Published in final edited form as:

Curr Opin Biotechnol. 2019 February 1; 55: 114–123. doi:10.1016/j.copbio.2018.09.008.

Repurposing enzymatic transferase reactions for targeted labeling and analysis of DNA and RNA

Migl Tomkuvien , Milda Mickut , Giedrius Vilkaitis, and Saulius Klimašauskas*

Institute of Biotechnology, Life Sciences Center, Vilnius University, Saul tekio av. 7, Vilnius LT-10257, Lithuania

Abstract

Produced as linear biopolymers from four major types of building blocks, DNA and RNA are further furnished with a range of covalent modifications. Despite the impressive specificity of natural enzymes, the transferred groups are often poor reporters and not amenable to further derivatization. Therefore, strategies based on repurposing some of these enzymatic reactions to accept derivatized versions of the transferrable groups have been exploited. By far the most widely used are *S*-adenosylmethionine-dependent methyltransferases, which along with several other nucleic acids modifying enzymes offer a broad selection of tagging chemistries and molecular features on DNA and RNA that can be targeted *in vitro* and *in vivo*. Engineered enzymatic reactions have been implemented in validated DNA sequencing-based protocols for epigenome analysis. The utility of chemo-enzymatic labeling is further enhanced with recent advances in physical detection of individual reporter groups on DNA using super-resolution microscopy and nanopore sensing enabling single-molecule multiplex analysis of genetic and epigenetic marks in minute samples. Altogether, a number of new powerful techniques are currently in use or on the brink of real benchtop applications as research tools or next generation diagnostics.

Introduction

DNA and RNA serve key functions of storing, copying and transmitting the genetic information in all known organisms. Originally produced as linear biopolymers from four major types of building blocks, they are further processed by furnishing with a range of covalent modifications required for their functions. Among numerous enzymatic reactions involved in nucleic acid metabolism, methylation is the most simple and ubiquitous natural modification playing a variety of regulatory and structural roles in the cell (reviewed in [1–3]). Selective methylation of specific targets in DNA and RNA is performed by numerous dedicated methyltransferase enzymes (MTases), most of which use the cofactor *S*-adenosyl-L-methionine (AdoMet) as the source of a methyl group.

Although the biological modifications are deposited by corresponding enzymes in a highly specific manner, the transferred groups are usually poor reporters themselves and/or are not

*Corresponding Author, Saulius Klimašauskas, Corresponding Author's Institution, Vilnius University, saulius.klimasauskas@bti.vu.lt.

Declarations of interest

M.M, G.V and S.K are inventors on related patents and patent applications

amenable to further derivatization. To unlock the biotechnological power of these enzymes, strategies based on repurposing enzymatic reactions to accept expanded (derivatized) versions of the transferable groups have been exploited (Figure 1). The broadest spectrum of engineered reactions has been achieved with AdoMet-dependent MTases. The latest version of the technology, termed methyltransferase-directed Transfer of Activated Groups, or mTAG [4], is widely exploited for targeted covalent tagging of DNA and RNA with demonstrated applications in (epi)genome and (epi)transcriptome analysis (please refer to recent review articles for comprehensive reading [5–8]). Several other DNA and RNA transferase reactions have been similarly redesigned and employed to varied degrees. Here we discuss the latest developments in this rapidly expanding field with a focus on i) engineering of suitable enzyme-cofactor pairs *in vitro* and *in vivo*, ii) analysis of the tagged DNA and RNA using next generation sequencing and advanced optical techniques and iii) applications of the newly created tools in biochemical and biomedical research including genomic/epigenomic as well as transcriptomic/epitranscriptomic studies.

Engineering AdoMet-dependent transmethylation reactions

Although initially the development of AdoMet analogs for MTase-directed labeling was based on coupling reactions [9], the most recent approach relies on synthetic cofactors in which the transferable sulfonium-bound methyl group of AdoMet is replaced with an activated extended side chain (Figure 1A Top and Figure 2) [4,10–13]. Typically, linear linkers are used to carry terminal amino or azide and alkyne functional groups, which permit subsequent selective labeling via amino-NHS ester or azide-alkyne cycloaddition (AAC) reactions, respectively [12,14]. In addition, groups carrying an alkene [15–18], a norbornene moiety (for bioorthogonal tetrazine ligation [13]), a photocrosslinker [19] and a photo-removable cage [20] have recently become available. An increasingly popular choice is one-step mTAG labeling with fluorophores or biotin [21–25].

To date, three routes for production of mTAG cofactors have been proposed (reviewed in [26]). Chemical synthesis is based on direct “recharging” of the reaction co-product AdoHcy with a desired side chain. This route affords multi-milligram amounts of cofactor analogs, usually as diastereomeric mixtures [12,27]. Cofactors with large reporter groups such as chromophores or biotin are often obtained by further appending the side chains of functionalized cofactors via suitable conjugation chemistries under mild conditions [24,25,28]. An alternative more recent strategy is based on the natural route of AdoMet production from methionine and ATP catalyzed by AdoMet synthetase (also known as methionine adenosyltransferase, MAT). Enzyme engineering and synthetic efforts in several laboratories led to an expanded repertoire of methionine analogs that can be converted to stereochemically pure AdoMet analogs, although the size of the transferable moiety is currently limited to 5–7 carbon units [29–31]. The MAT-directed conversion of respective methionine analogs bypasses the issues of cofactor shelf-life or cell impermeability inherent for most of the AdoMet analogs and is well suited for coupled one-pot production/tagging reactions *in vitro* [32]. Proof-of-principle demonstrations are also available for covalent *in vivo* tagging of proteins [30] and RNA [33]. Yet another approach exploits halogenases [34,35]; in particular, the SalI chlorinase naturally catalyzes chloride-dependent decomposition of AdoMet to methionine and 5'-chloroadenosine. In the reverse reaction, the

methionine analogs can be converted to extended cofactors *in vitro*, which can be *in situ* used for mTAG labeling [35]. *In vivo* application of the system may be limited due to its intolerance to chloride ions present in cellular environments.

