Electronic Supplementary Material (ESI) for RSC Chemical Biology.

Computational design and experimental characterization of a photo-controlled mRNA-cap guanine-N7 methyltransferase

Dennis Reichert[a,b], Helena Schepers[a], Julian Simke[c], Horst Lechner[d,e], Wolfgang Dörner[a], Birte Höcker[d], Bart Jan Ravoo[b,c], Andrea Rentmeister[a,b,*]

[a] Institute of Biochemistry, University of Münster, Correnstraße 36, 48149 Münster Germany

[b] Cells in Motion Interfaculty Center, University of Münster, 48149 Münster Germany

[c] Organic Chemistry Institute, University of Muenster, Corrensstraße 36, 48149 Münster, Germany

[d] Department for Biochemistry, University of Bayreuth, Universitätsstraße 30, 95447 Bayreuth, Germany.

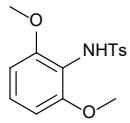
[e] Institute of Biochemistry Graz University of Technology, Petersgasse 10-12/II, 8010 Graz, Austria

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

Synthesis

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 was 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 was performed on an Agilent Infinity 1260 system (*Agilent Technologies*). ^{1}H/^{13}C-NMR spectra were recorded at 298 K on a Bruker NEO400 (400 MHz) or Agilent DD2 600 (600 MHz) spectrometer. Chemical shifts δ (in ppm) were given relative to undeuterated residues of the used deuterated solvents as internal standard. Multiplicities are indicated as s (singlet), d (doublet), t (triplet), q (quartet) and m (multiplet).

Synthesis of *N*-(3,5-dimethoxy phenyl)-4-methylbenzenesulfonamide (5)



2,6-Dimethoxyaniline (2.50 g, 16.3 mmol, 1.0 eq.) and TsCl (3.41 g, 17.9 mmol, 1.1 eq.) were dissolved in pyridine (7.00 mL), and the mixture was stirred for 1 h at 100 °C. After cooling to rt, the solution was diluted with Milli-Q water and kept in the fridge overnight. The solid was filtered and recrystallized from aq. EtOH. **5** was obtained as grey/silver needles.

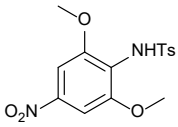
Yield: 4.62 g (15.0 mmol, 92%).

^{1}H-NMR: (Chloroform-*d*, 400 MHz) δ = 7.80-7.63 (m, 2H), 7.36-7.19 (m, 2H), 7.11 (t, *J* = 8.4 Hz, 1H), 6.48 (d, *J* = 8.5 Hz, 2H), 6.26 (s, 1H), 3.59 (s, 6H), 2.41 (s, 3H) ppm.

^{13}C-NMR: (Chloroform-*d*, 101 MHz) δ = 155.1, 143.0, 138.4, 128.9, 127.7, 127.5, 114.5, 104.4, 55.8, 21.6 ppm.

MS (*m/z*): (ESI, MeOH) Calculated for $[C_{15}H_{17}NO_4SNa]^+$: 330.0770; found 330.0792.

Synthesis of *N*-(3,5-dimethoxy-4-nitrophenyl)-4-methylbenzenesulfonamide (6)



5 (704 mg, 2.29 mmol, 1.0 eq.) was dissolved in a solution of nitric acid (0.645 mL) in H_2O (5.12 mL), and then acetic acid (5.12 mL) and $NaNO_2$ (17.9 mg, 0.259 mmol, 0.1 eq) were added subsequently. The mixture was stirred for 1 h at 100 °C. After precipitation with H_2O (10.0 mL), the crystals were filtered and recrystallized from aq. EtOH. The desired compound **6** was obtained as yellow crystals.

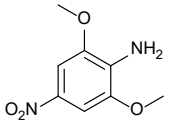
Yield: 576 mg (1.64 mmol, 72%).

