMet analogues.[62] However, a single mutation in these three enzymes was shown to have a dramatic impact on their ability to catalyze transalkylation reactions and, furthermore, led to a significant reduction in the rate at which they performed methylations, using AdoMet. Hence, these mutants were found to be capable of performing transalkylation reactions with synthetic AdoMet analogues in the presence of native AdoMet (e.g. in cell lysates).[63] In a subsequent study, Islam *et. al.* investigated the function of the mutations (Y1211A in EuHMT1, Y1154A in EuHMT2) which further improved the enzymatic transalkylation rates using the AdoMet analogues. The residues targeted for mutations were branded as gatekeeper amino acids that had blocked access to the cofactor binding pocket for the bulkier, synthetically-prepared AdoMet analogues.[20b]

10.1002/anie.201608625

Table 2. protein MTases shown to be compatible with various AdoMet analogues and the functionalities implemented on the AdoMet analogue.

	MTAG			Aziridine		
	None	Alkyne	Azide	None	Alkyne	Azide
PRMT1[a]		[20g, 20h]		[15, 58]	[58]	[58]
PRMT3		[20a]				
Dim-5		[20e, 20h]				
GLP1 (EuHMT1) [a]	[20b]	[59]				
G9a (EuHMT2) [a]	[20b]	[20b, 21, 28, 62, 64]				
METTL21A		[20b, 20c, 20h, 21, 28-29, 62, 64]	[60]			
METTL10		[29a]				
Set7/9		[29a]				
PrmC		[20h, 49d]				
SUV39H2		[20h]				
		[28, 62]				

[a] Mutant enzyme with improved activity in the presence of a AdoMet analogue

Small-molecule transalkylation

Small-molecule or natural product methylation is ubiquitous across all branches of the tree of life. Many AdoMet dependent MTases targeting natural products (NP-MTases) exist and they can generally be categorized based on their preference for oxygen, nitrogen, sulfur or carbon as their nucleophilic substrate.[65] As such, NP-MTases participate in the modification of a large numbers of structurally diverse small organic molecules, affecting their bio-availability, activity and reactivity in processes ranging from metabolism to signaling and biosynthesis.[65b] Many NP-MTases could be shown to feature a Rossman-like fold structural motif that is often extended by additional domains. These additions to the core enzyme structure ultimately allow individual NP-MTases to display wide ranging substrate specificity. Because of their versatility, NP-MTases, used in combination with AdoMet or its artificial analogues are particularly appealing in biocatalysis and production of fine chemicals where they can help to unlock synthetic routes that would not be accessible using traditional methods. Excellent overviews of such efforts are provided elsewhere.[2b, 66]

3. Current and future applications

The ever expanding repertoire of doubly-activated cofactor analogues and suitable natural- or engineered MTases has provided us with a unique and versatile set of tools for labeling the three major biopolymers.[67] While their most immediate applications might be found in the study of biological mechanisms underlying epigenetic modification and signaling, it would be hard to overstate the potential uses of such a highly specific biomolecular toolkit in all areas of research where directed, site specific labeling is required. In the following paragraphs, a selection of these applications is presented.

3.1. Selective enrichment of biomolecules

Specific enrichment of genetic material is particularly important in the context of high throughput, high coverage sequencing efforts.[68] Indeed, even though next-generation sequencing technologies have made whole genome sequence information relatively easy to obtain, the sheer amount of data produced can

10.1002/anie.201608625

be a confounding factor. Moreover, diseases such as cancer or viral infections will result in sub populations of cells or even single cells with a distinct genetic makeup against the background of the entire population, so called sub clonal genetic heterogeneities.[68b] By labeling target DNA with e.g. a biotin-containing AdoMet analogue, it can subsequently be captured using e.g. streptavidin functionalized particles.[19b] Alternatively, biomolecules can be labeled using clickable AdoMet analogues and coupled to azide or alkyne beads. This method was recently demonstrated using lambda DNA labeled with M.*Taq*I and a cofactor featuring a terminal alkyne.[69] A different study showed capturing of DNA by labeling with either amine or azide moieties.[50] Subsequently, the molecules were coupled to the surface of particles via either a biotin NHS-ester adduct or a biotin dibenzocyclooctyne moiety.

The use of MTases in the detection of non-methylated genomic DNA has been demonstrated by the group of Klimašauskas.[50] The MTase M.*Sss*I was used to label DNA featuring regions of unmethylated CpG sites with a cofactor that could be used to biotinylate the DNA. This biotinylated DNA could subsequently be selectively captured for downstream analysis using micro-array screening or sequencing.[50]

Next to the identification of genomic regions featuring low methylation, MTase based labeling strategies could also be applied to the separation of DNA of interest from a background of other genomic material in the sequencing of the Neanderthal genome. Here, environmental DNA of mostly bacterial origin needed to be removed. To achieve this, the researchers used the fact that mammalian DNA is frequently methylated at CpG sites, which occurs far less frequently in bacterial DNA. Treatment with restriction enzymes that specifically target CpG sites, largely destroys the bacterial DNA, leaving the endogenous DNA intact.[70]

Capturing biomolecules is not just limited to DNA, but can potentially also be used to specifically enrich RNA or proteins. Indeed, it has been shown that proteins can be captured by coupling a biotin tag to proteins with click chemistry.[71] This approach has been used by the group of Luo for the specific targeted capturing of proteins using MTase-based labeling (Figure 3).[20a, 20b, 28, 64]

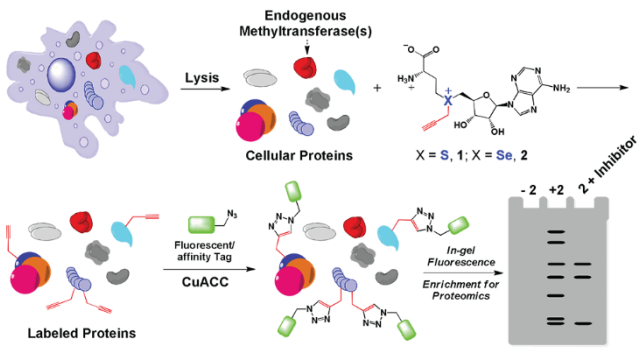


Figure 3. Schematic illustration of using AdoMet analogue as a reporter of protein methylation. The AdoMet analogue can be utilized by endogenous methyltransferases to label cellular proteins with an alkyne moiety. The modified proteins can then be coupled to a fluorescent tag using CuAAC for further characterization. From Islam *et al.*[28]