With respect to mTAG catalysts, both DNA and RNA MTases show a wide range of plasticity in accepting chemically expanded cofactors. On one end of the spectrum are highly promiscuous enzymes that require no engineering for the transfer of even the bulkiest groups (for example, the TaqI DNA adenine-*N*6 MTase (M.TaqI) [13,20–23], the HEN1 2'-*O*-RNA MTases from plants and animals [24,25,36] and the 5'-cap specific Ecm1 RNA MTase [13,19,37]). Meanwhile, structure-guided steric engineering of the cofactor pocket in DNA cytosine-5 MTases is required to achieve desired transalkylation activity [38,39] (Figure 3). On the other hand, the most compact functionalized groups such as allyl or propargyl are accepted by a broad range of natural enzymes [18,33,40]. However, the suitability of such cofactors for *in vivo* studies may be limited by their poor MTase selectivity (simultaneous acceptance by multiple cellular wt MTases).

DNA modifying enzymes for genomic and epigenomic analysis

Applications of DNA MTases in genomic analysis

DNA MTases generally target a specific nucleobase residue within a defined 2–8 bp recognition sequence producing m⁶A, m⁴C or m⁵C. Enzymes with over 350 distinct recognition sequences (REBASE, <http://rebase.neb.com>) have been identified suggesting that a broad repertoire of DNA sequences can potentially be targeted. However, a rather small fraction of representative enzymes have been isolated and characterized to different degrees *in vitro* [1] and subsequently a handful of those (based on their target specificity, stability and catalytic performance) have been recruited for DNA tagging applications [7] (Figures 3A and 4A). Following numerous demonstrations of targeted labeling of DNA and RNA *in vitro* or *ex vivo*, the potential of mTAG labeling is further unlocked with advances in instrumentation and techniques enabling optical detection of individual fluorophore molecules in biological samples. The key idea here is to determine physical positions of fluorescently-labeled target sites along the DNA contour (in linearly stretched single molecules) and analyze the resulting pattern to reveal the distribution of the target sequences in the genome. Such optical mapping (OM) has been implemented as an automated process using a limited set of available sequence-specific nicking nucleases for DNA labeling (commercialized by Bionano Genomics Inc [41]). Introduction of MTase-directed labeling, which makes no nicks in DNA, can open a new page in OM analysis [42] by i) extending the size of genomic maps to 20-100 Mbp and ii) expanding the selection of potentially available target specificities [14,43]. At this point, OM of mTAG labeled DNA has progressed to demonstrate its potential applicability for virus or macrosatellite genotyping [21,44]. Similarly, after mTAG biotin labeling, DNA sequences can be monitored using nanopore sensing. Biotin-bound streptavidin gives rise to current blockade signals during DNA passage through a quartz nanopore, which allowed simultaneous detection of two short genomes in a mixture [23]. Although presently both approaches are at the stage of technical refinement, the MTase-aided single molecule genotyping technologies may offer new versatile tools for genome analysis on the submegabase scale.

Applications of DNA MTases in epigenomic research

Modifications of cytosine (largely to m⁵C and hmC) in CG dinucleotides act as key epigenetic signals affecting gene expression and cellular differentiation in higher eukaryotes. Typically, 65–80% of the total 28 million CG sites are methylated in human tissues with characteristic epigenetic profiles observed in each tissue, cell type and disease condition, including cancers. Since the activity of DNA cytosine-5 MTases is blocked when the target cytosine residue is already modified, a CG-specific DNA C5-MTase, M.SssI, was engineered for two-step biotin labeling of unmodified CG dinucleotides. Subsequent enrichment of the covalently tagged DNA fraction followed by next generation sequencing yielded so-called “unmethylome” (inverse of the canonical DNA methylome) [39]. This medium resolution approach has been subsequently advanced by attaching a DNA oligonucleotide (instead of biotin) in the second step of labeling. Remarkably, it was found that a covalently tethered oligonucleotide can promote non-homologous priming of a DNA polymerase at the tagged sites to directly produce adjoining regions for their sequencing and precise genomic mapping at single CpG resolution [45]. The method named Tethered Oligonucleotide-Primed sequencing (TOP-seq), was validated for whole-genome epigenetic profiling of human tissues and cancer cell lines.

Notably, the M.TaqI adenine MTase is also blocked by C methylation within its recognition sequence (TCGA), permitting DNA unmethylome analysis in a subset of genomic CG sites. Moreover, a combination of sequence-specific [46] and methylation-sensitive mTAG labeling and multicolor readout yields a hybrid genetic/epigenetic map of chromosome segments spanning hundreds of kilobases [47]. Similarly, Gilboa et al. [22] demonstrated a proof of principle of epigenetic analysis using nanopore sensing. Again, M.TaqI was used to attach fluorophores at unmethylated TCGA sequences, in a mono-chromic or bi-colour manner. Electrical and optical signals were then recorded simultaneously as individual labeled DNA molecules were passing through solid-state nanopores. The demonstrated ability to detect and quantify multiple colors/signals from single DNA molecules holds promise of a high throughput automated multiplex analysis of distinct (epi)genetic features in clinical samples in near future (see accompanying review by Heck et al.).

In addition to their proper activity, DNA C5-MTases, which use a covalent mechanism for nucleophilic activation of their target cytosine residues, can also promote reactions involving non-cofactor-like substrates. In particular, the C5-MTase-activated 5-hydroxymethylcytosine (hmC) residues in DNA can undergo derivatization with aliphatic thiols yielding corresponding 5-alkylthiomethyl derivatives (Figure 1A Bottom and 4A Bottom) [48,49]. M.SssI-directed tagging of hmC with cysteamine and subsequent amine-selective biotin labeling was exploited for selective covalent capture of 5-hydroxymethylated-CG sites [48] (commercialized as EpiJET 5hmC Enrichment Kit by Thermo Fisher Scientific) and found laboratory applications [50,51]. Similar M.HhaI-directed labeling has been used for a proof-of-principle demonstration of a photo-electrochemical biosensor for hmC quantitation [52]. In epigenomic research, hmC, the product of enzymatic oxidation of m⁵C, is now extensively recognized both as a part of demethylation pathway and as an independent epigenetic mark (reviewed in [3,53,54]).

