Department of

Biological DNA Modification

Chief Scientist and Head

Prof. Saulius Klimašauskas, PhD, Dr. Habil.

phone: +370 5 260214; fax: +370 5 2602116

e-mail: klimasau@ibt.lt; saulius.klimasauskas@bti.vu.lt

http://www.ibt.lt/dmtl_en.html

Scientific staff

Giedrius Vilkaitis, PhD
Edita Kriukienė, PhD
Viktoras Masevičius, PhD
Rasa Rakauskaitė, PhD
Zita Liutkevičiūtė, PhD
Miglė Tomkuvienė, PhD
Juozas Gordevičius, PhD
Zdislav Staševskij, M.Sc.
Giedrė Urbanavičiūtė, M.Sc.
Alexandra Plotnikova, B.Sc.
Aleksandr Osipenko, M.Sc.
Simona Baranauskė, M.Sc.
Stasė Butkytė, M.Sc.
Milda Mickutė, M.Sc.
Janina Ličytė, M.Sc.
Audronė Rukšėnaitė, M.Sc.

Technical staff

Daiva Gedgaudienė

Undergraduate Students

Inga Burneikienė
Dominyka Grigaitė
Jurgita Špakovska
Indrė Grigaitytė
Milda Rudytė
Ingrida Olendraitė
Agnė Vailionytė
Ieva Stirblytė
Eglė Mazgelytė
Mantas Šarauskas
Laurynas Čiupas
Sigita Grigaitytė

AdoMet-dependent methyltransferases (MTases), which represent more than 3% of the proteins in the cell, catalyze the transfer of the methyl group from S-adenosyl-L-methionine (SAM or AdoMet) to N-, C-, O- or S-nucleophiles in DNA, RNA, proteins or small biomolecules. In DNA of mammals, cytosines are often methylated at the 5-position of the pyrimidine ring to give 5-methylcytosine (5mC). DNA methylation profiles are highly variable across different genetic loci, cells and organisms, and are dependent on tissue, age, sex, diet, and disease. Besides 5mC, certain genomic DNAs have been shown to contain substantial amounts of 5-hydroxymethyl-cytosine (hmC). It was demonstrated that hmC is predominantly produced via oxidation of 5mC residues by TET oxygenases and that it can be further oxidized to form 5-formylcytosine (fC) and 5-carboxylcytosine (caC) (Fig. 1). Current evidence suggests that although hmC, fC and caC are intermediates on the pathway of active DNA demethylation, and the multiplicity of epigenetic states may also play independent roles in embryonic development, brain function and cancer progression. Therefore, a full appreciation of the biological significance of epigenetic regulation in mammals will require the development of novel tools that allow hmC, 5mC and C to be distinguished unequivocally.

Novel cofactor-independent reactions of DNA methyltransferases

S-Adenosylmethionine-dependent DNA methyltransferases (MTases) perform direct methylation of cytosine to yield 5-methylcytosine (5mC), which serves as part of the epigenetic regulation mechanism in vertebrates. Previously we found that (i) DNA C5-MTases catalyze covalent addition of exogenous aliphatic aldehydes to their target residues in DNA, yielding corresponding 5-α-hydroxyalkylcytosines and (ii) can promote the reverse reaction – the removal of formaldehyde from hmC in DNA (Liutkevičiūtė et al., Nat. Chem. Biol., 2009, 5: 400–402). We have also discovered that bacterial C5-MTases can catalyze in vitro condensation of aliphatic thiols and selenols to 5-hydroxymethyl-cytosine in DNA yielding 5-chalcogenomethyl derivatives (Liutkevičiūtė et al., Angew. Chem. Int. Ed., 2011, 50: 2090–2093). Most recently, we have demonstrated that both bacterial and mammalian C5-MTases can catalyze the direct decarboxylation of caC yielding unmodified cytosine in DNA in vitro but are inert toward fC. The observed atypical enzymatic C–C bond cleavage reaction provides a