Specifically, an AdoMet analogue with alkyne functionality was used for targeted labeling of protein MTase targets. Then, biotin with an azide group was coupled to the proteins using click chemistry, and the proteins were subsequently captured using streptavidin beads. The presence of a cleavable azo-linker between the azide and biotin made it possible to separate the captured proteins from the beads after treatment with sodium dithionite ($Na_2S_2O_4$). Finally, the strategies presented here might also be a useful tool for the assembly of biomolecular structures[49b] and positioning of nanoparticles.[49a]

3.2 Genomic analysis

DNA mapping

The discovery of sequence specific restriction endonucleases (REases) enabled DNA sequence analysis long before the advent of single nucleotide sequencing.[72] When a DNA sample is digested by a known panel of REases, the ensuing characteristic fragment lengths can be analyzed using agarose gel electrophoresis, resulting in a 'DNA fingerprint' with applications in e.g. forensics.[73] Schwarz *et al*. further refined the technique by binding linearized DNA molecules to a surface prior to digestion, resulting in restriction or sequence 'maps' where additional information is contained in the relative location of the different fragments.[74] Site specific incorporation of fluorescent labels, rather than actual restriction of the DNA has further contributed to increase the information content of sequence maps.[75] Whereas nicking endonucleases and corresponding fluorescent nucleotide analogues were initially used here, they have since been supplanted by MTase-directed labeling, offering a number of advantages. Firstly, use of MTases keeps the DNA backbone intact, improving the stability of the DNA molecules and minimizing unwanted fragmentation. Furthermore, DNA MTases allow for a more direct labeling approach compared to nicking endonucleases, increasing the efficiency of transfer. Finally, through careful choice of the DNA MTases, high densities of labeling can be achieved. Indeed, a DNA MTase with a 4 base recognition sequence applied to a random DNA sequence would result in 1 label every 256 base pairs (4^4).

The past couple of years several approaches to DNA mapping using DNA MTases have been developed and excellent reviews on the subject exist.[76] These approaches can be roughly categorized in DNA mapping in nanofluidic channels and DNA mapping using super resolution microscopes (Figure 4).

MTase mapping in nanofluidic channels was first demonstrated by the group of Ebenstein[23] Here, phage T7 and phage Lambda DNA were labeled using the methyltransferase M.*Taq*I (recognition sequence: 5'-TCGA-3') and a synthetic AdoMet analogue carrying the fluorophore TAMRA. The labeled molecules were subsequently stretched in nanochannels after which intensity profiles were extracted using a fluorescence microscope. Because of the high density of fluorophores in combination with the use of diffraction limited microscopy, the exact location of labels could not be determined. Instead, the obtained intensity profiles were matched to several *in silico* generated intensity maps by cross correlation of both profiles. Thresholding the cross correlation scores allowed for the efficient

matching of bacteriophages to the correct sequence. More recently, the same group also demonstrated the ability to perform genomic mapping with sub-diffraction limit resolution in silicon nanochannels, greatly enhancing the information density.[77]

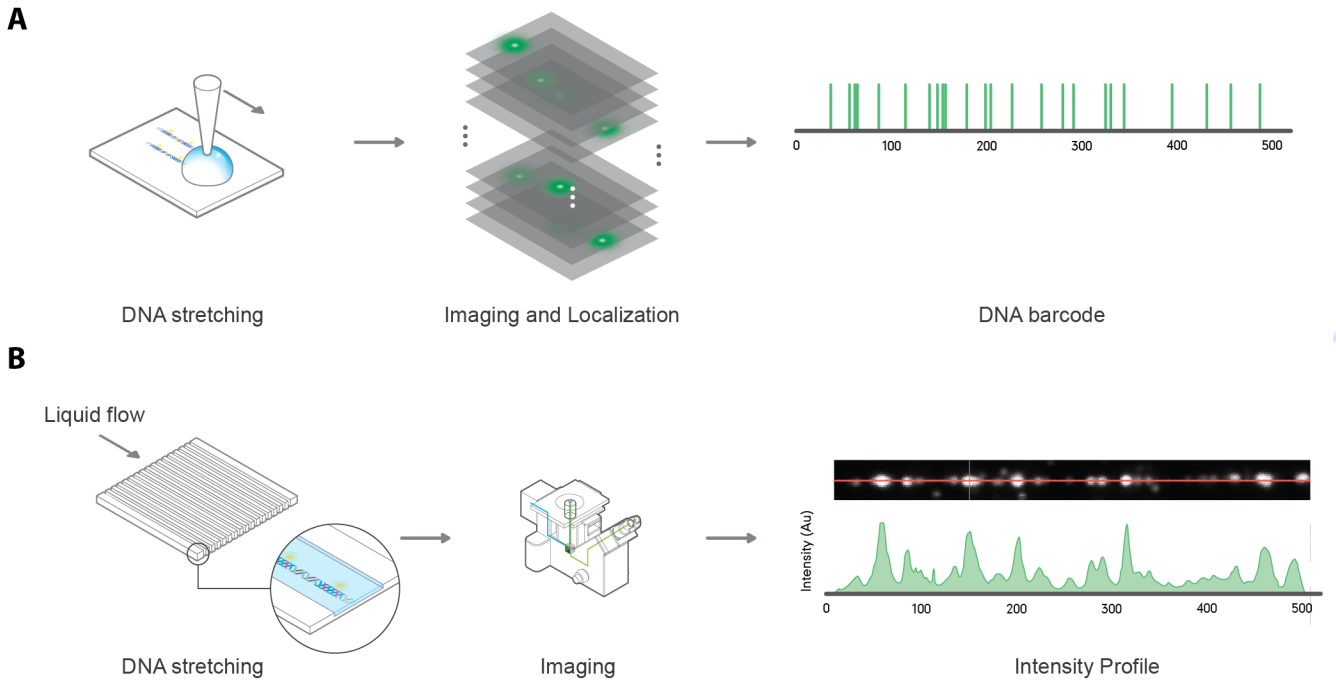


Figure 4: Illustration of the two main methods for reading the distribution of fluorophores on DNA molecules. A. For high resolution localization of fluorophores on the DNA, the DNA needs to be rigidly attached to the surface. B. For rapid imaging of DNA, the DNA molecules can be flown through nanochannels where they are elongated and subsequently, the intensity pattern can be recorded.