Other DNA modifying enzymes for epigenomic analysis

Besides MTases, only a few other enzymatic reactions have undergone engineering for biotechnological labeling of DNA. The most broadly used is the hmC-specific *beta*-glucosyltransferase (BGT) from bacteriophage T4. Engineered BGT reactions rely on the uridine 5'-diphosphate glucose cofactor in which the 6-position of the naturally transferable D-glucose moiety is functionalized with an amine, ketone or azide group [55–57] (Figure 1B). The latter click-compatible derivative (UDP-6-N₃-Glc) proved the most popular – it is already available commercially and can also be prepared *in situ* from inexpensive precursor compounds [58].

Recent analytical applications, with hmC as a biomarker (Figure 4B), include a biotin-based enrichment and next generation sequencing of cell-free DNA for cancer diagnostics and monitoring [59,60] (see also accompanying review [61]). Besides the elaborate genomic analyses, nanopore sensing permitted quantitation of the biotin-derivatized hmC in the murine genome [62], whereas optical quantitation of fluorescently labeled genomic hmC in minute blood samples proved valuable for diagnostics of colorectal and blood cancer [63] (see accompanying review by Heck et al.). In addition, a one-pot combination of TET oxygenase and BGT activities enabled detection of proximal m⁵C and hmC modifications using single-molecule FRET [64]. Recently, single-molecule dual optical mapping of genetic and epigenetic features in large fragments of human DNA has been demonstrated using commercial nanochannel arrays [65].

RNA modifying enzymes for transcriptome and epitranscriptome analysis

Every living cell contains multiple types of RNA molecules. A big variety of chemical groups as well as different positions in an RNA molecule can be modified by methyltransferases resulting in over 70 distinct modifications [66]. To date, a handful of these enzymes have been adapted for targeted covalent tagging of RNA (Figures 3B and 4A).

mRNA cap methyltransferases

Covalent labeling of the 5'-end of messenger RNA could provide tools for studies of trafficking and localization of genetic transcripts in eukaryotic cell. A substantial progress in the mRNA cap modification has been made by exploiting two enzymes: the GlcTgs2 MTase, which methylates the N² position of the m⁷G cap [15] and the Ecm1 cap guanine-N7 methyltransferase [37]. Ecm1 and an engineered variant of GlcTgs2 were used to deposit extended side chains from a range of AdoMet analogs onto *in vitro* transcribed G or m⁷G capped RNAs, respectively, [13,15,17,19,32,37,67] as well as dinucleotide cap variants m⁷GpppA or GpppA in bacterial [15] or eukaryotic [13,15,17,32,67] cell lysates. The tagged substrates were further analyzed after subsequent conjugation to functional reporters [13,15–17,19,32,37,67,68]. Using the Ecm1-directed labeling, it was demonstrated that photo-reactive arylazide or diazirine groups can be used to detect specific RNA-protein interaction *in vitro* [19]. Combining the above tagging reactions in a one-pot cascade format, dual labeling of the 5'-cap with different reporters has been demonstrated [68].

3' terminal RNA methyltransferases

Another type of RNA methyltransferases, HEN1 enzymes, modify the 2'-OH position of 3'-terminal nucleosides regardless of their identity. In cells, HEN1 transferases are guided by additional proteins to methylate distinct classes of regulatory RNAs such as miRNA and siRNA duplexes in plants [69] or single stranded piRNA and siRNA molecules in animals [25]. As mentioned above, the broad cofactor tolerance by both types of HEN1 MTases permits robust targeted deposition of a wide spectrum of functional or reporter groups including biotin or Cy3 fluorophore [24,25,36]. In particular, the AtHEN1 MTase was shown to be instrumental for DNA-addressed tagging of miRNA and siRNA strands. This reaction proved tolerant to the presence of additional chemical moieties (such as aptamers or fluorophores) on the probing DNA strand [24] permitting further development of two-reporter analytical systems for sequence-specific miRNA analysis in biological samples. Meanwhile a truncated variant of DmHen1 showed efficient derivatization of a wide range of ssRNA substrates regardless of their length (20–80 nt) or the nature of the 3'-terminal base, even in the presence of endogenous AdoMet [25] suggesting its potential applicability for *in vivo* RNA profiling experiments [70].

Internal RNA methyltransferases

The majority of biological RNA modifications occur at internal nucleosides [66]. It was shown that the substrates of tRNA:m²G26-methyltransferase Trm1 and tRNA:m²G10-methyltransferase Trm11 can be specifically modified with terminal alkynyl group and further visualized after Cu⁺-catalyzed AAC or even applied for single-molecule fluorescence experiment [71]. Synthetically tunable sequence-specific 2'-O-tagging and click labeling was demonstrated using an archaeal box C/D ribonucleoprotein complex [40]. Besides the described general labeling tools, profiling of internal m⁶A modifications in mRNA has recently gained increased attention as this modification has been associated with regulation of mRNA metabolism and cell differentiation [72]. However, the N⁶-methyl groups are “invisible” to general transcriptomic techniques and dedicated approaches for the identification of residues targeted by the heterodimeric METTL3-METTL14 MTase have recently been proposed. In one instance, it was shown that the action of the human MTase in the presence of an engineered cofactor can produce N⁶-allyladenosine residues; subsequent treatment with aqueous iodine leads to a N¹,N⁶-cyclic derivative, which is revealed by the appearance of misincorporated bases in RNA sequencing data [18]. Independently, METTL3-METTL14 was applied to modify its target sites with propargyl groups, which were then used for click-biotinylation and immobilization of the RNA to streptavidin beads. The created steric barrier prompted high frequency termination of reverse transcription at the modified nucleosides enabling their mapping [33].

