Supporting Information

Light-control of cap methylation and mRNA translation via genetic code expansion of Ecm1

Dennis Reichert[1], Henning D. Mootz[1] and Andrea Rentmeister[1,*]

[1]Institute of Biochemistry, Department of Chemistry and Pharmacy, University of Muenster, Corrensstr. 36, D-48149 Münster, Germany

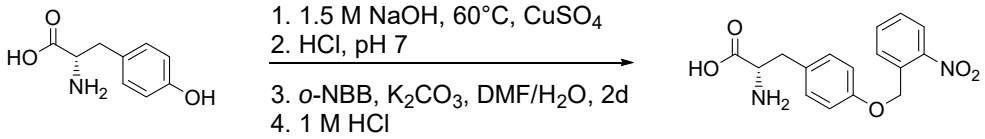
[*]E-mail: a.rentmeister@uni-muenster.de

Synthesis and Characterization of Compounds

General methods

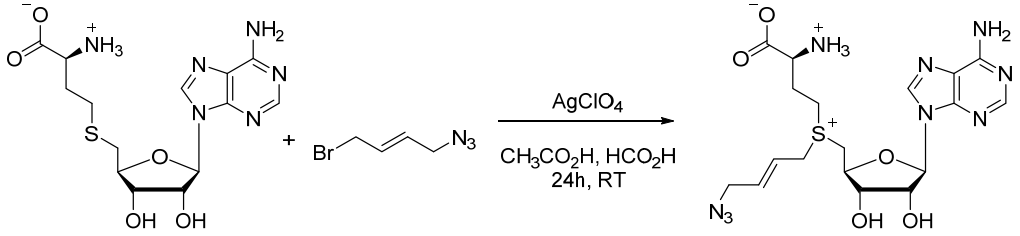
Chemicals were purchased from commercial suppliers in the highest possible grade unless otherwise stated. High-resolution mass spectra were obtained on a Bruker MicroTof (*Bruker Daltonik GmbH*). LC-MS analysis of intact proteins were performed using an UltiMate™ 3000 RS system (*Thermo Fisher Scientific*) connected to a maXis II UHR-qTOF mass spectrometer (*Bruker Daltonik GmbH*) with a standard ESI source (*Apollo, Bruker Daltonik GmbH*). HPLC analysis of reaction samples were performed on a Vanquish™ Horizon UHPLC system (*Thermo Fisher Scientific*).

Synthesis of *o*-nitrobenzyl tyrosine (ONBY)



The amino acid was synthesized as previously described[1] with minor changes. For this purpose, 1 g (5.5 mmol) L-tyrosine was mixed with 3 eq. of NaOH and 0.5 eq. of $CuSO_4$ in 10 mL water and stirred at 60 °C for 20 min, followed by quenching with 1 M HCl and filtering. After filtration, the complex of Cu(L-Tyr)$_2$ was dissolved in 75% aqueous DMF and 1 eq. K_2CO_3 and 0.75 eq. *o*-nitrobenzyl bromide (*o*-NBB) were added to the solution. The mixture was stirred for 2 days at RT and resulted in a light-blue suspension. After vacuum filtration, the complex was rinsed with 75% aqueous DMF followed by water, and acetone. Finally, the resulting solid was stirred with 1 M HCl for 1 h and then filtered and rinsed again with 1 M HCl, followed by water, and finally acetone, providing 730 mg of ONBY (yield 55 %) as a pale yellow solid. High resolution ESI-MS analysis: calc. mass for $C_{16}H_{16}N_2O_5$ $[M+H]^+$ 317.1132, found 317.1131 $[M+H]^+$.

Synthesis of 4-azidobut-2-enyl-*S*-adenosyl-L-methionine (AbSAM, 2c)



2b **3** **2c**

The synthesis was performed as previously described[2]. For this, *S*-adenosyl-L-homocysteine (SAH**, 2b**) (20 mg, 52 µmol) was solved in 10 mL of a freshly prepared mixture of formic and acetic acids (1:1) on ice. (*E*)-1-azido-4-bromobut-2-ene **3**, (180 mg, 1 mmol) and $AgClO_4$ (5.4 mg, 0.031 mM) were added to the reaction mixture. The reaction was allowed to warm to room temperature and stirred for 24 h at RT. After 24 h the reaction was quenched by adding 15 mL ddH_2O, the aqueous phase was washed three times with diethyl ether (3 × 5 mL) and subsequently lyophilised overnight. Subsequently, the crude product was purified by reversed-phase HPLC using a Nucleodur® C18Pyramid (5 µm, 125x10 mm, *Macherey Nagel*) column with a gradient of eluent A: 0.01% formic acid in water; eluent B: 0.01% formic acid in acetonitrile as mobile phase. The obtained diastereomeric mixture was lyophilized and dissolved in water + 0.01% TFA. Concentrations were determined by UV absorption with ε_{260}= 15,400 L mol^{-1} cm^{-1}. High resolution ESI-MS analysis: calc. mass for $C_{18}H_{26}N_9O_5S^+$: 480.1777 ($[M]^+$) found 480.1770 $[M]^+$

Cloning and Protein production

The amber stop codon TAG was introduced into the gene coding for Ecm1 in a pET28a expression vector using site directed mutagenesis. Positions coding for Y145, Y212 and Y284 in Ecm1 mutated. The plasmid coding for tRNA synthetase pEVOL-ONBY was kindly provided by Peter G. Schultz. For protein production, *E. coli* K12 UT5600 cells[3] were electroporated (1.8 kV, 5.4-5.8 ms) to introduce the respective plasmids coding for the tRNA synthetase, as well as Ecm1 ONBY variants. Cells were grown in 2xYT media at 37 °C at 180 rpm until an OD_{600} of 0.6 was reached. Subsequently, ONBY in a 1 M NaOH solution was added to the bacterial culture to a final concentration of 1 mM. Cells were incubated for another 15 min at 180 rpm before the expression was induced by addition of 0.4 mM IPTG and 0.2% arabinose. Subsequently, the cells were grown at 30 °C for 4 h and afterwards harvested by centrifugation (30 min, 4,000 x g, 4 °C).The pellets were frozen in liquid nitrogen and stored at –20 °C until further usage.