^{1}H-NMR: (Chloroform-*d*, 400 MHz) δ = 7.92-7.70 (m, 2H), 7.43 (s, 2H), 7.36-7.28 (m, 2H), 6.72 (s, 1H), 3.73 (s, 6H), 2.44 (s, 3H) ppm.

^{13}C-NMR: (Chloroform-*d*, 101 MHz) δ = 153.1, 146.0, 143.6, 138.5, 129.2, 127.2, 121.7, 100.5, 56.4, 21.7 ppm.

MS (*m*/*z*): (ESI, MeOH) Calculated for $[C_{15}H_{16}N_2O_6SNa]^+$: 375.0627; found 375.0621.

Synthesis of 3,5-dimethoxy-4-nitroaniline (7)



6 (7.97 g, 22.6 mmol, 1.0 eq.) was dissolved in 23.9 mL of a mixture of sulfuric acid/H_2O (10:1) and stirred at rt overnight. After that, H_2O (200 mL) was added and alkalized by adding conc. Ammonia. The resultant solid was filtered and recrystallized from aq. EtOH. The desired compound **7** was obtained as yellow-orange needles.

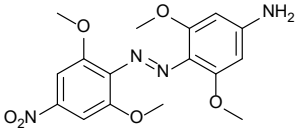
Yield: 4.24 g (21.4 mmol, 95%).

^{1}H-NMR: (Chloroform-*d*, 400 MHz) δ = 7.51 (s, 2H), 4.55 (br. s, 2H), 3.92 (s, 6H) ppm.

^{13}C-NMR: (Chloroform-*d*, 101 MHz) δ = 145.0, 137.5, 133.5, 101.2, 56.3 ppm.

MS (*m*/*z*): (ESI, MeOH) Calculated for $[C_8H_{10}N_2O_4Na]^+$: 221.0538; found 221.0529.

Synthesis of (*E*)-4-((2,6-dimethoxy-4-nitrophenyl)diazenyl)-3,5-dimethoxyaniline (8a)



7 (4.24 g, 21.4 mmol, 1.0 eq.) was dissolved in H_2O (4.10 mL) and HCl (37%, 5.30 mL) and cooled to 0 °C. $NaNO_2$ (1.48 g, 21.4 mmol, 1.0 eq.) in H_2O (8.60 mL) was added dropwise, and the mixture was stirred for 20 min at 0 °C. This diazonium salt was then transferred to a suspension of 3,5-dimethoxyaniline (3.28 g, 21.4 mmol, 1.0 eq.) in H_2O (137 mL) at 0 °C. The pH was adjusted to pH 8-9 by adding a saturated $NaHCO_3$ solution. The mixture was stirred for 18 h at rt. After filtration, the desired compound **8a** was purified *via* column chromatography (silica, DCM → DCM/MeOH 9:1) and obtained as a red solid.

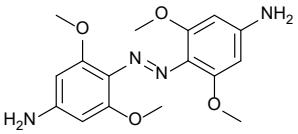
Yield: 2.58 g (7.12 mmol, 30%).

^{1}H-NMR: (Methanol-*d*4, 300 MHz) δ = 7.62 (s, 2H), 6.03 (s, 2H), 3.88 (s, 6H), 3.84 (s, 6H) ppm.

^{13}C-NMR: (Methanol-*d*4, 126 MHz) δ = 162.5, 159.0, 153.1, 147.7, 125.1, 101.9, 91.0, 57.1, 56.2 ppm.

MS (*m*/*z*): (ESI, MeOH) Calculated for $[C_{16}H_{19}N_4O_6H]^+$: 363.1299; found 363.1301.