Other tRNA modifying enzymes for post-transcriptional RNA labeling

Two other tRNA modifying enzymes have been successfully applied for sequence-specific RNA labeling using modified co-substrate analogs (Figures 1C, D and 4C). In both cases, enzyme-directed covalent modification of a target sequence (prokaryotic tRNA or its fragment) genetically fused to an RNA of interest was exploited. The incorporated functional handles or direct reporters were used to detect the presence and location of the

RNA in cellular extracts or fixed cells [73,74]. A tRNA^{Ile2}-agmatidine synthetase (Tias), which catalyses the attachment of the agmatine (4-aminobutylguanidine) moiety to C34 of *A. fulgidus* tRNA^{Ile2}, was used to modify RNA substrates with agmatine analogs bearing azide or alkyne functional groups [73]. *E. coli* tRNA guanine transglycosylase (TGT) was shown to use a shorter 17-nt fragment of tRNA^{Tyr} for derivatization of RNA. TGT functions by replacing G34 with the preQ1 nucleobase, which can be appended with a fluorophore or biotin beforehand [74,75] or in a subsequent step [76]. Notably, the latter system can even be configured for spatiotemporal regulation of translation in transfected HeLa cells: modification of the inserted target sequences in the 5' UTR with bulky light-sensitive groups leads to stalled translation, which is readily reversed by light-induced uncaging of mRNA [77].

Conclusions

Repurposing enzymatic reactions for covalent derivatization of selected targets in DNA and RNA (Figure 1) is coming of age, counting presently over 20 examples. By far the most widely exploited are MTases, followed by BGT (Figure 4). The mTAG technique [4,9] offers a very broad selection of tagging chemistries (Figure 2) as well as molecular features on DNA and RNA (Figure 3) that can be targeted in a highly specific manner *in vitro*. *In vivo* tagging of biological methylation sites has been attempted in model systems and currently appears to suffer from technical limitations: the size restrictions of accepted methionine analogs, competition by endogenous methionine and AdoMet and the MTase selectivity; bypassing these hurdles will permit deciphering deeper layers of the cellular mechanisms underlying transcriptional and translational regulation in mammals. Both types of reactions (DNA MTases and BGT) have been implemented in validated sequencing-based protocols for epigenome analysis ready for routine laboratory and, potentially, clinical applications [39,45,59,60] (see also accompanying reviews [61] and Heck et al.). However, the utility of mTAG labeling is further increased with recent technological advances in physical detection of individual tagged reporter molecules on DNA and RNA. Optical (super-resolution microscopy) and/or electrochemical (nanopore sensing) positioning of covalently tagged reporters along the contour of DNA (currently at 50–200 bp resolution) offers single-molecule multiplexed readout permitting high throughput automated mapping of genetic and epigenetic features in minute biological samples. The techniques are at different stages of technical maturity, awaiting further refinement and clinical validation. Two other RNA modifying enzymes (Figures 1C and D) have been successfully adapted for *in vivo* labeling of RNA at short genetically fused affinity tags (Figure 4C). Altogether, a handful of developed powerful techniques have collected enough of proofs of principle and are currently on the brink of real benchtop applications in both fundamental research and next generation diagnostics.

Acknowledgements

The authors wish to acknowledge support from the European Research Council [grant ERC-2016-ADG 742654 to S.K.], the Lithuanian Research Council [grant SEN-17/2015 to G.V.] and the H2020 COST consortium CM1406 for facilitating constructive discussion platforms.