Protein Purification

The cell pellet was thawed on ice and resuspended in 10 mL lysis buffer (20 mM HEPES, 150 mM NaCl, pH 7.4). Afterwards, the cells were lysed by sonication and the supernatant was cleared by centrifugation (30 min, 22,000 x g, 4 °C). Next, the His-tagged protein was purified via IMAC using a 1 mL HisTrap (*GE Healthcare*) column and lysis buffer containing 500 mM imidazole for elution. In order to obtain RNase-free proteins, an additional size exclusion chromatography step was added using Superdex 200 Increase 10/300 GL column (*GE Healthcare*) with gel filtration buffer (20 mM HEPES pH 7.4, 150 mM NaCl, 5 mM DTT, 10% glycerol). Protein elution was monitored by absorption at 280 nm and collected fractions were analyzed

by SDS-PAA gel electrophoresis (10% PAA, 200 V, 50 min). Fractions containing the desired protein were checked for RNase contamination. RNase-free fractions were pooled and concentrated using Amicon Ultra-4 centrifugal filters (MWCO 10 kDa), before the protein purity and concentration was determined by comparison with BSA standard on SDS-PAA gel.

Protein Mass Spectrometry

Proteins were reduced with 2 mM TCEP on ice for 15 minutes. Then, samples were acidified using a 5% formic acid solution and centrifuged (10 min, 21000 x g, 4 °C). Depending on the protein concentration, an appropriate volume of the supernatant was loaded on a C4 column (Advance Bio RP-mAb C4, 2.1 mm x 50 mm, 3.5 µm, *Agilent Technologies*) at a flow rate of 0.6 mL/min in 5% eluent B (eluent A: 0.1% formic acid in water; eluent B: 0.1% formic acid in acetonitrile). After a desalting period of 7 minutes at 5% B, a steep gradient was applied (5-60% B in 2 min). MS settings: capillary voltage 4500 V, end-plate offset 500 V, nebulizer 5.0 bar, dry gas 9.0 L/min, dry T=200 °C, mass range m/z 300-3000. Data was analyzed with DataAnalysis 4.4 (*Bruker Daltonik GmbH*) and deconvolution was performed using the MaxEnt algorithm implemented in the software.

Western Blot

Proteins were separated via SDS-PAA gel electrophoresis (10% PA, 200 V, 60 min) and transferred onto a nitrocellulose membrane Roti®-NC (*Roth*) in semi-dry transfer buffer with 80 mA for 60 min at rt. Before antibody treatment, the membrane was washed with 1x PBS with 0.1% Tween (PBST) and blocked in blocking buffer (PBS with 3% BSA) for 1 h at rt. Incubation with the 6x anti-His primary antibody was performed during shaking overnight at 4 °C. The membrane was washed three times with PBST (each 5 min at rt) and then incubated with polyclonal rabbit anti-Mouse/HRP-conjugated secondary antibody for 1 h at rt, followed by repeated washing (3x PBST, 5 min, at rt). For detection, the EZ-ECL Chemiluminescence detection kit (*Biological Industries*) was used and the results were analyzed with a Chemo Star Advanced Fluorescence & ECL Imager (*Intas*).

Activity test and illumination assay

Enzymatic modification of GpppA (0.5 mM) using AdoMet (1 mM) with 25 µM Ecm1 was performed in the presence of 4 µM MTAN and LuxS in 20 mM HEPES, 150 mM NaCl, pH 7.4 at 37 °C as described previously.[4] Samples were taken at indicated time points and proteins precipitated by adding 1/10 volume 1 M $HClO_4$ and centrifugation (10 min, 21000 x g, 4 °C). For light-induced activation, the samples were irradiated for 1 min at 365 nm (3 W, λ_{max} = 365 nm, *LED Engin*) using a custom made LED box. The samples were analyzed by UHPLC using a reversed-phase column (Nucleodur C18 Pyramid, 100x2 mm, 1.8 µm, *Macherey Nagel*). Samples were separated with a flow rate of 0.4 mL/min and a linear gradient from 20 mM NH_4HCO_3 (pH 6.0) to MeOH.

Kinetic characterization

Kinetic parameters V_{max} and K_M were determined as previously described.[5] Enzymatic modification of GpppA (0.025 – 1 mM) with 25 µM Ecm1, 0.5 mM AdoMet in the presence of 4 µM MTAN and LuxS in 20 mM Hepes, 150 mM NaCl, pH 7.4 at 37 °C for 1 h. The reaction was stopped by adding 1/10volume 1 M $HClO_4$. Conversions were analyzed by reversed-phase HPLC as described above and plotted against the concentration of the respective GpppA. The Michaelis-Menten equation in Graphpad Prism was used for fitting.

***In vitro* transcription**

First, the genes coding for eGFP, *Firefly luciferase* (FLuc) or *Renilla luciferase* (RLuc) mRNAs were amplified from pMRNA-vectors containing the respective sequence.[7, 8] After purification (NucleoSpin Gel and PCR Clean-up, *Macherey-Nage*l), the resulting linear dsDNA was used as template (200 ng) for *in vitro* T7 transcription performed in transcription buffer (40 mM Tris/HCl, 25 mM NaCl, 8 mM $MgCl_2$, 2 mM spermidine) by adding A/C/UTP mix (0.5 mM), GTP (0.25 mM), GpppG/m^7GpppG/ApppG or ARCA cap analog (1 mM), T7 RNA polymerase (50 U) (*Thermo Scientific*) and pyrophosphatase (0.1 U) (*Thermo Scientific*) for 4 h at 37 °C. Remaining DNA template was digested by addition of 2 U DNase I for 1 h at 37 °C and the mRNA was purified using the RNA Clean & Concentrator™-5 Kit (*Zymo Research*). To digest non-capped RNA, 10 U of the RNA 5'-polyphosphatase (*Epicentre*) as well as the supplied reaction buffer were added to purified mRNA. After an incubation period of 30 min at 37 °C, 5′–3′ exoribonuclease XRN1 (*NEB*) (0.5 U) and $MgCl_2$ (5 mM) were added. The reaction mixture was incubated for 60 min at 37 °C. Subsequently, the capped mRNA was purified using the RNA Clean & Concentrator™ -5 Kit (*Zymo Research*).