Due to the high density of labels, the work in our lab has focused on extracting high resolution localization information from the DNA molecules.[19c, 20f] To be able to extract high resolution localization information from the DNA it is important that the molecules are fixed on the surface. One method for stretching DNA molecules on a surface is based on flow stretching and attaching the stretched DNA on a poly-L-lysine surface. In one such example bacteriophage T7 (40 kbp) labeled with M.*Bse*CI (recognition sequence: 5'-ATCGAT-3') and a biotin containing aziridine analogue was coupled to streptavadin-coated quantum dots. The DNA molecules were subsequently stretched on the surface using a capillary flow.[49c] Due to the sparse labeling, the labels could be easily localized. However, the localization suggests there is a large variation in the positioning due to deposition inhomogeneity. An alternative method for stretching DNA is based on DNA combing, a method developed by Bensimon *et. al.* in the 90s.[78] With this approach, DNA transalkylation using AdoMet analogues carrying either terminal amine or alkyne groups was directed with the M.*Hha*I and M.*Taq*I (recognition sequences 5'-TCGA-3' and 5'-GCGC-3') methyltransferases. This gave a high density of modified sites which were subsequently labelled using the NHS-ester- or azide-derivatives of the Atto647N dye. Following deposition, high resolution were created based on stochastic photobleaching and localization of individual emitters.[76a] This resulted in an approximate resolution of 42nm (approximately 80 base pairs).[19c] Both these methods suggest promising new technique for extracting sequence information from DNA molecules by studying the positon of MTase-directed labels on DNA. One such example is the typing of bacteriophage molecules based on the barcode embedded in the molecule. Another area could be the detection of copy number variation, repeats of large genomic elements in the genomes which are hard to find by sequencing.[76a] Finally, the MTase-based labeling approach can easily be combined with other genomic analysis approaches such as fluorescence in situ hybridization (FISH).

DNA localization and spatially resolved transcriptomics

The attachment of reporter groups to biomolecules allows for their localization *in situ*. This has enabled researchers to put detailed functional analysis of biological processes in a spatial context, allowing biological function, i.e. molecular genetics and biochemistry, to be correlated with information on biological structure obtained from e.g. embryology and histology.[79]

Fluorescent labeling of nucleic acids such as DNA is one of the easiest routes to their localization in cells and many methods in fact exist.[80] However, methyltransferase directed transfer of

10.1002/anie.201608625

fluorescent groups offers the particular advantage that it allows for controlled targeting of the label and its covalent attachment.
In an example from the group of Weinhold, the covalent coupling of Cy3 dyes using an N-adenosylaziridine cofactor to pUC19 and pBR322 plasmids was used.[51] In this study, the labeled plasmids are successfully transfected in CHO-k1 cells. Of these transfected cells, 25% of the cells showed a high Cy3 fluorescence intensity in the nucleus, despite the absence of nuclear import sequence on the plasmids.[51]
It can be envisioned that MTase based labeling strategies could equally contribute to facilitate the *in situ* study of and large scale regulatory networks or efforts in highly multiplexed gene transcription profiling such as recently demonstrated by the team of Zhuang (Figure 5).[81] Here, the authors used an elaborate library of hybridization probes, fluorescently labeled and designed to target specific RNA species. By designing multiple, differently labeled probes, and multiple rounds of hybridization the authors were able to impart a unique color encoding for tens to even hundreds of individual RNA species *in situ* (Figure 5).

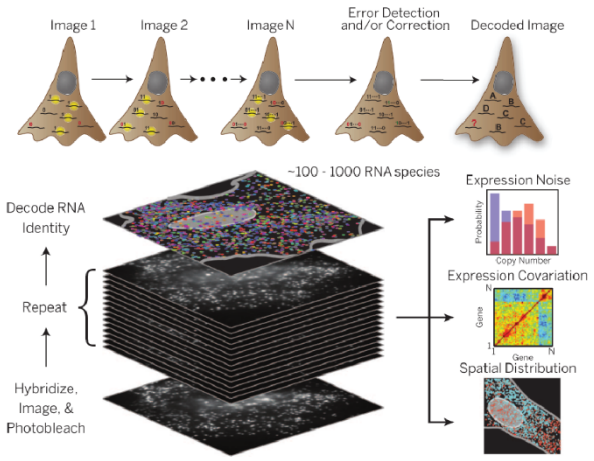


Figure 5. In the highly multiplexed RNA profiling method of the Zhuang group[81] hundreds of individual mRNA can correctly be identified, and localized in situ.

Although this example constitutes a great technical and scientific feat, the requirement to design and subsequently use hybridization probes can prove challenging and indeed, the authors had to apply significant error correction concepts borrowed from computer data encoding to their design of the probes. Furthermore, although this study allowed for simultaneous tracking of large numbers of biomolecular species, it was still only limited to RNA. Therefore, the possibility of achieving of highly directed MTase mediated labeling of all major classes of biomolecules opens up enticing new prospects to apply massively parallel observation of all species in an effort to truly unravel complex, spatially organized regulatory networks to elucidate cell-to-cell variations in the context of whole tissues.[79]

3.3 Epigenetic analysis

Epigenetic regulation is a collective term used to denote the entire spectrum of processes that modulate gene activity in an organism without actually altering the genetic sequence itself. Cytosine methylation (5mC) and stepwise conversion of 5mC to hydroxymethyl- (5hmC), formyl- (5fC) and finally carboxyl (5caC) cytosine in TET enzyme mediated demethylation are common in mammalian cells.[82] These modifications play important roles in embryonic development, stem-cell differentiation, genomic imprinting, neuronal function and cancer.[83] Additionally, histones and transcription factors can also be modified through methylation acylation or phosphorylation.[84] Because of their role in methylation, MTases, together with suitable AdoMet analogues are the ideal tools to study these diverse and transient phenomena.

DNA methylation

Bisulfite conversion is a commonly used method for the detection of 5mC as well as the other cytosine modifications. It relies on chemical modification of target species followed by PCR or sequencing based quantification.[83, 85] Unfortunately, the method suffers from the relatively low sensitivity, particularly for low abundance modifications as well as the relatively high error rates.[83] Furthermore, the method only provides ensemble averaged information whereas the stochastic nature of DNA methylation calls for approaches with single cell resolution.[86] Here, single molecule detection methods can offer a solution. In one example, the methylation status of single DNA molecules was probed using restriction enzymes, whose activity is blocked by the methylated target base. Combining this with optical mapping results in a methylation map of the DNA.[87] More recently, methyl-CpG binding proteins were used on DNA stretched in nanochannels.[88] Allowing for the direct detection of the methylation status of single molecules. A more direct method was used for the detection of hydroxymethyl cytosine. Here, the enzyme T4 β-glucosyltransferase was used to attach glucose molecules with a reactive azide moiety to the hydroxymethyl group, which could subsequently be coupled with fluorescent dyes.[89] For a more complete overview of the field the reader is referred to several excellent reviews.[76]
Like restriction enzymes, a methylated base also blocks the activity of methyltransferases. This characteristic was recently explored in a recent paper that showed profiling of the DNA unmethylome.[50] Unmethylated sites were labeled with the CpG MTase, M.*Sss*I, and an AdoMet analogue containing either an azide or amine functionality. The labeled DNA was then coupled to biotin (biotin-azide or biotin-NHS-ester) and extracted using streptavidin coated microbeads. After purification the DNA was analyzed using microarrays (Figure 6). This approach suggests a relatively straightforward way of enriching and subsequent sequencing of non-methylated DNA. Because of the sensitivity of the MTase-based method, very low amounts of DNA (100-300 ng) were needed to complete the analysis. Furthermore, the method was particularly sensitive to regions in the genome with low CpG presence, something that is difficult for other enrichment