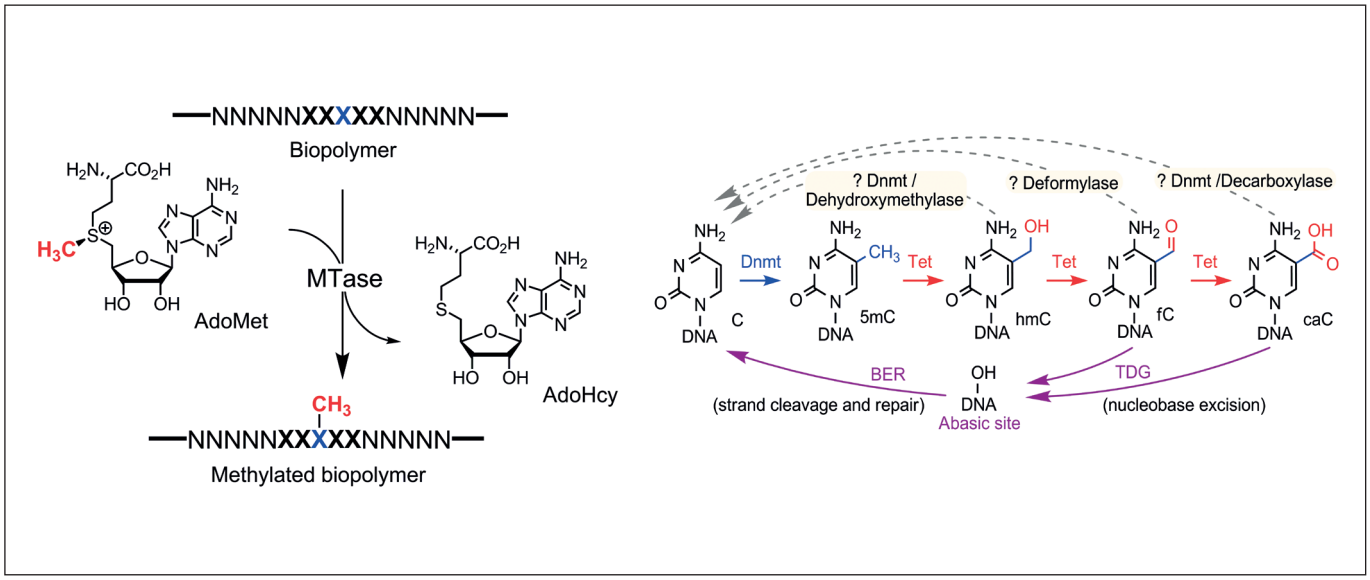

Figure 1. Biological Methylation and Demethylation. Left, targeted transfer of the methyl group (red) from the S-adenosyl-methionine cofactor (AdoMet) to a specific position of a biopoplymer by a methyltransferase (MTase) releasing S-adenosyl-homocysteine (AdoHcy) coproduct. Right, Biological DNA methylation and demethylation in vertebrates. Cytosine (C) is converted to 5-methylcytosine (5mC) by endogenous C5-MTases of the Dnmt1 and Dnmt3 families (blue); 5mC can be consecutively converted to 5-hydroxymethylcytosine (hmC), 5-formylcytosine (fC), and 5-carboxylcytosine (caC) by the TET oxygenases (red). Known reverse pathways (magenta) involve base excision repair (BER) of fC and caC by thymine DNA glycosylase (TDG), leading to transient formation of abasic sites in DNA. Dashed arrows denote the newly discovered 5-dehydroxymethylation and 5-decarboxylation reactions performed by C5-MTases and which may be carried out in vivo by Dnmts or other putative enzymes.