Enzymatic modification of GpppG-RNAs by Ecm1 and digestion

First, 500 ng GpppG-FLuc-mRNA was incubated with 0.2 µM wild type Ecm1 or denatured Ecm1, in presence of 0.6 µM SAM or 1 µM AbSAM in 20 mM HEPES, 150 mM NaCl, pH 7.4 for 60 min at 37 °C. Subsequently, modified mRNA was extracted by Phenol:Chloroform (5:1) and precipitated in EtOH. The resulting pellet was dissolved in ddH_2O and used for *in vitro* translation or click reaction. For quantification of the modified cap, the RNA was digested with 0.15 U nuclease P1 in 20 mM NH_4OAc, 0.1 mM $ZnCl_2$, pH 5.3 for 2 h at 37 °C followed by dephosphorylation using 1 U FastAP (*Thermo Scientific*) for 1 h at 37 °C. Subsequently, proteins were precipitated by adding 1/10 volume 1 M $HClO_4$ and centrifugation (10 min, 21000 x g, 4 °C)

Quantification of Cap nucleotides

LC-QqQ-MS analysis and quantification of cap nucleotides were performed on an Agilent 1260 Infinity II HPLC system (*Agilent Technologies*) using a Poroshell 120EC-C18 (150x3 mm, 2.7 µM, 20 °C, *Agilent Technologies*) column coupled to an Agilent Ultivo triple-quad mass spectrometer. Samples were separated at 0.6 mL/min using gradient from 0-30% eluent B (eluent A: 20 mM NH_4OAc (pH = 6.0), eluent B: CH_3CN). MS settings: gas temperature: 250 °C, gas flow: 7.0 L/min, nebulizer 40 psi, sheath gas temperature 375 °C, sheath gas flow 12 L/min, capillary voltage 9 V. For quantification, standard curves were prepared using standard mixes of GpppG and m^7GpppG in different concentrations (0.1 or 1 µM/) and subsequently used for calculation of the relative abundance of different cap using the software Agilent Mass Hunter Quantitative Analysis (for QqQ). For both cap nucleotides mass at 248.1 m/z was used as quantifier with two additional qualifiers.

Table.2: Setting mass spectrometer for cap quantification

Compound	Pecursor (m/z)	Product (m/z)	Fragmentor (V)	CAV(V)	CE (V)	Mode
GpppG	789.1	134.7	150	9	121	Positive
GpppG	789.1	151.8	150	9	57	Positive
GpppG	789.1	248.4	150	9	21	Positive
m^7GpppG	803.1	151.7	135	9	61	Positive
m^7GpppG	803.1	165.7	135	9	53	Positive
m^7GpppG	803.1	248.1	150	9	21	Positive

Strain-promoted azide-alkyne cycloaddition (SPAAC) with modified RNA

The click reaction was performed with enzymatically modified azidobutenyl-modified mRNA in presence of 10 µM DBCO-SRB (*Jena Bioscience)* for 1 h at 37 °C followed by precipitation in EtOH. Subsequently, the click-product was separated on a 7.5% TBE-PAA gel and visualized on a Typhoon Gel imaging system (*GE Healthcare*).

***In vitro* translation and Luminescence measurement**

For in vitro translation the Retic Lysate IVT™ kit (*Invitrogen*), a eukaryotic cell-free protein expression system, was used. Experiments were performed in a total volume of 15 µL and contained 10 ng ARCA-RLuc-mRNA, 40 ng of the indicated FLuc-mRNAs, 50 µM L-methionine and 150 mM potassium acetate. Samples were mixed with 8.5 µL of the *Reticulocyte* lysate and incubated for 90 min at 30 °C. Afterwards, 2 µL of the respective translation mix were used in a luminescence assay. The translation efficiencies of the reporter mRNAs were assessed in a dual luciferase assay using Beetle-juice Luciferase Assay Firefly (*pjk*) and Gaussia-Juice Luciferase Assay Kit (*pjk*). Luciferase activity was determined after adding 50 µL freshly prepared substrate solution to the translation mixture. Luminescence was monitored using a Tecan infinite® m1000pro microplate reader with an acquisition time of 3 s (duration of signal acquisition).