Synthesis of (*E*)-4,4'-(diazene-1,2-diyl)bis(3,5-dimethoxyaniline) (8b)



8a (256 mg, 0.707 mmol, 1.0 eq.) was dissolved in a mixture of THF/H_2O (3:1, 28 mL) and $Na_2S{\times}9H_2O$ (509 mg, 2.12 mmol, 3.0 eq.) was added. The solution was stirred for 17 h at 80 °C. The solvent was removed under reduced pressure, and the residue was dissolved in EtOAc. The organic phase was washed with H_2O (20.0 mL) and brine (2×20.0 mL) and dried over $MgSO_4$. The solvent was evaporated, and the crude product was purified by column chromatography (silica, EtOAc/MeOH/NEt_3 100:20:5), yielding **8b** as a purple solid.

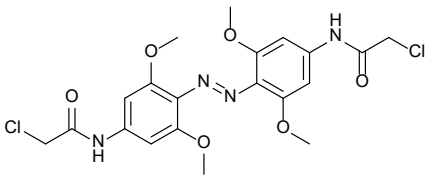
Yield: 102 mg (0.307 mmol, 43%).

^{1}H-NMR: (DMSO-*d*6, 500 MHz) δ = 5.96 (s, 4H), 3.68 (s, 12H) ppm.

^{13}C-NMR: (DMSO-*d*6, 126 MHz) δ = 154.8, 151.3, 128.7, 90.9, 55.9 ppm.

MS (*m*/*z*): (ESI, MeOH) Calculated for $[C_{16}H_{21}N_4O_4H]^+$: 333.1563; found 333.1570.

Synthesis of (*E*)-*N,N'*-(diazene-1,2-diylbis(3,5-dimethoxy-4,1-phenylene))bis(2-chloroacetamide) (3)



8b (15.2 mg, 45.7 µmol, 1.0 eq.) was dissolved in a mixture of $CHCl_3$ (1.50 mL) and NEt_3 (31.7 µL, 229 µmol, 5.0 eq.) and cooled to 0 °C. Chloroacetylchloride (18.2 µL, 229 µmol, 5.0 eq.) was added dropwise, and the solution was stirred for 1 d at rt. Then, EtOAc (15.0 mL) was added, and the mixture was washed with H_2O (15.0 mL) and extracted with EtOAc (3×15.0 mL). The organic layers were combined, dried over $MgSO_4$, and the solvent was removed under reduced pressure. After purification *via* column chromatography (silica, EtOAc), **3** was obtained as a light reddish-(brown) solid.

Yield: 8.97 mg (18.5 µmol, 41%).

^{1}H-NMR: (DMSO-*d*6, 500 MHz) δ = 10.47 (s, 2H), 7.10 (s, 4H), 4.29 (s, 4H), 3.71 (s, 12H) ppm.

^{13}C-NMR: (DMSO-*d*6, 126 MHz) δ = 165.0, 152.3, 139.9, 130.1, 96.3, 56.0, 43.7 ppm.

MS (*m*/*z*): (ESI, MeOH) Calculated for $[C_{20}H_{22}Cl_2N_4O_6H]^+$: 485.0995; found 485.0981.

Experimental

Cloning and recombinant expression

The gene coding for Ecm1 wild-type enzyme previously cloned into a pRSET-A vector (N-terminal His-Tag)[1] was used as a template for site-directed mutagenesis to obtain all Ecm1 variants (Table S1). For protein production, *E. coli* Tuner cells were grown shaking at 37 °C to an optical density OD_{600} = 0.6. Subsequently, expression was induced by adding 0.4 mM Isopropyl-β-D-thiogalactopyranoside (IPTG), 1 % glucose (wt), and 2% ethanol (v/v) followed by incubation overnight at 18 °C. Next, cells were harvested by centrifugation (30 min, 4000 x *g*, 4 °C), the resulting cell pellet was resuspended in 10 mL binding buffer (50 mM Tris, 150 mM NaCl, 20 mM Imidazole, pH 8.5), and the cells were lyzed by sonification (10 min, 5 s pulse, 10 s pause on ice). Subsequently, cell debris was removed by centrifugation (30 min, 22000 x *g*, 4 °C), and the desired protein isolated by affinity chromatography using 1 mL HisTrap™ FF column with ÄKTApurifier chromatography system (*GE Healthcare*) and 50 mM Tris, 150 mM NaCl, 500 mM Imidazole, pH 8.5 for elution. Fractions containing the desired protein were concentrated and dialyzed against storing buffer (50 mM Tris, 150 mM NaCl, 5 mM DTT, 10 % Glycerol (v/v), pH 8.5). Finally, purity and concentration were determined by comparison with a BSA standard on SDS-PAA gel.