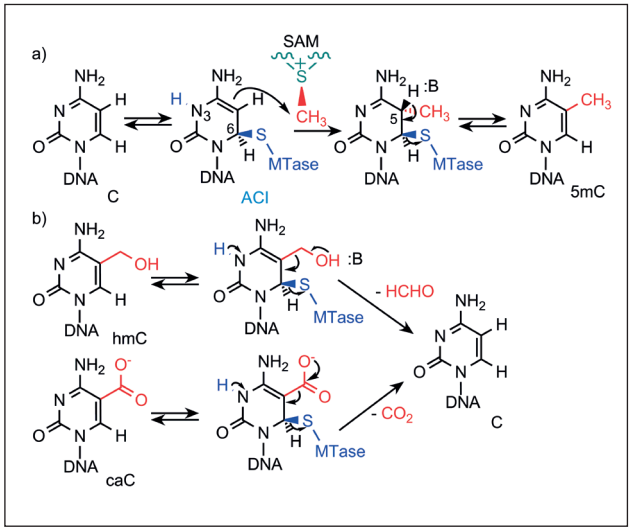

Figure 2. Mechanisms of covalent catalysis by DNA cytosine-5 MTases. A) Biological DNA methylation involves a covalent addition of a catalytic Cys residue at the C6 position of the cytosine, leading to an activated covalent intermediate (ACI) and sebsequnt tranfer of the activated methyl group from the SAM cofactor. B) Atypical C-C bond cleavage reactions catalyzed by MTases, dehydroxymethylation (upper) and decarboxylation (lower), also use the covalent activation mechanism.

plausible precedent for a direct reversal of caC to the unmodified state in DNA and offers a unique approach for sequence-specific analysis of genomic caC [5].

The MTase-directed decarboxylation reaction occurs under mild conditions and retains the cognate sequence- and nucleotide-specificity of the enzyme. The M.SssI MTase is also similarly active toward hmC, but not toward 5mC and fC. Based on these findings we developed a new method for CpG-specific analysis of genomic caC residues in which hmC sites are blocked by treatment of DNA with T4-β-glucosyltransferase (BGT), whereas C, 5mC, and fC sites are selectively eliminated by R.MspI cleavage and end-processing [5].

Engineering the catalytic reaction of methyltransferases for targeted covalent labeling of DNA

The ability of most MTases to catalyze highly specific covalent modifications of biopolymers makes them attractive molecular tools, provided that the transfer of larger chemical entities can be achieved. Our long standing effort is aimed at **redesigning the enzymatic methyltransferase reactions for targeted covalent deposition of desired functional or reporter groups** onto biopolymer molecules such as **DNA and RNA**. We had synthesized a series of model AdoMet analogs with sulfonium-bound extended side chains replacing the methyl group and showed that allylic and propargylic side chains can be efficiently transferred by DNA MTases with high sequence- and base-specificity (Dalhoff et al, Nature Chem. Biol., 2006, 2: 31–32; Klimašauskas and Weinhold, Trends Biotechnol., 2007, 25: 99–104) in collaboration with the group of Prof. Elmar Weinhold (RWTH Aachen, Germany). Using DNA MTases along with their novel cofactors that carry useful functional or reporter groups, we demonstrated that our new approach name mTAG (methyltransferase-directed Transfer of Activated Groups) can be used for sequence-specific functionalization and labeling of a wide variety of model and natural DNA substrates (see below). To further optimize the efficacy of the mTAG reactions we performed steric engineering of the cofactor pocket in the M.HhaI C5-MTase by systematic replacement of three non-essential positions (Lukinavičius et al., Nucleic Acids Res., 2012, 40, 11594–11602). Analogous replacements of two conserved residues in M.SssI C5-MTase also resulted in improved transalkylation activity [1] attesting a general applicability of the homology-guided engineering to the C5-MTase family and expanding the repertoire of sequence-specific tools for covalent in vitro and ex vivo labeling of DNA.