Supplementary Figures

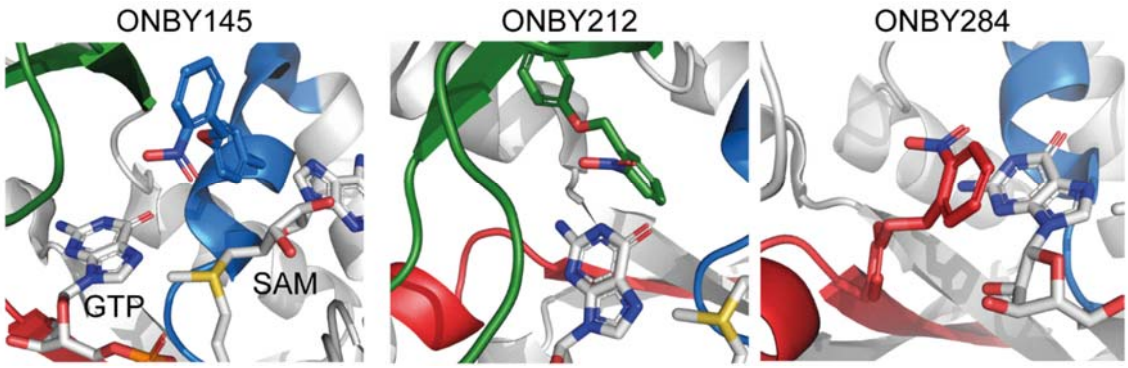
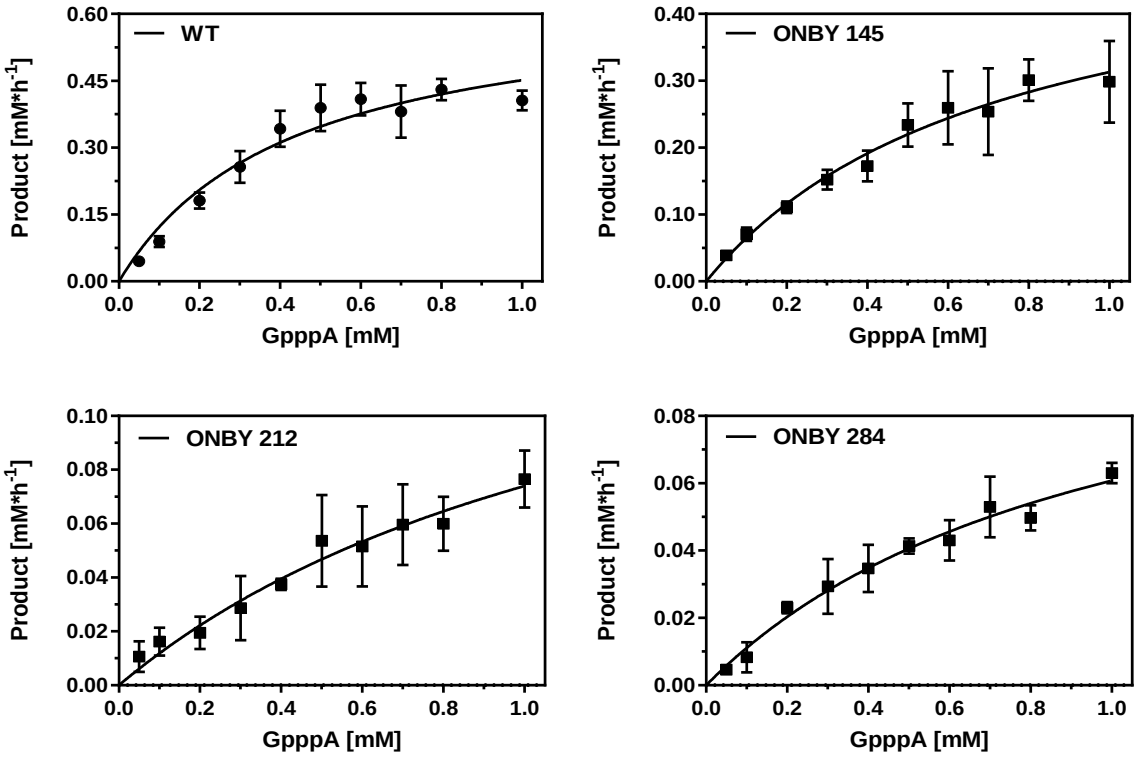


Figure S1: Model of ONBY substitutions for Y145, Y212 or Y284, respectively based on crystal structure of Ecm1 (PDB 1RI2) with GTP and SAM, shown in sticks. The structure was edited with the PyMol Builder tool and the ONB group introduced at the respective positions.



	WT	ONBY 145	ONBY 212	ONBY 284
V_{max} [mM*h^{-1}]	0.644 ± 0.079	0.553 ± 0.060	0.131 ± 0.024	0.121 ± 0.017
K_M [mM]	0.43 ± 0.12	0.77 ± 0.15	0.98 ± 0.29	0.99 ± 0.23
k_{cat} [h^{-1}]	25.8	22.1	5.3	4.8
k_{cat}/K_M [(h*mM)$^{-1}$]	60.3	28.6	5.4	4.9
rel. activity [%]	100	47	9	8

Figure S2: Kinetic characterization of Ecm1 ONBY variants**.** Product formation was monitored by HPLC and plotted against substrate concentration. Non-linear fit was performed using the Michaelis-Menten fit model in Graphpad prism. Error bars depicted show SD from independent experiments (n=3).