10.1002/anie.201608625

techniques such as MeDIP (Methylated DNA immune-precipitation) and MBD (Methyl-CpG-binding domain capture).[50] MTase-based labeling of methylated DNA could easily be combined with single molecule detection in nanochannels.[76b] Since MTase-based labeling can be targeted towards either adenine or cytosine bases, the method could be combined with optical mapping in order to localize the regions of (un)methylation. Epigenome mapping could then be a dual color approach, one color for the map of the DNA and one color to create a map of the (un)methyl map.

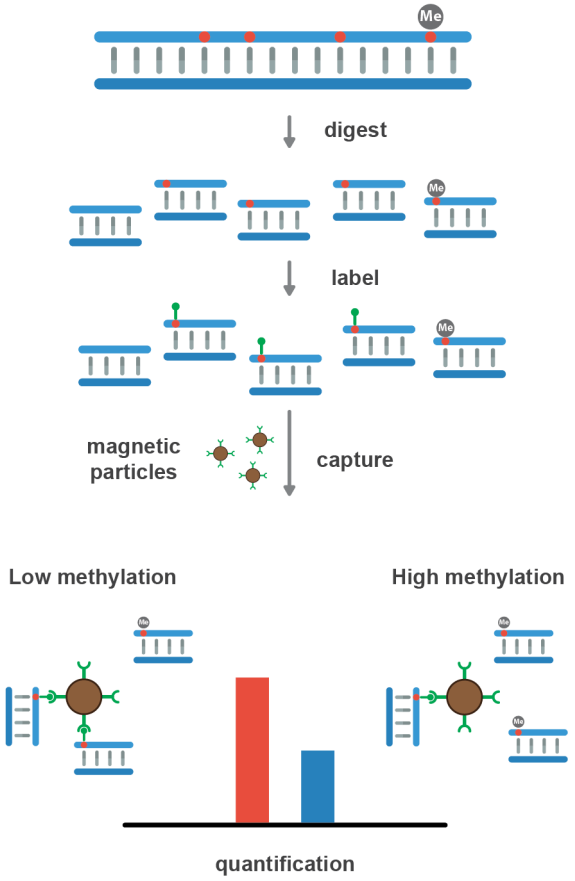


Figure 6. General principle for capture based unmethylome detection, illustrated with a quantitative PCR assay

Protein methylation

The target substrates of protein methyltransferases have been extensively studied by the group of Luo through the development of an assay termed Bio-orthogonal Profiling of Protein Methylation (BPPM). Cells are transfected with a mutated protein MTase showing reduced preference for natural AdoMet compared to the analogue. Upon lysis, the mixture is incubated with a doubly-activated AdoMet analogue with an azide linker. Target proteins are then coupled to a clickable dye such as e.g. dibenzylcyclooctyne coupled dyes and identified by mass spectrometry (MS) analysis or SDS-page (Figure 3).[21, 60-61] The approach was used for detecting the substrates of human protein MTase PRMT3, a protein MTase that preferably resides within the cytoplasm.[20a] Interestingly, while the protein MTase localizes in the cytoplasm, 23% of the modifications were found in the nucleus, suggesting a broader role for PRMT3.

More recently, the team set out to study methylation status of histones in live cells by hijacking the enzymatic synthesis of AdoMet to create AdoMet analogues instead. A modified version of methionine carrying a clickable analogue (with a terminal alkyne) was transported into the cells and, together with ATP, processed by an engineered MAT to form the modified cofactor. This modified cofactor was then used by a mutant of G9a which successfully transferred the alkyne group to the histones. The transfer of alkyne group to histones was subsequently confirmed using LC/MS.[64] The authors subsequently used this method to attach clickable biotin to the labeled chromatin. This allowed them to enrich the labeled chromatin using streptavidin coated beads. After cleavage of the linker, the DNA attached to the histones could be purified and sent for sequencing.[64] Indeed, qPCR analysis confirmed the presence of several genes that resides on the targeted histones. Because of the structural similarity in different MTases, this method is expected to be applicable across a wide array of protein MTases with varying targets. For more details on the detection and analysis of protein methylation the reader is referred to one of several reviews on the topic.[90]

Conclusions and future outlook

Since the conception of labeling biomolecules using MTases, much work has been done. The method offers a simple solution for placing reporter molecules specifically and covalently on target biomolecules. So far the method has been convincingly shown with over 20 MTases, consisting of 9 DNA MTases, both adenine and cytosine DNA MTases, 4 RNA MTases and 11 protein MTases. Several mutated versions of DNA or protein MTases have been developed with improved activity for the various AdoMet analogues, and in some cases a preference of the MTase for the AdoMet analogue as opposed to the natural cofactor AdoMet. This preference of the MTase allows the method to be used in *in vivo* systems where the natural cofactor AdoMet is still present in the sample mixture.

Two major groups of AdoMet analogues are currently available for MTase-based labeling. The aziridine based cofactors are synthesized chemically and can be used in MTase directed labeling reactions in which the whole compound is transferred to the substrate. The largest group of AdoMet analogues are the double-activated AdoMet analogues in which the transferrable methyl group of AdoMet is replaced by an unsaturated alkyl group that are used in mTag labeling of biomolecules. These analogues are typically synthesized from its precursor AdoHcy. Interestingly, these AdoMet analogue can sometimes also be synthesized by hijacking the same enzymes that create the natural cofactor AdoMet in cells, greatly simplifying the chemical route.

This approach for targeted labeling of biomolecules can subsequently be used for all sorts of applications. The attachment of biotin or other chemical groups makes it possible to specifically capture low quantities of DNA, mRNA or proteins. Furthermore,

the method offers a way to covalently attach organic fluorescent molecules to target molecules, which makes it useful for detection using microscopy methods. Finally, the targeted labeling of DNA offers a way to analyze sequence information beyond the standard sequence reads, either by imaging far longer DNA molecules or by focusing on the epigenetic code embedded in the DNA molecule.