Improved synthetic AdoMet analogues for methyltransferase-directed labeling of DNA

Although clearly useful in certain applications, our previously described series of AdoMet analogues carrying sulfonium-bound 4-substituted but-2-ynyl side chains (Lukinavicius et al., J. Amer. Chem. Soc., 2007, 129: 2758), exhibited short lifetimes in physiological buffers. Examination of the reaction kinetics and products showed that their fast inactivation followed a differnt pathway than observed for AdoMet and rather involved a pH-dependent addition of a water molecule to the side chain. This side reaction was eradicated by synthesis of a series of cofactor analogues in which the separation between an electronegative group and the triple bond was increased from one to three carbon units. The designed hex-2-ynyl moiety-based cofactor analogues with terminal amino, azide, or alkyne groups showed a markedly improved enzymatic transalkylation

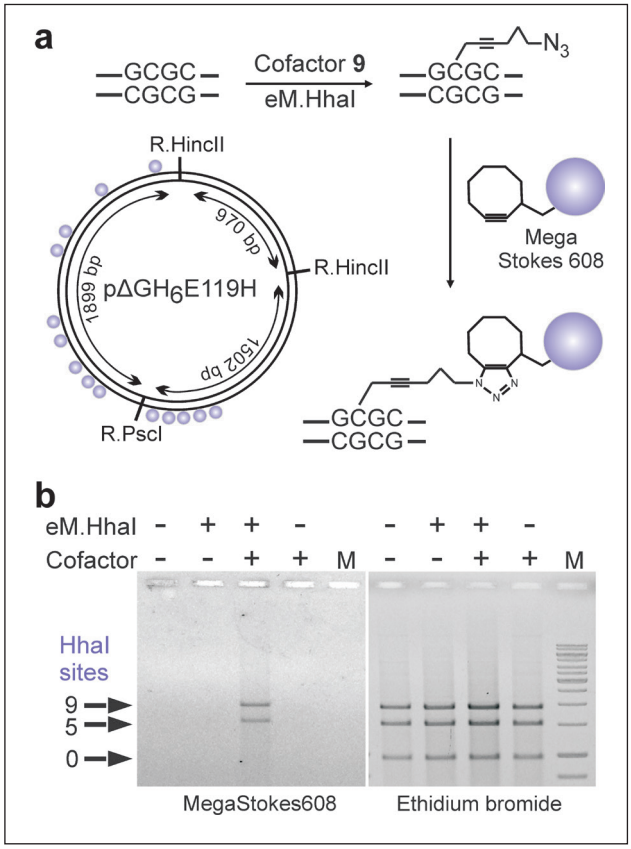

Figure 3. Sequence-specific two-step mTAG labeling of plasmid DNA using copper-free click reaction in crude cell extract. a) Escherichia coli ER2267 cells carrying the pΔGH6E119H plasmid were harvested, and crude lysate was treated with eM.HhaI (engineered HhaI MTase) and an azide cofactor analogue for 3 h at 37 °C. Alkyne MegaStokes 608 dye was added, and incubation continued for another 3 h. b) Image of agarose gel electrophoresis of the labeled plasmid fragmented with R.HincII and R.PscI endonucleases. Number of HhaI sites is indicated on the left of the fragment-pointing arrows. Control samples lacked M.HhaI or cofactor as indicated. M, DNA size marker.