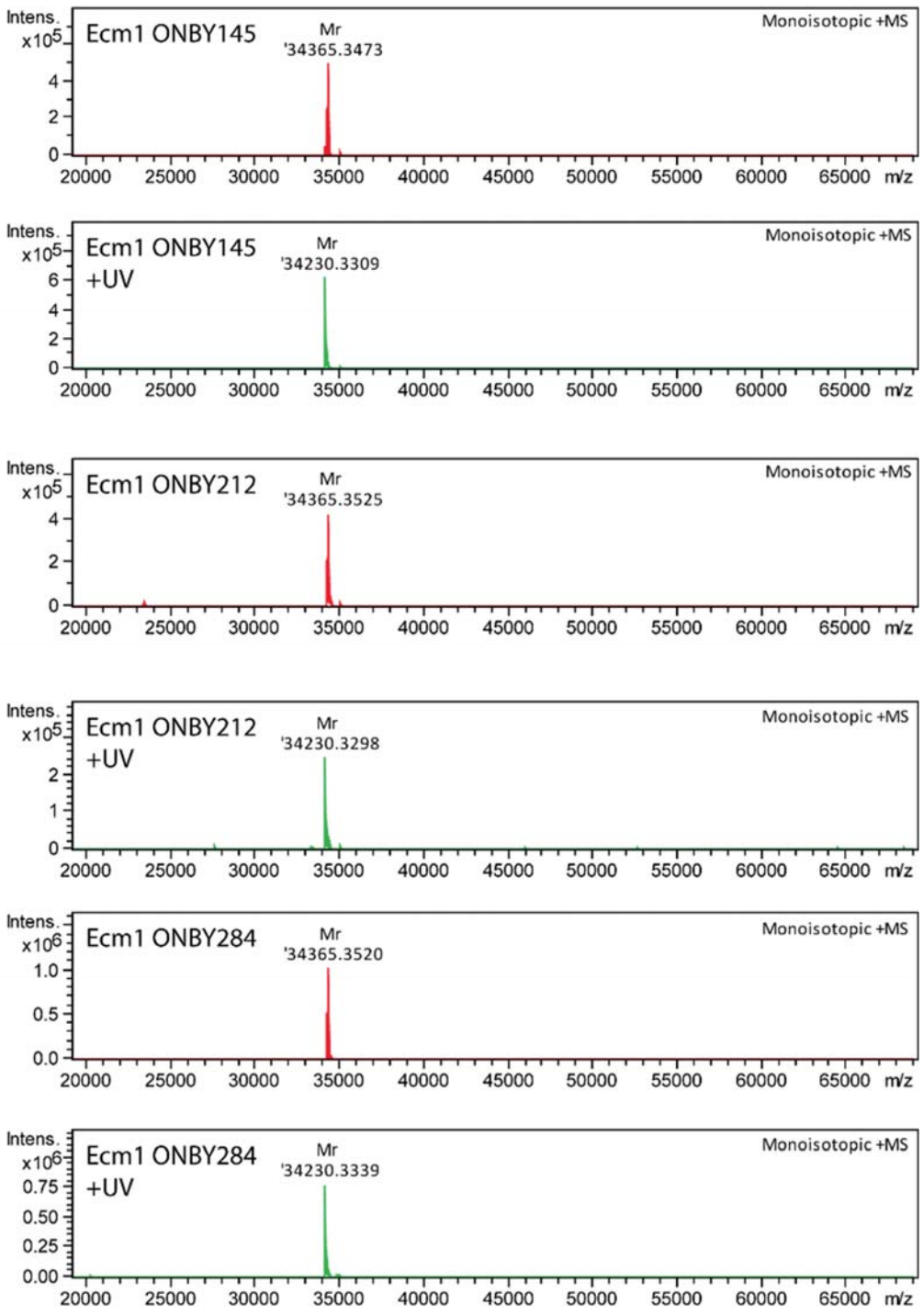


Figure S3: LC/MS analysis of purified Ecm1 ONBY variants. Monoisotopic mass of Ecm1 variants ONBY145, ONBY212 and ONBY284. $M_{calc.}$ 34365.3618 are shown before irradiation (red) and after irradiation for 1 min at 365 nm (green). Monoisotopic mass of wild type enzyme $M_{calc.}$ 34230.3298

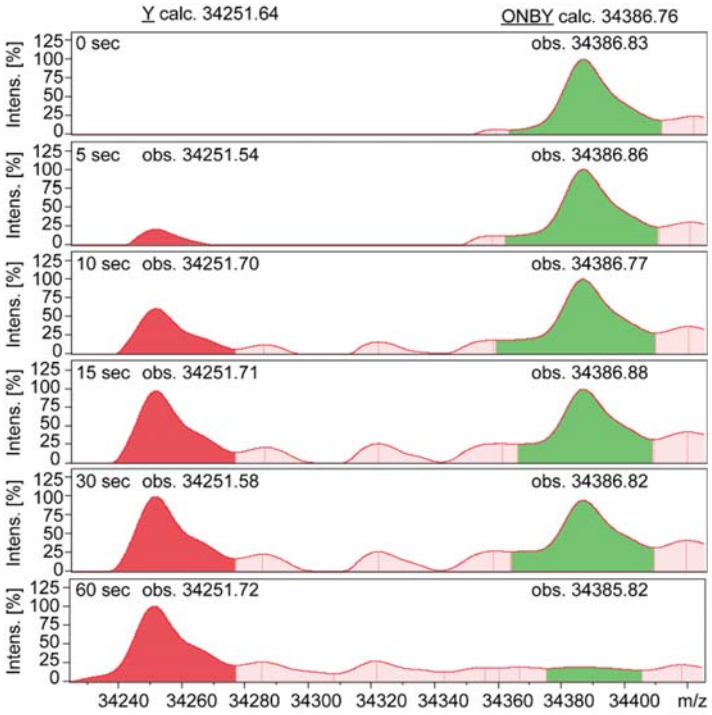


Figure S4: Quantification of photo-deprotection as determined for ONBY212. Individual samples were irradiated for 0–60 s and subsequently measured on LC/MS. The peak area containing the average mass of caged ONBY variant (green) and decaged tyrosine variant (red) was determined and used to calculate the deprotection yield as the area ratio of uncaged/caged protein. Average mass of Ecm1 ONBY212 $M_{calc.}$ 34386.76 and wild type $M_{calc.}$ 34251.64.