The development of the methyltransferases as tools for biotechnology is a story that showcases the impact basic chemistry can have in this field. The emerging applications we have discussed highlight the potential of the MTase-directed reactions as a general strategy for targeting and labelling specific sites complex biological samples. There remains significant work to do to bring these tools to the non-specialist laboratory but the promise we have outlined is a strong driver for future development.

Acknowledgements

J.H. would like to acknowledge the European Research Council under the European Union's Seventh Framework Program (FP7/2007-2013), ERC Grant Agreement no. 291593 FLUOROCODE, Horizon 2020 Grant Agreement no. 686271 Ad gut. The Flemish government is acknowledged for the long term structural funding 'Methusalem' grant 1404 CASAS^2, Fonds voor Wetenschapplijk Onderzoek Vlaanderen (FWO) for grants G0484.12, G0683.15 and research infrastructure grant GOH6316N. J.H. and K.P.F.J. would like to thank Industrieel Onderzoeksfonds (IOF) for grant C32/16/005. K.P.F.J. would additionally like to thank FWO Flanders for post-doctoral scholarships 12C2814N and 12C2817N. RKN and JH gratefully acknowledge the European Union's Horizon 2020 research and innovation program under grant agreement No 634890 (BeyondSeq). RKN gratefully acknowledges the Engineering and Physical Sciences Research Council (EPSRC) for grant EP/N020901/1.

Conflict of interest statement: J.H., R.K.N. and V.L. have founded a company which may market applications of artificial AdoMet analogues for biofunctionalization.

Keywords: S-Adenosyl Methionine • Methyltransferase • DNA • DNA mapping • DNA functionalization

[1] A. B. Novikoff, *Science* **1945**, *101*, 209-215.

[2] a. H. L. Schubert, R. M. Blumenthal, X. Cheng, *Trends in Biochemical Sciences* **2003**, *28*, 329-335; b. A.-W. Struck, M. L. Thompson, L. S. Wong, J. Micklefield, *ChemBioChem* **2012**, *13*, 2642-2655.

[3] X. Cheng, R. J. Roberts, *Nucleic Acids Research* **2001**, *29*, 3784-3795.

[4] G. Lukinavičius, M. Tomkuvienė, V. Masevičius, S. Klimašauskas, *ACS Chemical Biology* **2013**, *8*, 1134-1139.

[5] a. X. Zhang, T. C. Bruice, *Proceedings of the National Academy of Sciences* **2006**, *103*, 6148-6153; b. D. V. Santi, C. E. Garrett, P. J. Barr, *Cell* **1983**, *33*, 9-10; c. M. O'Gara, S. Klimašauskas, R. J. Roberts, X. Cheng, *Journal of Molecular Biology* **1996**, *261*, 634-645.

[6] X. Cheng, *Annual Review of Biophysics and Biomolecular Structure* **1995**, *24*, 293-318.

[7] J. Zhang, Y. G. Zheng, *ACS Chemical Biology* **2015**, *11*, 583-597.

[8] a. G. L. Cantoni, *Journal of the American Chemical Society* **1952**, *74*, 2942-2943; b. A. Gross, S. Geresh, G. M. Whitesides, *Applied Biochemistry and Biotechnology* **1983**, *8*, 415-422.

[9] a. M. Thomsen, S. B. Vogensen, J. Buchardt, M. D. Burkart, R. P. Clausen, *Organic & Biomolecular Chemistry* **2013**, *11*, 7606-7610; b. S. Singh, J. Zhang, T. D. Huber, M. Sunkara, K. Hurley, R. D. Goff, G. Wang, W. Zhang, C. Liu, J. Rohr, S. G. Van Lanen, A. J. Morris, J. S. Thorson, *Angewandte Chemie International Edition* **2014**, *53*, 3965-3969.

[10] a. M. Pignot, C. Siethoff, M. Linscheid, E. Weinhold, *Angewandte Chemie International Edition* **1998**, *37*, 2888-2891; b. G. Pljevaljčić, F. Schmidt, E. Weinhold, *ChemBioChem* **2004**, *5*, 265-269.

[11] G. Pljevaljcic, M. Pignot, E. Weinhold, *Journal of the American Chemical Society* **2003**, *125*, 3486-3492.

[12] a. L. R. Comstock, S. R. Rajski, *Journal of the American Chemical Society* **2005**, *127*, 14136-14137; b. L. R. Comstock, S. R. Rajski, *Nucleic Acids Research* **2005**, *33*, 1644-1652.

[13] R. L. Weller, S. R. Rajski, *Organic Letters* **2005**, *7*, 2141-2144.

[14] a. R. L. Weller, S. R. Rajski, *ChemBioChem* **2006**, *7*, 243-245; b. Y. Du, C. E. Hendrick, K. S. Frye, L. R. Comstock, *ChemBioChem* **2012**, *13*, 2225-2233; c. C. Zhang, R. L. Weller, J. S. Thorson, S. R. Rajski, *Journal of the American Chemical Society* **2006**, *128*, 2760-2761.

[15] T. Osborne, R. L. Weller Roska, S. R. Rajski, P. R. Thompson, *Journal of the American Chemical Society* **2008**, *130*, 4574-4575.

[16] D. K. Ho, J. C. Wu, D. V. Santi, H. G. Floss, *Archives of Biochemistry and Biophysics* **1991**, *284*, 264-269.

[17] a. C. Dalhoff, G. Lukinavicius, S. Klimasauskas, E. Weinhold, *Nature Protocols* **2006**, *1*, 1879-1886; b. C. Dalhoff, G. Lukinavicius, S. Klimasauskas, E. Weinhold, *Nature Chemical Biology* **2006**, *2*, 31-32.

[18] G. Lukinavičius, V. Lapienė, Z. Staševskij, C. Dalhoff, E. Weinhold, S. Klimašauskas, *Journal of the American Chemical Society* **2007**, *129*, 2758-2759.

[19] a. M. Tomkuvienė, B. Clouet-d'Orval, I. Černiauskas, E. Weinhold, S. Klimašauskas, *Nucleic Acids Research* **2012**, *40*, 6765-6773; b. A. Plotnikova, A. Osipenko, V. Masevičius, G. Vilkaitis, S. Klimašauskas, *Journal of the American Chemical Society* **2014**, *136*, 13550-13553; c. R. K. Neely, P. Dedecker, J.-i. Hotta, G. Urbanaviciute, S. Klimasauskas, J. Hofkens, *Chemical Science* **2010**, *1*, 453-460; d. G. Lukinavičius, A. Lapinaitė, G. Urbanavičiūtė, R. Gerasimaitė, S. Klimašauskas, *Nucleic Acids Research* **2012**.