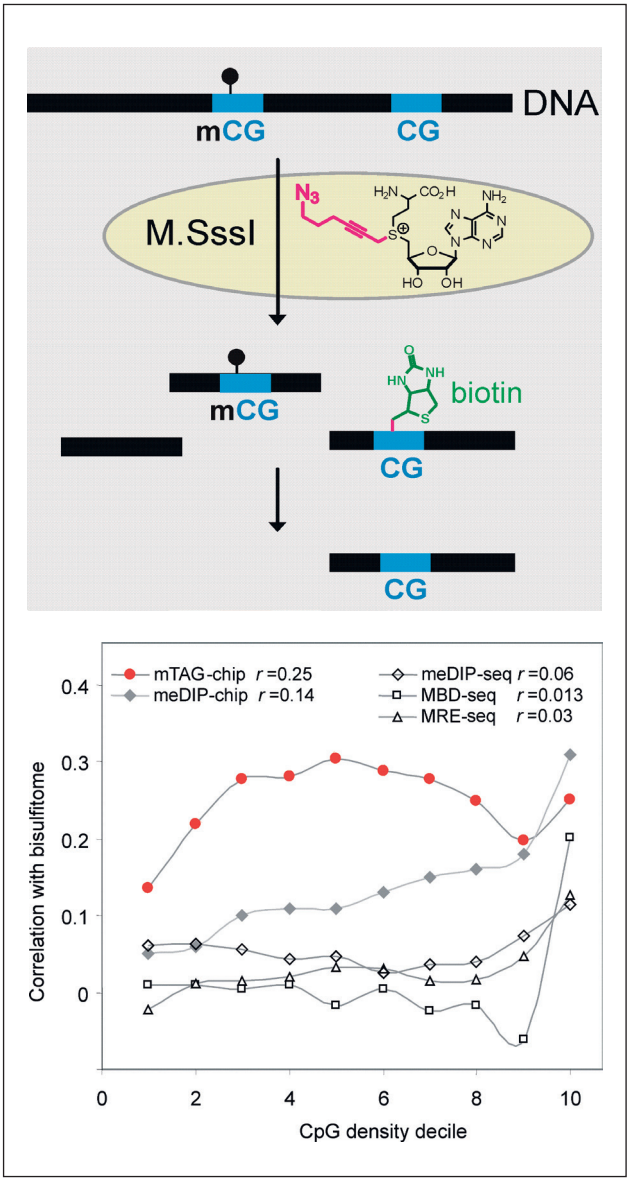

Figure 4. mTAG labeling-based enrichment and analysis of unmethylated CpG sites in the genome.. Left, (a) schematic outline of the analytical procedure. Right, Comparison of the mTAG approach with existing methylome profiling methods. Pearson correlations between experimental mTAG-chip, MeDIP-chip of IMR90 gDNA and published data of MeDIP-seq, MBD-seq, MRE-seq analysis of H1 gDNA were determined for 1 kb tiles on Chr4 against the corresponding MethylC-seq data and stratified according to local CpG density. Aggregate correlation numbers (r) obtained with each analytical procedure are shown above the plots.

activity and proved well suitable for methyltransferase-directed sequence-specific labeling of DNA in vitro and in bacterial cell lysates [2].

Molecular tools for genome analysis: epigenome profiling by covalent capture of unmodified CpG sites

Dynamic patterns of cytosine-5 methylation of CpG dinucleotide and its successive oxidated forms (see Figure 1) are part of epigenetic regulation in eukaryotes, including humans, which contributes to normal phenotypic variation and disease risk. However, no method alone can grasp the breath and chemical complexity of the mammalian epigenome. In collaboration with Prof. Art Petronis, CAMH, Canada, we went on to develop an approach for the mapping of unmodified regions of the genome, which we call the unmethylome [1]. Our technique exploits mTAG-based covalent biotin tagging of unmodified CpG sites using an engineered version of the M.SssI MTase followed by affinity enrichment and interrogation on tiling microarrays or next generation sequencing.

Control experiments and pilot studies of human genomic DNA from cultured cells and tissues demonstrate that, along with providing a unique cross-section through the chemical landscape of the epigenome, the mTAG-based approach offers high precision and robustness as compared with existing affinity-based techniques.

Covalent Labeling of miRNA and siRNA Duplexes Using HEN1 Methyltransferase

miRNAs are important regulators of mammalian gene expression. Individual miRNAs exhibit characteristic expression patterns that vary in normal and disease cells and therefore miRNA profiles provide valuable information for diagnosis and development of novel therapies. Current methods for small RNA discovery largely rely on size-dependent cloning of nucleic acids strands followed sequencing. However, such approaches lack the specificity required to discriminate against similarly sized other types of cellular RNA and DNA molecules and their degradation fragments.