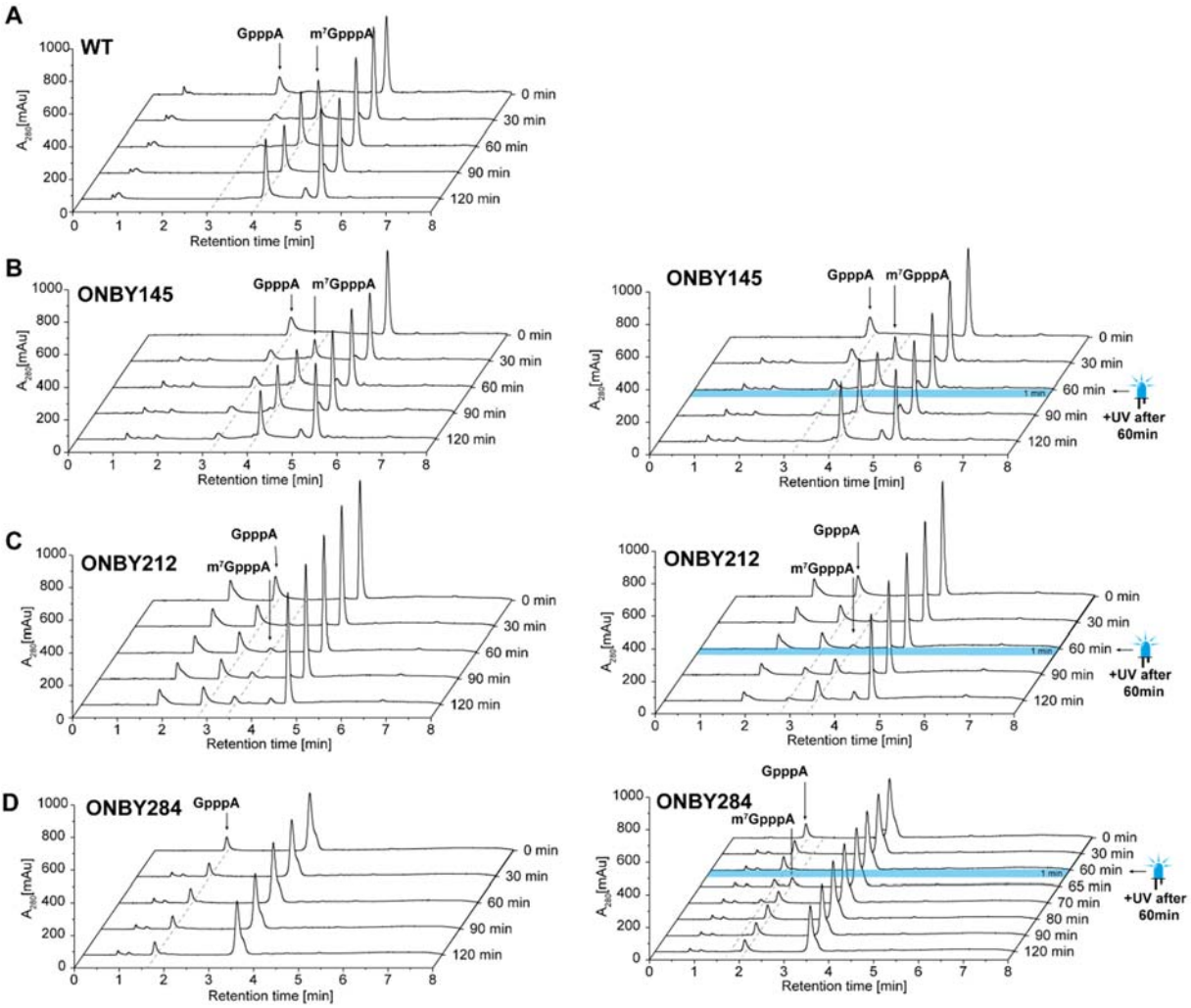


Figure S5: HPLC analysis of light-induced cap methylation catalyzed by wild type Ecm1 (**A**) and respective variants ONBY145 (**B**), ONBY212 (**C**) and ONBY284 (**D**) using the same reaction conditions. The wild type enzyme showed full conversion of 5′ cap analog GpppA after 60 min, where only product m^7GpppA was detected in the reaction mixture. For the respective ONBY variants the independent reactions were kept in the dark (-UV) or were irradiated after 60 min reaction time (+UV, indicated by LED).

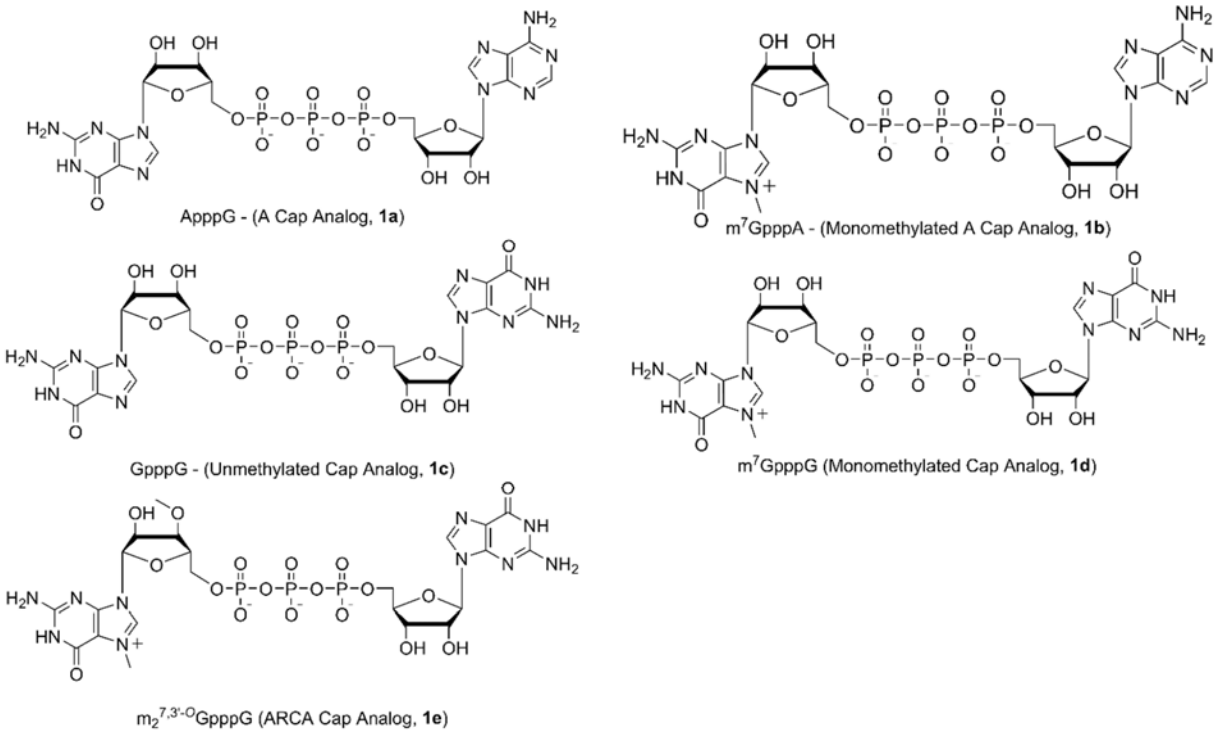


Figure S6: mRNA 5′ cap analogs used for *in vitro* transcription and biotransformation.