[20] a. H. Guo, R. Wang, W. Zheng, Y. Chen, G. Blum, H. Deng, M. Luo, *ACS Chemical Biology* **2014**, *9*, 476-484; b. K. Islam, Y. Chen, H. Wu, I. R. Bothwell, G. J. Blum, H. Zeng, A. Dong, W. Zheng, J. Min, H. Deng, M. Luo, *Proceedings of the National Academy of Sciences* **2013**, *110*, 16778-16783; c. K. Islam, W. Zheng, H. Yu, H. Deng, M. Luo, *ACS Chemical Biology* **2011**, *6*, 679-684; d. Y. Motorin, J. Burhenne, R. Teimer, K. Koynov, S. Willnow, E. Weinhold, M. Helm, *Nucleic Acids Research* **2011**, *39*, 1943-1952; e. W. Peters, S. Willnow, M. Duisken, H. Kleine, T. Macherey, K. E. Duncan, D. W. Litchfield, B. Lüscher, E. Weinhold, *Angewandte Chemie International Edition* **2010**, *49*, 5170-5173; f. C. Vranken, J. Deen, L. Dirix, T. Stakenborg, W. Dehaen, V. Leen, J. Hofkens, R. K. Neely, *Nucleic Acids Research* **2014**, *42*, e50; g. R. Wang, W. Zheng, H. Yu, H. Deng, M. Luo, *Journal of the*

American Chemical Society **2011**, *133*, 7648-7651; h. S. Willnow, M. Martin, B. Lüscher, E. Weinhold, *ChemBioChem* **2012**, *13*, 1167-1173.

[21] K. Islam, I. Bothwell, Y. Chen, C. Sengelaub, R. Wang, H. Deng, M. Luo, *Journal of the American Chemical Society* **2012**, *134*, 5909-5915.

[22] B. W. K. Lee, H. G. Sun, T. Zang, B. J. Kim, J. F. Alfaro, Z. S. Zhou, *Journal of the American Chemical Society* **2010**, *132*, 3642-3643.

[23] A. Grunwald, M. Dahan, A. Giesbertz, A. Nilsson, L. K. Nyberg, E. Weinhold, T. Ambjörnsson, F. Westerlund, Y. Ebenstein, *Nucleic Acids Research* **2015**, *43*, e117.

[24] D. F. Iwig, S. J. Booker, *Biochemistry* **2004**, *43*, 13496-13509.

[25] J. L. Hoffman, *Biochemistry* **1986**, *25*, 4444-4449.

[26] a. T. D. Huber, F. Wang, S. Singh, B. R. Johnson, J. Zhang, M. Sunkara, S. G. Van Lanen, A. J. Morris, G. N. Phillips, J. S. Thorson, *ACS Chemical Biology* **2016**; b. T. D. Huber, B. R. Johnson, J. Zhang, J. S. Thorson, *Current Opinion in Biotechnology* **2016**, *42*, 189-197.

[27] D. F. Iwig, A. T. Grippe, T. A. McIntyre, S. J. Booker, *Biochemistry* **2004**, *43*, 13510-13524.

[28] I. R. Bothwell, K. Islam, Y. Chen, W. Zheng, G. Blum, H. Deng, M. Luo, *Journal of the American Chemical Society* **2012**, *134*, 14905-14912.

[29] a. T. Shimazu, J. Barjau, Y. Sohtome, M. Sodeoka, Y. Shinkai, *PLoS ONE* **2014**, *9*, e105394; b. J. M. Winter, G. Chiou, I. R. Bothwell, W. Xu, N. K. Garg, M. Luo, Y. Tang, *Organic Letters* **2013**, *15*, 3774-3777.

[30] I. R. Bothwell, M. Luo, *Organic Letters* **2014**, *16*, 3056-3059.

[31] W. A. Smit, M. Z. Krimer, E. A. Vorob'eva, *Tetrahedron Letters* **1975**, *16*, 2451-2454.

[32] M. Ramadan, N. K. Bremner-Hay, S. A. Carlson, L. R. Comstock, *Tetrahedron* **2014**, *70*, 5291-5297.

[33] A. P. Townsend, S. Roth, H. E. L. Williams, E. Stylianou, N. R. Thomas, *Organic Letters* **2009**, *11*, 2976-2979.

[34] V. Masevičius, M. Nainytė, S. Klimašauskas, in *Current Protocols in Nucleic Acid Chemistry*, John Wiley & Sons, Inc., **2001**.

[35] G. de la Haba, G. A. Jamieson, S. H. Mudd, H. H. Richards, *Journal of the American Chemical Society* **1959**, *81*, 3975-3980.

[36] a. G. D. Markham, D. W. Parkin, F. Mentch, V. L. Schramm, *Journal of Biological Chemistry* **1987**, *262*, 5609-5615; b. G. L. Cantoni, *Journal of Biological Chemistry* **1953**, *204*, 403-416.

[37] a. G. L. Gilliland, G. D. Markham, D. R. Davies, *Journal of Biological Chemistry* **1983**, *258*, 6963-6964; b. G. D. Markham, *Journal of Biological Chemistry* **1981**, *256*, 1903-1909; c. G. D. Markham, *Journal of Biological Chemistry* **1986**, *261*, 1507-1509; d. G. D. Markham, E. W. Hafner, C. W. Tabor, H. Tabor, *Journal of Biological Chemistry* **1980**, *255*, 9082-9092; e. G. D. Markham, T. S. Leyh, *Journal of the American Chemical Society* **1987**, *109*, 599-600; f. C. Zhang, G. D. Markham, R. LoBrutto, *Biochemistry* **1993**, *32*, 9866-9873; g. R. J. Parry, A. Minta, *Journal of the American Chemical Society* **1982**, *104*, 871-872.

[38] J. M. Lipson, M. Thomsen, B. S. Moore, R. P. Clausen, J. J. La Clair, M. D. Burkart, *ChemBioChem* **2013**, *14*, 950-953.

[39] A. S. Eustaquio, F. Pojer, J. P. Noel, B. S. Moore, *Nature Chemical Biology* **2008**, *4*, 69-74.

[40] C. Vranken, A. Fin, P. Tufar, J. Hofkens, M. D. Burkart, Y. Tor, *Organic & Biomolecular Chemistry* **2016**, *14*, 6189-6192.

[41] S. T. Rao, M. G. Rossmann, *Journal of Molecular Biology* **1973**, *76*, 241-256.

[42] J. L. Martin, F. M. McMillan, *Current Opinion in Structural Biology* **2002**, *12*, 783-793.