References

1. Jurkowska RZ, Jeltsch A. Mechanisms and Biological Roles of DNA Methyltransferases and DNA Methylation: From Past Achievements to Future Challenges. *Adv Exp Med Biol.* 2016; 945:1–17. [PubMed: 27826832]
2. Nachtergaele S, He C. The emerging biology of RNA post-transcriptional modifications. *RNA Biol.* 2017; 14:156–163. [PubMed: 27937535]
3. Carell T, Kurz MQ, Müller M, Rossa M, Spada F. Non-canonical Bases in the Genome: The Regulatory Information Layer in DNA. *Angew Chem Int Ed Engl.* 2018; 57:4296–4312.
4. Lukinavičius G, Lapien V, Staševskij Z, Dalhoff C, Weinhold E, Klimašauskas S. Targeted labeling of DNA by methyltransferase-directed transfer of activated groups (mTAG). *J Am Chem Soc.* 2007; 129:2758–2759. [PubMed: 17309265] [*First publication describing targeted mTAG labeling of DNA.]
5. Zhang J, Zheng YG. SAM/SAH Analogs as Versatile Tools for SAM-Dependent Methyltransferases. *ACS Chem Biol.* 2016; 11:583–597. [PubMed: 26540123]
6. Deen J, Vranken C, Leen V, Neely RK, Janssen KPF, Hofkens J. Methyltransferase-Directed Labeling of Biomolecules and its Applications. *Angew Chem Int Ed Engl.* 2017; 56:5182–5200. [PubMed: 27943567]
7. Tomkuvien M, Kriukien E, Klimašauskas S. DNA Labeling Using DNA Methyltransferases. *Adv Exp Med Biol.* 2016; 945:511–535. [PubMed: 27826850]
8. Anhauser L, Rentmeister A. Enzyme-mediated tagging of RNA. *Curr Opin Biotechnol.* 2017; 48:69–76. [PubMed: 28399446]
9. Klimašauskas S, Weinhold E. A new tool for biotechnology: AdoMet-dependent methyltransferases. *Trends Biotechnol.* 2007; 25:99–104. [PubMed: 17254657]
10. Dalhoff C, Lukinavičius G, Klimašauskas S, Weinhold E. Direct transfer of extended groups from synthetic cofactors by DNA methyltransferases. *Nat Chem Biol.* 2006; 2:31–32. [PubMed: 16408089]
11. Willnow S, Martin M, Lüscher B, Weinhold E. A Selenium-Based Click AdoMet Analogue for Versatile Substrate Labeling with Wild-Type Protein Methyltransferases. *ChemBioChem.* 2012; 13:1167–1173. [PubMed: 22549896]
12. Lukinavičius G, Tomkuvien M, Masevičius V, Klimašauskas S. Enhanced chemical stability of adomet analogues for improved methyltransferase-directed labeling of DNA. *ACS Chem Biol.* 2013; 8:1134–1139. [PubMed: 23557731]
13. Muttach F, Muthmann N, Reichert D, Anhauser L, Rentmeister A. A benzylic linker promotes methyltransferase catalyzed norbornene transfer for rapid bioorthogonal tetrazine ligation. *Chem Sci.* 2017; 8:7947–7953. [PubMed: 29619168]
14. Lauer MH, Vranken C, Deen J, Frederickx W, Vanderlinden W, Wand N, Leen V, Gehlen MH, Hofkens J, Neely RK. Methyltransferase-directed covalent coupling of fluorophores to DNA. *Chem Sci.* 2017; 8:3804–3811. [PubMed: 28580113]
15. Schulz D, Holstein JM, Rentmeister A. A chemo-enzymatic approach for site-specific modification of the RNA cap. *Angew Chem Int Ed Engl.* 2013; 52:7874–7878. [PubMed: 23794451]
16. Stummer D, Herrmann C, Rentmeister A. Quantum Chemical Calculations and Experimental Validation of the Photoclick Reaction for Fluorescent Labeling of the 5' cap of Eukaryotic mRNAs. *ChemistryOpen.* 2015; 4:295–301. [PubMed: 26246991]
17. Holstein JM, Stummer D, Rentmeister A. Enzymatic modification of 5'-capped RNA with a 4-vinylbenzyl group provides a platform for photoclick and inverse electron-demand Diels-Alder reaction. *Chem Sci.* 2015; 6:1362–1369. [PubMed: 29560223]
18. Shu X, Dai Q, Wu T, Bothwell IR, Yue Y, Zhang Z, Cao J, Fei Q, Luo M, He C, et al. N(6)-Allyl adenosine: A New Small Molecule for RNA Labeling Identified by Mutation Assay. *J Am Chem Soc.* 2017; 139:17213–17216. [PubMed: 29116772]
19. Muttach F, Masing F, Studer A, Rentmeister A. New AdoMet Analogues as Tools for Enzymatic Transfer of Photo-Cross-Linkers and Capturing RNA-Protein Interactions. *Chemistry.* 2017; 23:5988–5993. [PubMed: 28042932]

20. Anhauser L, Muttach F, Rentmeister A. Reversible modification of DNA by methyltransferase-catalyzed transfer and light-triggered removal of photo-caging groups. *Chem Commun.* 2018; 54:449–451.
21. Grunwald A, Dahan M, Giesbertz A, Nilsson A, Nyberg LK, Weinhold E, Ambjornsson T, Westerlund F, Ebenstein Y. Bacteriophage strain typing by rapid single molecule analysis. *Nucleic Acids Res.* 2015; 43:e117. [PubMed: 26019180]
22. Gilboa T, Torfstein C, Juhasz M, Grunwald A, Ebenstein Y, Weinhold E, Meller A. Single-Molecule DNA Methylation Quantification Using Electro-optical Sensing in Solid-State Nanopores. *ACS Nano.* 2016; 10:8861–8870. [PubMed: 27580095] [*First demonstration of single-molecule analysis of mTAG-labeled DNA using nanopore sensing.]
23. Chen K, Juhasz M, Gularek F, Weinhold E, Tian Y, Keyser UF, Bell NAW. Ionic Current-Based Mapping of Short Sequence Motifs in Single DNA Molecules Using Solid-State Nanopores. *Nano Lett.* 2017; 17:5199–5205. [PubMed: 28829136]
24. Osipenko A, Plotnikova A, Nainyt M, Masevi ius V, Klimašauskas S, Vilkaitis G. Oligonucleotide-Addressed Covalent 3'-Terminal Derivatization of Small RNA Strands for Enrichment and Visualization. *Angew Chem Int Ed Engl.* 2017; 56:6507–6510. [PubMed: 28471013]
25. Mickut M, Nainyt M, Vasiliauskait L, Plotnikova A, Masevi ius V, Klimašauskas S, Vilkaitis G. Animal Hen1 2'-O-methyltransferases as tools for 3'-terminal functionalization and labelling of single-stranded RNAs. *Nucleic Acids Res.* 2018; doi: 10.1093/nar/gky514 [*The first example of methyltransferase-directed covalent modification of an entire major class of biological macromolecules – single stranded RNAs.]
26. Huber TD, Johnson BR, Zhang J, Thorson JS. AdoMet analog synthesis and utilization: current state of the art. *Curr Opin Biotechnol.* 2016; 42:189–197. [PubMed: 27506965]
27. Masevi ius V, Nainyt M, Klimašauskas S. Synthesis of S-Adenosyl-L-Methionine Analogs with Extended Transferable Groups for Methyltransferase-Directed Labeling of DNA and RNA. *Curr Protoc Nucleic Acid Chem.* 2016; 64:1.36.1–13. [PubMed: 26967468]
28. Deen J, Wang S, Van Snick S, Leen V, Janssen K, Hofkens J, Neely RK. A general strategy for direct, enzyme-catalyzed conjugation of functional compounds to DNA. *Nucleic Acids Res.* 2018; 46:e64. [PubMed: 29546351]
29. Singh S, Zhang J, Huber TD, Sunkara M, Hurley K, Goff RD, Wang G, Zhang W, Liu C, Rohr J, et al. Facile chemoenzymatic strategies for the synthesis and utilization of S-adenosyl-(L)-methionine analogues. *Angew Chem Int Ed Engl.* 2014; 53:3965–3969. [PubMed: 24616228]
30. Wang R, Islam K, Liu Y, Zheng W, Tang H, Lailler N, Blum G, Deng H, Luo M. Profiling genome-wide chromatin methylation with engineered posttranslation apparatus within living cells. *J Am Chem Soc.* 2013; 135:1048–1056. [PubMed: 23244065]
31. Dippe M, Brandt W, Rost H, Porzel A, Schmidt J, Wessjohann LA. Rationally engineered variants of S-adenosylmethionine (SAM) synthase: reduced product inhibition and synthesis of artificial cofactor homologues. *Chem Commun.* 2015; 51:3637–3640.
32. Muttach F, Rentmeister A. A Biocatalytic Cascade for Versatile One-Pot Modification of mRNA Starting from Methionine Analogues. *Angew Chem Int Ed Engl.* 2016; 55:1917–1920. [PubMed: 26694947]
33. Hartstock K, Nilges BS, Ovcharenko A, Cornelissen NV, Pullen N, Lawrence-Dorner AM, Leidel SA, Rentmeister A. Enzymatic or In Vivo Installation of Propargyl Groups in Combination with Click Chemistry for the Enrichment and Detection of Methyltransferase Target Sites in RNA. *Angew Chem Int Ed Engl.* 2018; 57:6342–6346. [PubMed: 29461645] [**This publication describes the first in vivo mTAG labeling and analysis of RNA methyltransferase target sites in human cells.]
34. Thomsen M, Vogensen SB, Buchardt J, Burkart MD, Clausen RP. Chemoenzymatic synthesis and in situ application of S-adenosyl-L-methionine analogs. *Org Biomol Chem.* 2013; 11:7606–7610. [PubMed: 24100405]
35. Lipson JM, Thomsen M, Moore BS, Clausen RP, La Clair JJ, Burkart MD. A tandem chemoenzymatic methylation by S-adenosyl-L-methionine. *Chembiochem.* 2013; 14:950–953. [PubMed: 23650044]