Protein mass spectrometry

First, disulfide bonds were reduced in the presence of 2 mM TCEP incubating on ice for 15 min. 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 3.5 bar, dry gas 8.0 L/min, dry T=200 °C, mass range m/z 300-3000. Data were analyzed with DataAnalysis 4.4 (*Bruker Daltonik GmbH*), and deconvolution was performed using the MaxEnt algorithm implemented in the software.

Activity assay cysteine free variants

Enzymatic modification of GpppA (0.5 mM) using SAM (1 mM) with 5 mol% Ecm1 was performed in the presence of 4 µM MTAN and LuxS in HEPES buffer (20 mM HEPES, 150 mM NaCl, pH 7.4) at 37 °C. Reactions were stopped at specific time-points (0–5 min) by adding 1/10 volume 1 M $HClO_4$ followed by centrifugation (10 min, 21000 x *g*, 4 °C). The conversion of GpppA was analyzed on an Infinity 1260 HPLC system (*Agilent Technologies*) using a C18 reversed-phase column (Nucleodur C18 Pyramid, 125x2mm, 5 µm, *Macherey Nagel*) for separation. Afterwards, the substrate consumption [µM min^{-1}] was calculated by linear regression and divided by the reaction volume and applied protein mass, yielding the specific enzyme activity U_s [µmol min^{-1} mg^{-1}].

Activity assay with designed variants

Enzymatic methylation of GpppA (0.25 mM) with SAM (0.5 mM) as methyl donor by 5 mol% Ecm1 was performed in the presence of 4 µM MTAN and LuxS in 50 mM Tris-HCl, 150 mM NaCl (pH 8.5) at 25 °C. Reactions were stopped at specific time-points (0–23 h) by adding 1/10 volume 1 M $HClO_4$ followed by centrifugation (10 min, 21000 x *g*, 4 °C). All samples were analyzed on an Infinity 1260 HPLC system (*Agilent Technologies*) using a C18 reversed-phase column (Nucleodur C18 Pyramid, 125x2mm, 5 µm, *Macherey Nagel*) for separation. The conversion of GpppA was monitored by integrating the peak areas of methylated and unmethylated cap.

Cysteine accessibility assay

10 µM enzyme were treated with 200 µM TCEP, and modified with 25-fold excess Fluoresceine-5-maleimide (*Thermo Fisher*, 10 mM in DMF) in 50 mM Tris-HCl (pH 7.0) for 2 hours at RT. Subsequently, samples were separated using a 10 % SDS-PAA gel. The fluorescence signal of Fluoresceine conjugated cysteine residues were monitored on a Typhoon FLA9500 gel imaging system (*GE Healthcare*) with a wavelength λ_{ex} =473 nm, λ_{em} ≥ 510 nm.

Modification under native conditions

A 100 µL solution of 45 µM protein in 50 mM Tris-HCl, 150 mM NaCl (pH 8.5) with 10 % glycerol was concentrated in a centrifugal concentrator (*Sartorius*) with a 10 kDa cut off at 15.000 x *g* to a volume of 20 µL. After adding 8 µL of a 2 mM TCEP-solution in 50 mM Tris-HCl, 150 mM NaCl (pH 8.5), the solution was incubated for 10 min at room temperature. Before each addition of **3**, the solution was irradiated for 30 min at 615 nm for obtaining azobenzene **3** in the *cis*-form. Then, the azobenzene solution was added stepwise after 0, 2, 4, and 6 hours, resulting in a final molar ratio of 8 eq. The reaction mixture was incubated at 19 °C and shaking at 300 rpm. After each addition, samples were taken and quenched with 1:1 volume of 5% formic acid. Denatured protein was precipitated by centrifugation (10 min, 21000 x *g*, 4 °C), and the cross-linking yield was determined by mass spectrometry. The residual protein solution is loaded onto a PD Spin-Trap G-25^{TM} (*GE healthcare*) and eluted with 150 µL storage buffer. Finally, the enzyme mixture was concentrated on a centrifugal concentrator (*Sartorius*) to a volume of 40 µL and used for light dependent activity assays