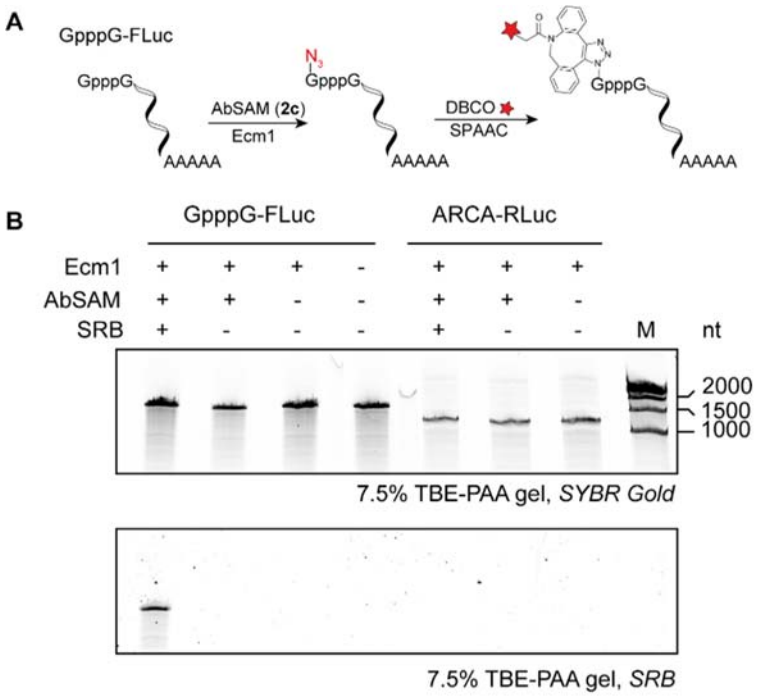


Figure S7: A) Scheme illustrating Ecm1 catalyzed cap modification of GpppG -FLuc mRNA. Azide modified reporter can be labeled through a SPAAC reaction with DBCO-SRB **B)** Luciferase mRNA (FLuc, RLuc) were analyzed by dPAGE (7.5% PAA gel) and visualized using SRB fluorescence (λex= 532 nm) and SYBR Gold staining

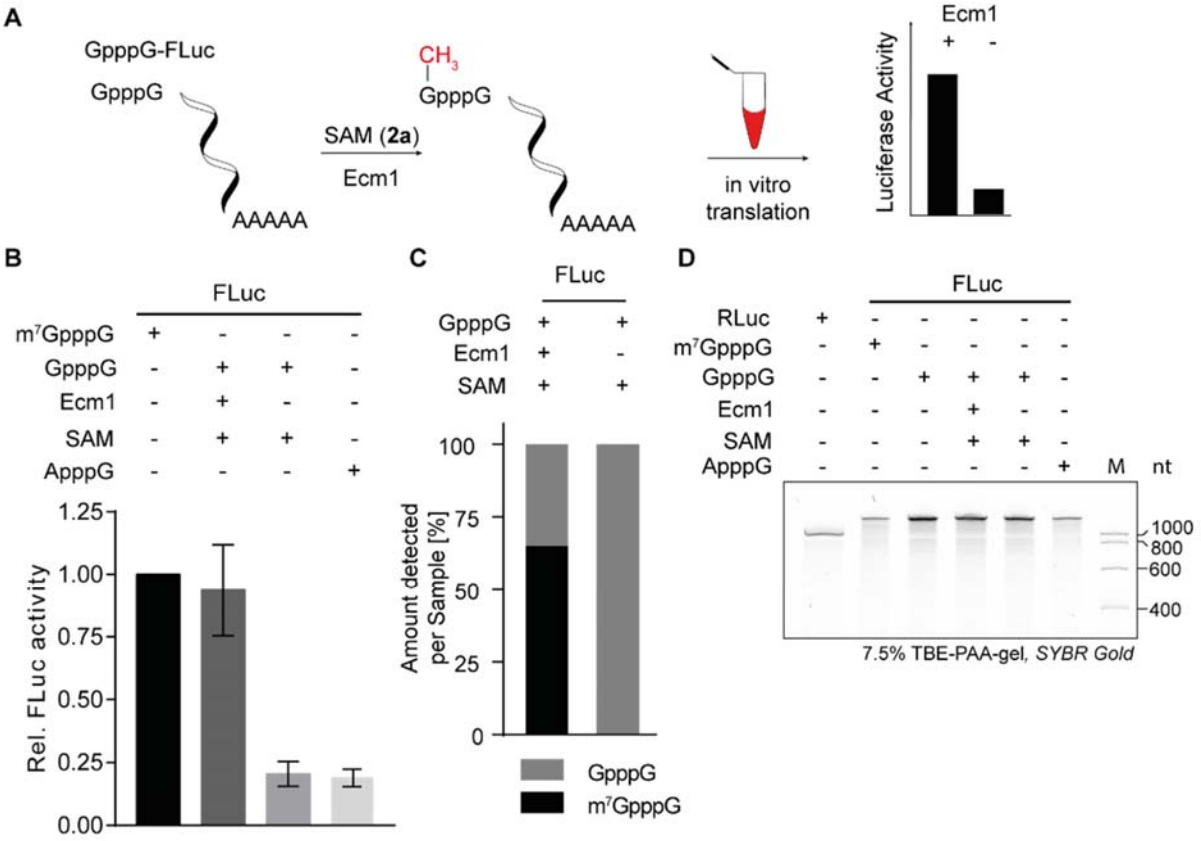


Figure S8: A) Scheme illustrating cap methylation of GpppG -FLuc mRNA catalyzed by Ecm1. **B)** Translation efficiency of differently capped FLuc reporter RNA were assessed by luminescence measurements using the Beetle Firefly Luciferase Assay Kit. FLuc luminescence was referred to ARCA-RLuc signal and normalized to m^7GpppG-FLuc (100%). Error bars depicted show SEM from five independent experiments. **C)** LC-QQQ analysis of caps modified with Ecm1 or denatured Ecm1. After enzymatic modification RNA was P1 digested to nucleosides and the amount of cap dinucleotides quantified. **D)** mRNA was analyzed on a 7.5 % TBE gel to confirm integrity before application in *in vitro* translation.