36. Plotnikova A, Osipenko A, Masevičius V, Vilkaitis G, Klimašauskas S. Selective covalent labeling of miRNA and siRNA duplexes using HEN1 methyltransferase. *J Am Chem Soc.* 2014; 136:13550–13553. [PubMed: 25170533]
37. Holstein JM, Anhauser L, Rentmeister A. Modifying the 5'-Cap for Click Reactions of Eukaryotic mRNA and To Tune Translation Efficiency in Living Cells. *Angew Chem Int Ed Engl.* 2016; 55:10899–10903. [PubMed: 27511141]
38. Lukinavičius G, Lapinaitis A, Urbanavičiūtė G, Gerasimaitis R, Klimašauskas S. Engineering the DNA cytosine-5 methyltransferase reaction for sequence-specific labeling of DNA. *Nucleic Acids Res.* 2012; 40:11594–11602. [PubMed: 23042683]
39. Kriukienė E, Labrie V, Khare T, Urbanavičiūtė G, Lapinaitis A, Koncėvičius K, Li D, Wang T, Pai S, Ptak C, et al. DNA unmethylome profiling by covalent capture of CpG sites. *Nat Commun.* 2013; 4
40. Tomkuvienė M, Clouet-d'Orval B, Berniauskas I, Weinhold E, Klimašauskas S. Programmable sequence-specific click-labeling of RNA using archaeal box C/D RNP methyltransferases. *Nucleic Acids Res.* 2012; 40:6765–6773. [PubMed: 22564896]
41. Das SK, Austin MD, Akana MC, Deshpande P, Cao H, Xiao M. Single molecule linear analysis of DNA in nano-channel labeled with sequence specific fluorescent probes. *Nucleic Acids Res.* 2010; 38:e177–e177. [PubMed: 20699272]
42. Neely RK, Dedecker P, Hotta J-i, Urbanavičiūtė G, Klimašauskas S, Hofkens J. DNA fluorocode: A single molecule, optical map of DNA with nanometre resolution. *Chem Sci.* 2010; 1:453.
43. Vranken C, Deen J, Dirix L, Stakenborg T, Dehaen W, Leen V, Hofkens J, Neely RK. Super-resolution optical DNA Mapping via DNA methyltransferase-directed click chemistry. *Nucleic Acids Res.* 2014; 42:e50. [PubMed: 24452797]
44. Jeffet J, Kobo A, Su T, Grunwald A, Green O, Nilsson AN, Eisenberg E, Ambjörnsson T, Westerlund F, Weinhold E, et al. Super-Resolution Genome Mapping in Silicon Nanochannels. *ACS Nano.* 2016; 10:9823–9830. [PubMed: 27646634]
45. Staševskij Z, Gibas P, Gordevičius J, Kriukienė E, Klimašauskas S. Tethered Oligonucleotide-Primed Sequencing, TOP-Seq: A High-Resolution Economical Approach for DNA Epigenome Profiling. *Mol Cell.* 2017; 65:554–564 e556. [PubMed: 28111014] [**This publication describes a new technique that combines mTAG labeling and innovative priming of DNA polymerase reads at the tagged sites which permitted high-resolution economical whole-genome profiling of methylation sites in human tissues and cancer cell lines.]
46. Cao H, Hastie AR, Cao D, Lam ET, Sun Y, Huang H, Liu X, Lin L, Andrews W, Chan S, et al. Rapid detection of structural variation in a human genome using nanochannel-based genome mapping technology. *GigaScience.* 2014; 3:34. [PubMed: 25671094]
47. Grunwald A, Sharim H, Gabrieli T, Michaeli Y, Torchinsky D, Arieli R, Juhasz M, Wagner KR, Pevsner J, Reifengerger J, et al. Reduced representation optical methylation mapping (R2OM2). *bioRxiv.* 2017
48. Liutkevičiūtė Z, Kriukienė E, Grigaitytė I, Masevičius V, Klimašauskas S. Methyltransferase-directed derivatization of 5-hydroxymethylcytosine in DNA. *Angew Chem Int Ed Engl.* 2011; 50:2090–2093. [PubMed: 21344558]
49. Serva S, Lagunavicius A. Direct conjugation of peptides and 5-hydroxymethylcytosine in DNA. *Bioconjug Chem.* 2015; 26:1008–1012. [PubMed: 25982286]
50. Strepėkaitis D, Alzbutas G, Astromskas E, Lagunavičius A, Sabaliauskaitis R, Arbačiauskas K, Lazutka J. Analysis of DNA Methylation and Hydroxymethylation in the Genome of Crustacean *Daphnia pulex*. *Genes.* 2015; 7:1.
51. Irwin RE, Thakur A, KM ON, Walsh CP. 5-Hydroxymethylation marks a class of neuronal gene regulated by intragenic methylcytosine levels. *Genomics.* 2014; 104:383–392. [PubMed: 25179375]
52. Yang Z, Shi Y, Liao W, Yin H, Ai S. A novel signal-on photoelectrochemical biosensor for detection of 5-hydroxymethylcytosine based on in situ electron donor producing strategy and all wavelengths of light irradiation. *Sens Actuator B-Chem.* 2016; 223:621–625.