Determination of photostationary states (PSS)

The photostationary state (PSS) for the *E*-*Z*- and the *Z*-*E*-isomerization was determined by ^{1}H-NMR spectroscopy. For this, a solution of the respective azobenzene (500 µM) was first prepared in DMSO-*d*6 and then the initial dark adapted spectrum was measured. To reach PSS_{E-Z} the sample was illuminated for 30 min with red light (λ = 615 nm) and a second spectrum was recorded. For PSS_{Z-E}, the sample was illuminated for a second time (blue light, λ = 450 nm, 45 min) and a last spectrum was measured. In order to determine the ratio between both isomers at the PSS now, either the proton signals in the aromatic region or the methyl groups at the benzene were integrated.

Computational Methods

The computational part of this work was performed according to a protocol previously published[2, 3] using Rosetta suite 3.10.

The tetra-*ortho*-methoxy azobenzene derivative **3** was built in Avogadro[4] using a crystal structure of the Cambridge Crystallographic Database (CCDC)[5] in *cis*-form (AZBENC01) and *trans*-form (AZBENC12) as starting conformation. Next, the average dihedral angles were investigated using the CCDC Mogul software.[5] These were used to generate a conformational library of the respective ligand.

Computational search for suitable cross-linking sites

The crystal structure of Ecm1 (PDB 1RI5) was used to create the cysteine-free mutant Δ4C (Table S1) using the Pymol mutagenesis tool as a starting point for the design. The resulting structure was initially optimized by Rosetta FastRelax[6] with all-heavy-atom constraints. The lowest score structure was then used as a starting point in the subsequent computational search for cross-linking sites. Next, the RosettaMatch[7] protocol was used to find pairs of cysteine substitutions capable of accommodating the azobenzene derivative **3** in the *cis* conformation. A RosettaScripts[8] protocol for design was applied to optimize the binding geometry, repacking, and minimizing nearby sidechains. No sequence changes (apart from the two cysteines) were permitted. Rotamer optimization of sidechains in proximity to the cross-linked azobenzene was allowed in the Rosetta packing simulations. A set of designs with different placements of cysteine residues were chosen for further experimental characterization.

Simulation of *cis*-to-*trans* isomerization

The model with the lowest energy and lowest constraint score with the azobenzene cross-linked in *cis* conformation after the FastRelax protocol was used as the starting conformation for subsequent *trans*-isomerization simulations. During the simulation, azobenzene derivatives were mutated to their *trans* isomers, followed by Rosetta FastRelax application to optimize the geometry and energy of the models. Regions containing the cysteine substitutions were set flexible to enable conformational changes during isomerization.