[43] R. Gerasimaitė, G. Vilkaitis, S. Klimašauskas, *Nucleic Acids Research* **2009**, *37*, 7332-7341.

[44] Q. Lin, F. Jiang, P. G. Schultz, N. S. Gray, *Journal of the American Chemical Society* **2001**, *123*, 11608-11613.

[45] M Meselson, a. R Yuan, J. Heywood, *Annual Review of Biochemistry* **1972**, *41*, 447-466.

[46] M. G. Goll, T. H. Bestor, *Annual Review of Biochemistry* **2005**, *74*, 481-514.

[47] a. S. Klimasauskas, S. Kumar, R. J. Roberts, X. Cheng, *Cell* **1994**, *76*, 357-369; b. B. Holz, E. Weinhold, S. Klimasauskas, S. Serva, *Nucleic Acids Research* **1998**, *26*, 1076-1083.

[48] R. J. Roberts, X. Cheng, *Annual Review of Biochemistry* **1998**, *67*, 181-198.

[49] a. G. Braun, M. Diechtierow, S. Wilkinson, F. Schmidt, M. Hüben, E. Weinhold, N. O. Reich, *Bioconjugate Chemistry* **2008**, *19*, 476-479; b. S. Wilkinson, M. Diechtierow, R. A. Estabrook, F. Schmidt, M. Hüben, E. Weinhold, N. O. Reich, *Bioconjugate Chemistry* **2008**, *19*, 470-475; c. S. Kim, A. Gottfried, R. R. Lin, T. Dertinger, A. S. Kim, S. Chung, R. A. Colyer, E. Weinhold, S. Weiss, Y. Ebenstein, *Angewandte Chemie International Edition* **2012**, *51*, 3578-3581; d. G. M. Hanz, B. Jung, A. Giesbertz, M. Juhasz, E. Weinhold, **2014**, e52014.

[50] E. Kriukienė, V. Labrie, T. Khare, G. Urbanavičiūtė, A. Lapinaitė, K. Koncevičius, D. Li, T. Wang, S. Pai, C. Ptak, J. Gordevičius, S.-C. Wang, A. Petronis, S. Klimašauskas, *Nature Communications* **2013**, *4*.

[51] F. H. G. Schmidt, M. Hüben, B. Gider, F. Renault, M.-P. Teulade-Fichou, E. Weinhold, *Bioorganic & Medicinal Chemistry* **2008**, *16*, 40-48.

[52] F. Kunkel, R. Lurz, E. Weinhold, *Molecules* **2015**, *20*, 19723.

[53] a. M. T. Bedford, S. G. Clarke, *Molecular Cell* **2009**, *33*, 1-13; b. C. Martin, Y. Zhang, *Nature Reviews Molecular Cell Biology* **2005**, *6*, 838-849.

[54] D. D. Le, D. G. Fujimori, *Current Opinion in Chemical Biology* **2012**, *16*, 507-515.

[55] a. S. L. Berger, *Nature* **2007**, *447*, 407-412; b. T. Kouzarides, *Cell* **2007**, *128*, 802.e801-802.e802.

[56] Y.-H. Lee, M. R. Stallcup, *Molecular Endocrinology* **2009**, *23*, 425-433.

[57] R. A. Copeland, M. E. Solomon, V. M. Richon, *Nature Reviews Drug Discovery* **2009**, *8*, 724-732.

[58] S. J. Hymbaugh Bergman, L. R. Comstock, *Bioorganic & Medicinal Chemistry* **2015**, *23*, 5050-5055.

[59] O. Binda, M. Boyce, J. S. Rush, K. K. Palaniappan, C. R. Bertozzi, O. Gozani, *ChemBioChem* **2011**, *12*, 330-334.

[60] G. Blum, K. Islam, M. Luo, *Current protocols in chemical biology* **2013**, *5*, 45-66.

[61] G. Blum, I. R. Bothwell, K. Islam, M. Luo, in *Current Protocols in Chemical Biology*, John Wiley & Sons, Inc., **2009**.

[62] R. Wang, G. Ibanez, K. Islam, W. Zheng, G. Blum, C. Sengelaub, M. Luo, *Molecular BioSystems* **2011**, *7*, 2970-2981.

[63] M. Luo, *ACS Chemical Biology* **2012**, *7*, 443-463.

[64] R. Wang, K. Islam, Y. Liu, W. Zheng, H. Tang, N. Lailler, G. Blum, H. Deng, M. Luo, *Journal of the American Chemical Society* **2013**, *135*, 1048-1056.

[65] a. L. A. Wessjohann, J. Keim, B. Weigel, M. Dippe, *Current Opinion in Chemical Biology* **2013**, *17*, 229-235; b. D. K. Liscombe, G. V. Louie, J. P. Noel, *Nature Product Reports* **2012**, *29*, 1238-1250.

[66] H. Stecher, M. Tengg, B. J. Ueberbacher, P. Remler, H. Schwab, H. Griengl, M. Gruber-Khadjawi, *Angewandte Chemie International Edition* **2009**, *48*, 9546-9548.

[67] S. Klimašauskas, E. Weinhold, *Trends in Biotechnology* **2007**, *25*, 99-104.

[68] a. L. Mamanova, A. J. Coffey, C. E. Scott, I. Kozarewa, E. H. Turner, A. Kumar, E. Howard, J. Shendure, D. J. Turner, *Nature Methods* **2010**, *7*, 111-118; b. M. W. Schmitt, E. J. Fox, M. J. Prindle, K. S. Reid-Bayliss, L. D. True, J. P. Radich, L. A. Loeb, *Nature Methods* **2015**, *12*, 423-425.

[69] A. B. Artyukhin, Y.-H. Woo, *Analytical Biochemistry* **2012**, *425*, 169-174.

[70] R. E. Green, J. Krause, A. W. Briggs, T. Maricic, U. Stenzel, M. Kircher, N. Patterson, H. Li, W. Zhai, M. H.-Y. Fritz, N. F. Hansen, E. Y. Durand, A.-S. Malaspinas, J. D. Jensen, T. Marques-Bonet, C. Alkan, K. Prüfer, M. Meyer, H. A. Burbano, J. M. Good, R. Schultz, A. Aximu-Petri, A. Butthof, B. Höber, B. Höffner, M. Siegemund, A. Weihmann, C. Nusbaum, E. S. Lander, C. Russ, N. Novod, J. Affourtit, M. Egholm, C. Verna, P. Rudan, D. Brajkovic, Ž. Kucan, I. Gušic, V. B. Doronichev, L. V. Golovanova, C. Lalueza-Fox, M. de la Rasilla, J. Fortea, A. Rosas, R. W. Schmitz, P. L. F. Johnson, E. E. Eichler, D. Falush, E. Birney, J. C. Mullikin, M. Slatkin, R. Nielsen, J. Kelso, M. Lachmann, D. Reich, S. Pääbo, *Science* **2010**, *328*, 710-722.