53. Klungland A, Robertson AB. Oxidized C5-methyl cytosine bases in DNA: 5-Hydroxymethylcytosine; 5-formylcytosine; and 5-carboxycytosine. *Free Radic Biol Med*. 2017; 107:62–68. [PubMed: 27890639]
54. Kriukien E, Liutkevič Z, Klimašauskas S. 5-Hydroxymethylcytosine--the elusive epigenetic mark in mammalian DNA. *Chem Soc Rev*. 2012; 41:6916–6930. [PubMed: 22842880]
55. Song CX, Sun Y, Dai Q, Lu XY, Yu M, Yang CG, He C. Detection of 5-hydroxymethylcytosine in DNA by transferring a keto-glucose by using T4 phage beta-glucosyltransferase. *Chembiochem*. 2011; 12:1682–1685. [PubMed: 21656634]
56. Song CX, Szulwach KE, Fu Y, Dai Q, Yi C, Li X, Li Y, Chen CH, Zhang W, Jian X, et al. Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine. *Nat Biotechnol*. 2011; 29:68–72. [PubMed: 21151123] [*Development and application of 6-azido UDP-Glc cofactor for covalent tagging of hmC residues in mammalian DNA.]
57. Dai N, Bitinaite J, Chin H-G, Pradhan S, Corrêa IR. Evaluation of UDP-GlcN Derivatives for Selective Labeling of 5-(Hydroxymethyl)cytosine. *ChemBioChem*. 2013; 14:2144–2152. [PubMed: 24106095]
58. Nifker G, Levy-Sakin M, Berkov-Zrihen Y, Shahal T, Gabrieli T, Fridman M, Ebenstein Y. One-Pot Chemoenzymatic Cascade for Labeling of the Epigenetic Marker 5-Hydroxymethylcytosine. *ChemBioChem*. 2015; 16:1857–1860. [PubMed: 26147952]
59. Song CX, Yin S, Ma L, Wheeler A, Chen Y, Zhang Y, Liu B, Xiong J, Zhang W, Hu J, et al. 5-Hydroxymethylcytosine signatures in cell-free DNA provide information about tumor types and stages. *Cell Res*. 2017; 27:1231–1242. [PubMed: 28820176] [*Successful adaptation of enzyme-directed covalent labeling for cancer epigenotyping in cell-free DNA.]
60. Li W, Zhang X, Lu X, You L, Song Y, Luo Z, Zhang J, Nie J, Zheng W, Xu D, et al. 5-Hydroxymethylcytosine signatures in circulating cell-free DNA as diagnostic biomarkers for human cancers. *Cell Res*. 2017; 27:1243–1257. [PubMed: 28925386] [*Successful adaptation of enzyme-directed covalent labeling for cancer epigenotyping in cell-free DNA.]
61. Danait K, Jarmalait S, Kriukien E. Epigenomic technologies for deciphering circulating tumor DNA. *Curr Opin Biotech*. 2018; 55:23–29. [PubMed: 30081280]
62. Zahid OK, Zhao BS, He C, Hall AR. Quantifying mammalian genomic DNA hydroxymethylcytosine content using solid-state nanopores. *Sci Rep*. 2016; 6
63. Gilat N, Tabachnik T, Shwartz A, Shahal T, Torchinsky D, Michaeli Y, Nifker G, Zirkin S, Ebenstein Y. Single-molecule quantification of 5-hydroxymethylcytosine for diagnosis of blood and colon cancers. *Clin Epigenetics*. 2017; 9:70. [PubMed: 28725280]
64. Song CX, Diao J, Brunger AT, Quake SR. Simultaneous single-molecule epigenetic imaging of DNA methylation and hydroxymethylation. *Proc Natl Acad Sci U S A*. 2016; 113:4338–4343. [PubMed: 27035984]
65. Gabrieli T, Sharim H, Nifker G, Jeffet J, Shahal T, Arielly R, Levi-Sakin M, Hoch L, Arbib N, Michaeli Y, et al. Epigenetic Optical Mapping of 5-Hydroxymethylcytosine in Nanochannel Arrays. *ACS Nano*. 2018; 12:7148–7158. [PubMed: 29924591] [**The first work to demonstrate single-molecule dual optical mapping of genetic and epigenetic features in large fragments of human DNA.]
66. Boccaletto P, Machnicka MA, Purta E, Piatkowski P, Baginski B, Wirecki TK, de Crecy-Lagard V, Ross R, Limbach PA, Kotter A, et al. MODOMICS: a database of RNA modification pathways. 2017 update. *Nucleic Acids Res*. 2018; 46:D303–D307. [PubMed: 29106616]
67. Holstein JM, Schulz D, Rentmeister A. Bioorthogonal site-specific labeling of the 5'-cap structure in eukaryotic mRNAs. *Chem Commun*. 2014; 50:4478–4481.
68. Holstein JM, Muttach F, Schiefelbein SHH, Rentmeister A. Dual 5' Cap Labeling Based on Regioselective RNA Methyltransferases and Bioorthogonal Reactions. *Chemistry*. 2017; 23:6165–6173. [PubMed: 27869340] [*First demonstration of a one-pot cascade reaction for tandem mTAG modification of mRNA starting from methionine analogs.]
69. Baranausk S, Mickut M, Plotnikova A, Finke A, Venclovas , Klimašauskas S, Vilkaitis G. Functional mapping of the plant small RNA methyltransferase: HEN1 physically interacts with HYL1 and DICER-LIKE 1 proteins. *Nucleic Acids Res*. 2015; 43:2802–2812. [PubMed: 25680966]