Ligand constraint files

ZCl constraint file: **@ZCl_const_new.cst**

```
CST::BEGIN
        TEMPLATE::      ATOM_MAP: 1 atom_name: C10 C9 N3
        TEMPLATE::      ATOM_MAP: 1 residue3: ZCl

        TEMPLATE::      ATOM_MAP: 2 atom_name: SG CB CA
        TEMPLATE::      ATOM_MAP: 2 residue3: CYS

        CONSTRAINT::    distanceAB:        1.8  0.10    60.0  1      0
        CONSTRAINT::    angle_A:         109.00  6.00    10.00 360.00 0
        CONSTRAINT::    angle_B:         120.00  3.00    10.00 360.00 0
        CONSTRAINT::    torsion_A:         0.00  180.00  10.00 360.00 18
        CONSTRAINT::    torsion_B:        70.00  20.00   10.00 360.00 2
        CONSTRAINT::    torsion_AB:       70.00  20.00   10.00 120.00 2
CST::END

CST::BEGIN
        TEMPLATE::      ATOM_MAP: 1 atom_name: C18 C17 N4
        TEMPLATE::      ATOM_MAP: 1 residue3: ZCl

        TEMPLATE::      ATOM_MAP: 2 atom_name: SG CB CA
        TEMPLATE::      ATOM_MAP: 2 residue3: CYS

        CONSTRAINT::    distanceAB:        1.8   0.10    60.0  1      0
        CONSTRAINT::    angle_A:         109.00   6.00    10.00 360.00 0
        CONSTRAINT::    angle_B:         120.00   3.00    10.00 360.00 0
        CONSTRAINT::    torsion_A:         0.00   180.00  10.00 360.00 18
        CONSTRAINT::    torsion_B:        70.00   20.00   10.00 360.00 2
        CONSTRAINT::    torsion_AB:       70.00   20.00   10.00 120.00 2
CST::END
```

ECl constraint file: **@ECl_const_new.cst**

```
CST::BEGIN
        TEMPLATE::      ATOM_MAP: 1 atom_name: C18 C17 N3
        TEMPLATE::      ATOM_MAP: 1 residue3: ECl

        TEMPLATE::      ATOM_MAP: 2 atom_name: SG CB CA
        TEMPLATE::      ATOM_MAP: 2 residue3: CYS

        CONSTRAINT::    distanceAB:        1.8   0.10    60.0  1     0
        CONSTRAINT::    angle_A:         109.00   6.00    20.00 360.00 0
        CONSTRAINT::    angle_B:         120.00   3.00    20.00 360.00 0
        CONSTRAINT::    torsion_A:         0.00   180.00  20.00 360.00 18
        CONSTRAINT::    torsion_B:        70.00   20.00   20.00 360.00 2
        CONSTRAINT::    torsion_AB:       70.00   20.00   20.00 120.00 2
CST::END
CST::BEGIN
        TEMPLATE::      ATOM_MAP: 1 atom_name: C20 C19 N4
        TEMPLATE::      ATOM_MAP: 1 residue3: ECl

        TEMPLATE::      ATOM_MAP: 2 atom_name: SG CB CA
        TEMPLATE::      ATOM_MAP: 2 residue3: CYS

        CONSTRAINT::    distanceAB:        1.8  0.10    60.0  1     0
        CONSTRAINT::    angle_A:         109.00  6.00    20.00 360.00 0
        CONSTRAINT::    angle_B:         120.00  3.00    20.00 360.00 0
        CONSTRAINT::    torsion_A:         0.00  180.00  20.00 360.00 18
        CONSTRAINT::    torsion_B:        70.00  20.00   20.00 360.00 2
        CONSTRAINT::    torsion_AB:       70.00  20.00   20.00 120.00 2
CST::END
```

Match algorithm

```
$ROSETTA3/bin/match.linuxgccrelease @substrate.flags @matching.flags
-s input_pdb/1ri5_dCysV_relaxed.pdb -match:scaffold_active_site_residues_for_geomcsts Zcl_pos.cst
```

@matching.flags

```
-packing
-ex1
-ex2
-extrachi_cutoff 0
-use_input_sc
-match:filter_colliding_upstream_residues
-match:filter_upstream_downstream_collisions
-match:updown_collision_tolerance 0.5
-match::bump_tolerance 0.5
-match_grouper SameSequenceAndDSPositionGrouper
-output_format PDB
-out:path:pdb ./output
-enumerate_ligand_rotamers
-only_enumerate_non_match_redundant_ligand_rotamers
-match:euclid_bin_size 2.5
-match:euler_bin_size 20.0
-out::file::output_virtual
-mute protocols.idealize
-dynamic_grid_refinement
-match:consolidate_matches
```