We exploited the mTAG strategy and the specificity of the HEN1 methyltransferase in which miRNA duplexes are selectively targeted for labeling in mixtures of different types of RNA or DNA. We observe a highly efficient labeling of miRNA strands with biotin of fluorophores using a two-step procedure. The two-step approach offers the flexibility in selecting desired labeling chemistries and reporter groups by simply choosing different cofactors and matching chemical probes. We also demonstrate the first single-step labeling of RNA using a novel synthetically produced AdoMet analogue that carries a biotin reporter. The latter labeling modality may be more expensive but will be most useful when speed and simplicity is desired. Altogether, these features make our new molecular tools an extremely valuable addition to the existing analytical toolbox of small RNA paving the way to developing numerous novel analytical techniques. This is the first report of selective labeling of a whole class of cellular RNAs rather than defined nucleotides or sequences [6].

PhD student Milda Mickute caught in preparation for RNA methylation analysis

Roles of the plant 2'-O-methyltransferase HEN1 in microRNA biogenesis

Arabidopsis HEN1 belongs to a family of RNA 2'-O-methyltransferases widespread in both prokaryotes and eukaryotes that share conservative catalytic domain and transfer methyl group on 3'-terminal nucleotide of RNA molecules. The unique feature of the plant homologues is that they covalently modify double-stranded microRNAs or siRNAs. To achieve its biological function the plant HEN1 carries four additional domains preceding the catalytic methyltransferase domain. We experimentally showed that the two double-stranded RNA binding domains of HEN1 considerably but unequally contribute to the binding of mature miRNA/miRNA* duplex, and mapped residues in each domain responsible for this function. Detailed enzymatic and mutational analysis of the RNA interacting domains provided new important insights into the underlying mechanisms of double-stranded RNA specific recognition by 2'-O-methyltransferases [3, 7].

We also showed for the first time direct physical interactions of HEN1 with two proteins essential for plant microRNAs biogenesis, ribonuclease III-type enzyme DICER-LIKE 1 and double-stranded RNA binding protein HYL1, and the lack of direct interactions with SERRATE protein, a third component of the presumed plant miRNA processing complex. We further mapped a central domain of previously unknown function in HEN1 is a key factor for tight binding with HYL1 protein. Based on these findings, we propose model of plant microR-

NAs biogenesis wherein for the first time it was assigned a physical assembly for HEN1 methyltransferase with other proteins involved in maturation of microRNAs (Figure 5). Since HEN1 has been implicated and is the only invariable protein in biogenesis of all small RNA types in plants (miRNA, ta-siRNA, nat-siRNA, ls-siRNA, hc-siRNA and etc.), our studies thus pave the way to understanding the roles of RNA 2'-O-methylation in maturation of other classes of cellular small RNAs [7].