70. Alberti C, Manzenreither RA, Sowemimo I, Burkard TR, Wang J, Mahofsky K, Ameres SL, Cochella L. Cell-type specific sequencing of microRNAs from complex animal tissues. *Nat Methods*. 2018; 15:283–289. [PubMed: 29481550]
71. Motorin Y, Burhenne J, Teimer R, Koynov K, Willnow S, Weinhold E, Helm M. Expanding the chemical scope of RNA:methyltransferases to site-specific alkynylation of RNA for click labeling. *Nucleic Acids Res*. 2011; 39:1943–1952. [PubMed: 21037259]
72. Guo M, Liu X, Zheng X, Huang Y, Chen X. m6A RNA Modification Determines Cell Fate by Regulating mRNA Degradation. *Cell Reprogram*. 2017; 19:225–231. [PubMed: 28682669]
73. Li F, Dong J, Hu X, Gong W, Li J, Shen J, Tian H, Wang J. A covalent approach for site-specific RNA labeling in Mammalian cells. *Angew Chem Int Ed Engl*. 2015; 54:4597–4602. [PubMed: 25694369]
74. Alexander SC, Busby KN, Cole CM, Zhou CY, Devaraj NK. Site-Specific Covalent Labeling of RNA by Enzymatic Transglycosylation. *J Am Chem Soc*. 2015; 137:12756–12759. [PubMed: 26393285]
75. Zhou CY, Alexander SC, Devaraj NK. Fluorescent turn-on probes for wash-free mRNA imaging via covalent site-specific enzymatic labeling. *Chem Sci*. 2017; 8:7169–7173. [PubMed: 29081948]
76. Ehret F, Zhou CY, Alexander SC, Zhang D, Devaraj NK. Site-Specific Covalent Conjugation of Modified mRNA by tRNA Guanine Transglycosylase. *Mol Pharm*. 2018; 15:737–742. [PubMed: 28749687]
77. Zhang D, Zhou CY, Busby KN, Alexander SC, Devaraj NK. Light-Activated Control of Translation by Enzymatic Covalent mRNA Labeling. *Angew Chem Int Ed Engl*. 2018; 57:2822–2826. [PubMed: 29380476]

Highlights

- Enzymatic transfer of chemically extended groups enables targeted covalent functionalization of DNA and RNA
- Methyltransferase-directed labeling of DNA and RNA provides novel tools for (epi)genomic and (epi)transcriptomic studies
- Detection of individual covalently-bound reporters on DNA permits single-molecule analysis of (epi)genomes

Biological R moieties are shown in grey, transferred functional groups – blue, reporters – magenta. A, Top. AdoMet analog and methyltransferase (MTase)-dependent modification of DNA and RNA (for “R” variants see Figure 2). Bottom. Non-canonical modification of hmC residues in DNA with aliphatic thiols by DNA cytosine-5 methyltransferases. HisTag and StrepTag – peptidyl affinity tags. B. UDP-glc analog and beta-glucosyltransferase (BGT)-dependent modification of hmC residues in DNA. C. Agmatine analog and tRNA^{Ile2}–

agmatidine synthetase (Tias)-dependent modification of RNA. D. PreQ1 derivative and tRNA guanine transglycosylase (TGT)-dependent modification of RNA.

Figure 2. Structure of synthetic AdoMet analogs containing activated side chains with functional or reporter groups.

Incorporation of an activating π -electron system (a multiple C-C bond or an aromatic ring, shown in green) next to the transferable carbon atom facilitates MTase-directed transfer of the side chain. Besides the mandatory activating group, linear linkers (shown in black) are used to carry terminal alkyne, azide, amino, alkene or norbornene functional groups (shown in blue). Reporter groups such as fluorophores or biotin (shown in magenta) can be further added to a transferable group (via coupling chemistries mentioned above) to yield cofactors suitable for one-step mTAG labeling. Some transferrable side chains combine an activating

and functional group (terminal alkene or alkyne, photo-reactive aryl-azide or diazine groups, photo-removable 2-nitrobenzyl) in a single moiety (shown in brown).

A. DNA modification

B. RNA modification

Figure 3. Substrate specificity of AdoMet-dependent methyltransferases recruited for targeted covalent derivatization of DNA and RNA.

“R” – extended transferable group. Enzymes that require engineering for mTAG reactions are marked with an asterisk. A. DNA MTases with their respective sequence recognition sites (the target base underlined) and the positions of base modification. B. RNA MTases with their respective target sites in the mRNA cap, an internal sequence or a 3'-terminal nucleoside.

Figure 4. Schematic survey of transferase-directed covalent derivatization of nucleic acids and its applications.

A. Site-specific methyltransferase (MTase)-directed tagging of DNA and RNA: B. Beta-glucosyltransferase (BGT)-directed tagging of 5-hydroxymethylcytosine in DNA. C. tRNA-modifying enzyme (tRNA guanine transglycosylase, TGT, or tRNA^{Ile2}-agmatidine synthetase, Tias)-directed tagging of RNA.