[71] Y.-Y. Yang, M. Grammel, A. S. Raghavan, G. Charron, H. C. Hang, *Chemistry & Biology* **2010**, *17*, 1212-1222.

[72] a. P. Gill, A. J. Jeffreys, D. J. Werrett, *Nature* **1985**, *318*, 577-579; b. D. Botstein, R. L. White, M. Skolnick, R. W. Davis, *American Journal of Human Genetics* **1980**, *32*, 314-331; c. J. Sambrook, J. Williams, P. A. Sharp, T. Grodzicker, *Journal of Molecular Biology* **1975**, *97*, 369-390.

[73] W. A. M. Loenen, D. T. F. Dryden, E. A. Raleigh, G. G. Wilson, N. E. Murray, *Nucleic Acids Research* **2014**, *42*, 3-19.

[74] a. D. Schwartz, X. Li, L. Hernandez, S. Ramnarain, E. Huff, Y. Wang, *Science* **1993**, *262*, 110-114; b. W. Cai, H. Aburatani, V. P. Stanton, D. E. Housman, Y. K. Wang, D. C. Schwartz, *Proceedings of the National Academy of Sciences* **1995**, *92*, 5164-5168.

[75] M. Xiao, A. Phong, C. Ha, T.-F. Chan, D. Cai, L. Leung, E. Wan, A. L. Kistler, J. L. DeRisi, P. R. Selvin, P.-Y. Kwok, *Nucleic Acids Research* **2007**, *35*, e16-e16.

[76] a. R. K. Neely, J. Deen, J. Hofkens, *Biopolymers* **2011**, *95*, 298-311; b. M. Levy-Sakin, Y. Ebenstein, *Current Opinion in Biotechnology* **2013**, *24*, 690-698.

[77] J. Jeffet, A. Kobo, T. Su, A. Grunwald, O. Green, A. N. Nilsson, E. Eisenberg, T. Ambjornsson, F. Westerlund, E. Weinhold, D. Shabat, P. K. Purohit, Y. Ebenstein, *ACS Nano* **2016**.

[78] a. J. Deen, W. Sempels, R. De Dier, J. Vermant, P. Dedecker, J. Hofkens, R. K. Neely, *ACS Nano* **2015**, *9*, 809-816; b. A. Bensimon, A. Simon, A. Chiffaudel, V. Croquette, F. Heslot, D. Bensimon, *Science* **1994**, *265*, 2096-2098.

[79] N. Crosetto, M. Bienko, A. van Oudenaarden, *Nature Reviews Genetics* **2015**, *16*, 57-66.

[80] K. Rombouts, K. Braeckmans, K. Remaut, *Bioconjugate Chemistry* **2016**, *27*, 280-297.

[81] K. H. Chen, A. N. Boettiger, J. R. Moffitt, S. Wang, X. Zhuang, *Science* **2015**.

[82] a. W. A. Pastor, L. Aravind, A. Rao, *Nature Reviews Molecular Cell Biology* **2013**, *14*, 341-356; b. Z. D. Smith, A. Meissner, *Nature Reviews Genetics* **2013**, *14*, 204-220.

[83] H. Wu, Y. Zhang, *Nature Structural and Molecular Biology* **2015**, *22*, 656-661.

[84] A. P. Feinberg, B. Tycko, *Nature Reviews Cancer* **2004**, *4*, 143-153.

[85] J. G. Herman, J. R. Graff, S. Myöhänen, B. D. Nelkin, S. B. Baylin, *Proceedings of the National Academy of Sciences of the United States of America* **1996**, *93*, 9821-9826.

[86] a. M. Levy-Sakin, A. Grunwald, S. Kim, N. R. Gassman, A. Gottfried, J. Antelman, Y. Kim, S. O. Ho, R. Samuel, X. Michalet, R. R. Lin, T. Dertinger, A. S. Kim, S. Chung, R. A. Colyer, E. Weinhold, S. Weiss, Y. Ebenstein, *ACS Nano* **2014**, *8*, 14-26; b. C. A. Aguilar, H. G. Craighead, *Nature Nanotechnology* **2013**, *8*, 709-718.

[87] G. E. Ananiev, S. Goldstein, R. Runnheim, D. K. Forrest, S. Zhou, K. Potamousis, C. P. Churas, V. Bergendahl, J. A. Thomson, D. C. Schwartz, *BMC Molecular Biology* **2008**, *9*, 1-14.

[88] S. Fang Lim, A. Karpusenko, J. J. Sakon, J. A. Hook, T. A. Lamar, R. Riehn, *Biomicrofluidics* **2011**, *5*, 034106-034106-034108.

[89] a. Y. Michaeli, T. Shahal, D. Torchinsky, A. Grunwald, R. Hoch, Y. Ebenstein, *Chemical Communications* **2013**, *49*, 8599-8601; b. C.-X. Song, K. E. Szulwach, Y. Fu, Q. Dai, C. Yi, X. Li, Y. Li, C.-H. Chen, W. Zhang, X. Jian, J. Wang, L. Zhang, T. J. Looney, B. Zhang, L. A. Godley, L. M. Hicks, B. T. Lahn, P. Jin, C. He, *Nature Biotechnology* **2011**, *29*, 68-72.

[90] a. R. Wang, M. Luo, *Current Opinion in Chemical Biology* **2013**, *17*, 729-737; b. W. Fischle, D. Schwarzer, *ACS Chemical Biology* **2016**, *11*, 689-705.

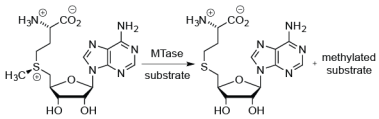
10.1002/anie.201608625

WILEY-VCH

Entry for the Table of Contents

REVIEW

Methyltransferases (MTases) compose a large family of enzymes which have the ability to methylate a diverse set of targets, ranging from the three major biopolymers DNA, RNA and protein to small molecules. In this review, various strategies for labelling and functionalizing biomolecules using AdoMet-dependent MTases and AdoMet analogues are discussed along with current and future applications.



Jochem Deen, Charlotte Vranken, Volker Leen, Robert K. Neely, Kris P. F. Janssen and Johan Hofkens**

Page No. – Page No.

Methyltransferase directed labeling of biomolecules and its applications