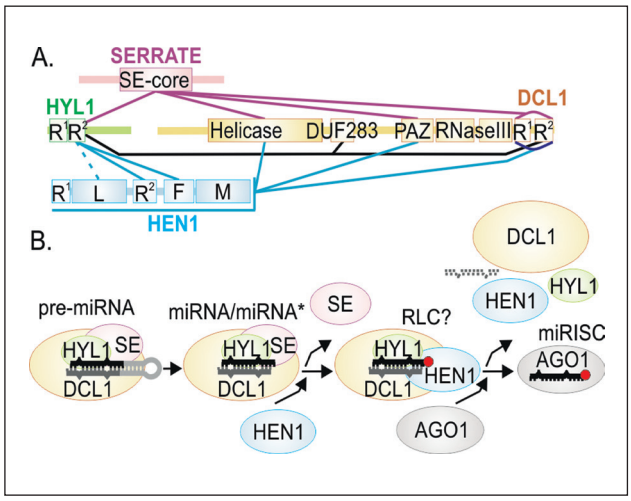

Figure 5. HEN1 interaction network and the proposed model of late stages of miRNA biogenesis. (A) Protein-protein interaction network involving SE, HYL1, DCL1 and HEN1. Cyan lines show interactions experimentally determined in this work, purple and black lines depict those reported previously. (B) Proposed model of miRNA biogenesis envisions that after the miRNR/miRNR duplex is cut out of its precursor, SE is expelled and HEN1 methyltransferase is bound in the microprocessor complex to form a HYL1-HEN1-DCL1 complex, which might represent the still unidentified plant RISC-loading complex (RLC). This complex directs HEN1 methylation (red circle) to the target miRNA strand (black) thus marking it for incorporation into AGO1 complex.*

Biosynthesis of selenoproteins with genetically-encoded photocaged selenocysteines

Engineering and in-cell production of recombinant proteins with desired catalytic capacity is widely exploited for structural and functional studies and for practical applications in medicine and industry. L-Selenocysteine (Sec), the 21st amino acid, endows engineered proteins with new valuable properties due to its enhanced chemical reactivity (higher nucleophilicity, lower pKa, and a lower redox potential) as

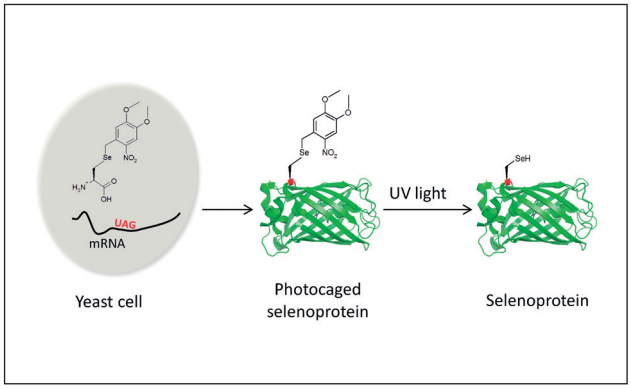

Figure 6. Strategy for incorporation of a photocaged selenocysteine into a genetically-encoded position of a recombinant protein in yeast cells followed by its photochemical decaging in vitro.

compared to cysteine. Despite its high technological potential, targeted incorporation of Sec into recombinant proteins is far from trivial. We developed the first general approach for efficient biosynthesis of selenoproteins containing photocaged selenocysteine residues at genetically predetermined positions. We explored a novel strategy based on a photolabile (4,5-dimethoxy-2-nitrobenzyl, DMNB) group to protect Sec in producing cells and during protein isolation using a yeast expression system originally designed to incorporate DMNB-Ser residues in proteins (Figure 6). We achieved efficient incorporation of DMNB-Sec in a model protein, EGFP, and also demonstrated an efficient photolytic removal of the protecting group from the Se atom, which has not been previously described for any protein or a synthetic peptide. Examples of light-controlled dimerization and site-specific labeling of such recombinant proteins further illustrate robustness and practical utility of the new technique. The generality of this approach is attested by our recent successful production of a HpaII DNA cytosine-5 methyltransferase fusion protein, in which an essential catalytic Cys is replaced with Sec. This paves the way to direct comparison of S- and Se-nucleophiles in the natural and atypical reactions (see above) potentially leading to design of improved molecular tools for genome studies [8].