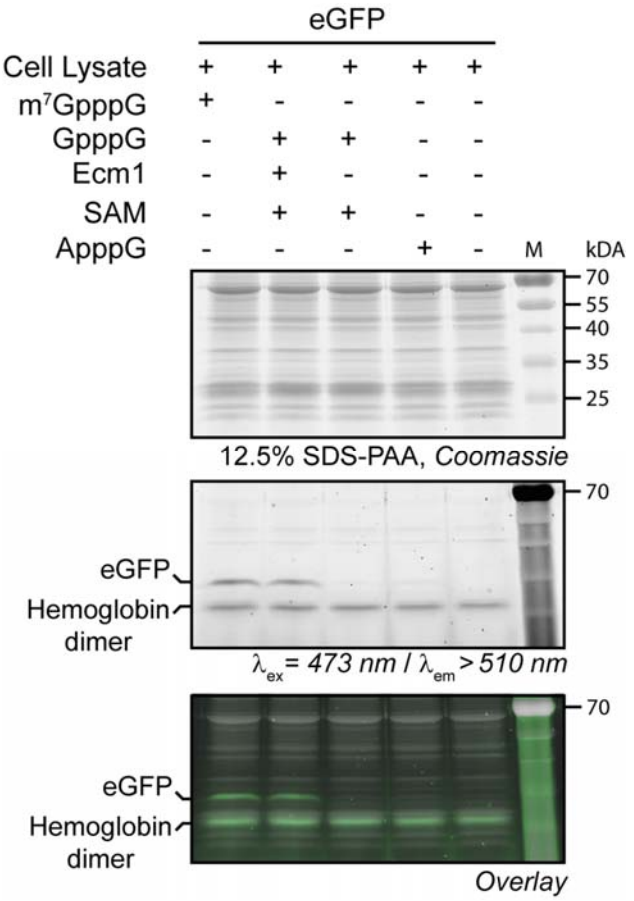


Figure S9: SDS-PAA gel and fluorescence scan of eGPF produced by *in vitro* translation. Differently capped eGFP-reporter mRNA were added to the translation mix and reacted for 90 min at 30 °C, then separated on 12.5 % SDS-PAA gel. In gel fluorescence signal was visualized on a Typhoon Gel imaging system (*GE Healthcare*).

Amino acid sequences of produced Ecm1 ONBY variants

A) **ONBY145:** MEGKKEEIREHYNSIRERGRESRQRSKTINIRNANNFIKACLIRLYTKRGDSVLDLGC
GKGGDLLKYERAGIGEYYGVDIAEVSINDARVRARNMKRRFKVFFRAQDSYGRHMDLGKEFDVI
SSQFSFH**Y**AFSTSESLDIAQRNIARHLRPGGYFIMTVPSRDVILERYKQGRMSNDFYKIELEKMED
VPMESVREYRFTLLDSVNNCIEYFVDFTRMVDGFKRLGLSLVERKGFIDFYEDEGRRNPELSKK
MGLGCLTREESEVVGIYEVVVFRKLVPESDALLEHHHHHH

B) **ONBY212:** MEGKKEEIREHYNSIRERGRESRQRSKTINIRNANNFIKACLIRLYTKRGDSVLDLGC
GKGGDLLKYERAGIGEYYGVDIAEVSINDARVRARNMKRRFKVFFRAQDSYGRHMDLGKEFDVI
SSQFSFHYAFSTSESLDIAQRNIARHLRPGGYFIMTVPSRDVILERYKQGRMSNDFYKIELEKMED
VPMESVRE**Y**RFTLLDSVNNCIEYFVDFTRMVDGFKRLGLSLVERKGFIDFYEDEGRRNPELSKK
MGLGCLTREESEVVGIYEVVVFRKLVPESDALLEHHHHHH

C) **ONBY284:** MEGKKEEIREHYNSIRERGRESRQRSKTINIRNANNFIKACLIRLYTKRGDSVLDLGC
GKGGDLLKYERAGIGEYYGVDIAEVSINDARVRARNMKRRFKVFFRAQDSYGRHMDLGKEFDVI
SSQFSFHYAFSTSESLDIAQRNIARHLRPGGYFIMTVPSRDVILERYKQGRMSNDFYKIELEKMED
VPMESVREYRFTLLDSVNNCIEYFVDFTRMVDGFKRLGLSLVERKGFIDFYEDEGRRNPELSKK
MGLGCLTREESEVVGI**Y**EVVVFRKLVPESDALLEHHHHHH

References

1. J. C. Miller, S. K. Silverman, P. M. England, D. A. Dougherty and H. A. Lester, *Neuron*, 1998, **20**, 619-624.
2. K. Islam, I. Bothwell, Y. Chen, C. Sengelaub, R. Wang, H. Deng and M. Luo, *J. Am. Chem. Soc.*, 2012, **134**, 5909-5915.
3. J. K. Bocker, W. Dorner and H. D. Mootz, *Chem. Commun.*, 2019, **55**, 1287-1290.
4. J. Leiter, D. Reichert, A. Rentmeister and R. Micura, *ChemBioChem*, 2020, **21**, 265-271.
5. S. Hausmann, S. Zheng, C. Fabrega, S. W. Schneller, C. D. Lima and S. Shuman, *J. Biol. Chem.*, 2005, **280**, 20404-20412.
6. G. V. Semisotnov, N. A. Rodionova, O. I. Razgulyaev, V. N. Uversky, A. F. Gripas and R. I. Gilmanshin, *Biopolymers*, 1991, **31**, 119-128.
7. L. Anhauser, S. Huwel, T. Zobel and A. Rentmeister, *Nucleic Acids Res.*, 2019, **47**.
8. J. M. Holstein, L. Anhauser and A. Rentmeister, *Angew. Chem. Int. Ed.*, 2016, **55**, 10899-10903.