COLLABORATION

Prof. Art Petronis, CAMH, Toronto U., Canada
Prof. Ting Wang, Washington University School of Medicine, St. Louis, MO, U.S.A.
Prof. Sun-Chong Wang, National Central University, Chungli, Taiwan
Prof. Dr. Elmar Weinhold, RWTH Aachen, Germany
Dr. Andreas Finke, Max Planck Institute for Plant Breeding Research, Cologne, Germany
Dr. Marie-Pierre Chapot-Chartier ir Dr. Saulius Kulakauskas, INRA Institut Micalis, Jouy-en-Josas, France
Prof. Shoji Tajima, Institute for Protein Research, Osaka University, Japan
Dr. Andrey Kulbachinskiy, IMG, Moscow, Russia

FUNDING

National Institutes of Health/National Human Genome Research Institute (USA)
European Social Fund under the Global Grant measure
Research Council of Lithuania
Bilateral French-Lithuanian Collaborative Programme 'Gilibert'
COST actions: TD0905 „Epigenetics: Bench to Bedside" and CM1303 „Systems Biocatalysis".

CONTRACTS

Thermo Fisher Scientific Baltics

SELECTED PUBLICATIONS 2013-2014

1. **Kriukienė E.,** Labrie V., Khare T., **Urbanavičiūtė G., Lapinaitė A.,** Koncevičius K., Li D., Wang T., Pai S., Ptak C., Gordevičius J., Wang Sun-Chong, **Petronis A., Klimašauskas S.** DNA unmethylome profiling by covalent capture of CpG sites. Nature Communications 2013, 4:2190.

2. **Lukinavičius G., Tomkuvienė M., Masevičius V., Klimašauskas S.** Enhanced chemical stability of AdoMet analogs for improved methyltransferase-directed labeling of DNA. ACS Chem Biol 2013, 8: 1134-1139.

3. **Plotnikova A., Baranauskė S., Osipenko A., Klimašauskas S., Vilkaitis G.** Mechanistic insights into small RNA recognition and modification by the HEN1 methyltransferase. Biochem J 2013, 453(2):281-290.

4. Miropolskaya N., Esyunina D., **Klimašauskas S.,** Nikiforov V., Artsimovitch I., Kulbachinskiy A. Interplay between the trigger loop and the F loop during RNA polymerase catalysis. Nucleic Acids Res 2014, 42(1):544-552.

5. **Liutkevičiūtė Z., Kriukienė E., Ličytė J., Rudytė M., Urbanavičiūtė G., Klimašauskas S.** Direct decarboxylation of 5-carboxylcytosine by DNA C5-methyltransferases. J Am Chem Soc 2014, 136(16):5884 5887.

6. **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(39):13550-13553.

7. **Baranauskė S., Mickutė M., Plotnikova A.,** Finke A., **Venclovas Č., Klimašauskas S. and Vilkaitis G.** Functional mapping of the plant small RNA methyltransferase: HEN1 physically interacts with HYL1 and DICER-LIKE 1 proteins. Nucleic Acids Res , in press.

8. **Rakauskaitė R., Urbanavičiūtė G., Rukšėnaitė A., Liutkevičiūtė Z., Juškėnas R., Masevičius V.**, and **Klimašauskas S.** Biosynthetic selenoproteins with genetically-encoded photocaged selenocysteines. Chem Commun in press.

GRANTED PATENTS

Klimašauskas S., Liutkevičiūtė Z., Kriukienė E. Derivatization of biomolecules by covalent coupling of non-cofactor compounds using methyltransferases. EP2414528 (B1), US8822146 (B2). 2014/09/02

Klimašauskas S., Liutkevičiūtė Z., Kriukienė E. Conversion of alpha-hydroxyalkylated residues in biomolecules using methyltransferases. US8889352 (B2). 2014/11/18

PATENT APPLICATIONS

Klimašauskas S., Rakauskaitė R., Masevičius V. Production of selenoproteins. LT2013069; PCT/LT2014/000009.

Klimašauskas S., Staševskij Z. Nucleic acid production and sequence analysis. US2013130249 (A1); WO2013072515 (A1). 2013/05/23

Klimašauskas S., Vilkaitis G., Plotnikova A. Analysis of small RNA. PCT/LT2013/000009. 2013/05/30