@Zcl_pos.cst

```
N_CST 2
1: 58 109 112 115
2: 109 183 186
```

@substrate.flags

```
-extra_res_fa ZCl_conf.params
-match:lig_name ZCl
-match:geometric_constraint_file Zcl_const_new.cst
```

EnzDes algorithm

```
<ROSETTASCRIPTS>
<SCOREFXNS>
        <ScoreFunction name="beta_csts" weights="beta_nov16_cst.wts" />
        <ScoreFunction name="beta" weights="beta_nov16.wts" />
        <ScoreFunction name="constraints_only" weights="empty.wts" >
                <Reweight scoretype="atom_pair_constraint"       weight="1.0" />
                <Reweight scoretype="coordinate_constraint"      weight="1.0" />
                <Reweight scoretype="angle_constraint"           weight="1.0" />
                <Reweight scoretype="dihedral_constraint"        weight="1.0" />
                <Reweight scoretype="res_type_constraint"        weight="1.0" />
        </ScoreFunction>
</SCOREFXNS>
<RESIDUE_SELECTORS>
                <ResidueName name="AZO" residue_name3="ZC1"/>
</RESIDUE_SELECTORS>

<TASKOPERATIONS>
                <IncludeCurrent name="include_curr" />
                <InitializeFromCommandline name="init" />

                <SetCatalyticResPackBehavior name="fix_cysazo" fix_catalytic_aa="1"/>
                <DetectProteinLigandInterface name="dpli" cut1="4" cut2="6" cut3="8" cut4="10"
                design="0"/>
</TASKOPERATIONS>
<SIMPLE_METRICS>
                <TotalEnergyMetric name="total_energy_beta_csts" scorefxn="beta_csts"
                custom_type="total_beta_csts"/>
                <TotalEnergyMetric name="total_energy_beta" scorefxn="beta" custom_type="total_beta"/>
                <TotalEnergyMetric name="constraint_energy" scorefxn="constraints_only"
                custom_type="constraints_only"/>
</SIMPLE_METRICS>
<FILTERS>
        <Rmsd name="ca_rmsd_chymase" chains="A" threshold="1000000" superimpose="1"/>
</FILTERS>
<MOVERS>
        <AddOrRemoveMatchCsts name="match_cst" cst_instruction="add_new"/>

        <EnzRepackMinimize name="cst_opt" scorefxn_repack="beta_csts" scorefxn_minimize="beta_csts"
        cst_opt="1" design="0" repack_only="1" fix_catalytic="1" minimize_rb="0" minimize_bb="0"
        minimize_sc="1" min_in_stages="0" backrub="0" cycles="1"
        task_operations="include_curr,init,dpli"/>

        <EnzRepackMinimize name="enzrepack_min" scorefxn_repack="beta_csts" scorefxn_minimize="beta_csts"
        cst_opt="0" design="0" repack_only="1" fix_catalytic="1" minimize_rb="1" minimize_bb="0"
        minimize_sc="1" min_in_stages="0" backrub="0" cycles="1"
        task_operations="include_curr,init,dpli,fix_cysazo"/>

        <RunSimpleMetrics name="run_energy_metrics" metrics="total_energy_beta_csts, total_energy_beta,
        constraint_energy" prefix="energy_"/>
</MOVERS>

<APPLY_TO_POSE>
</APPLY_TO_POSE>

<PROTOCOLS>
        <Add mover="match_cst" />
        <Add mover="cst_opt" />
        <Add mover="enzrepack_min" />
        <Add mover="run_energy_metrics" />
        <Add filter="ca_rmsd_chymase" />
</PROTOCOLS>

</ROSETTASCRIPTS>
```

@designflags

```
-run:preserve_header
-ex1
-ex2
-extrachi_cutoff 1
-linmem_ig 10
-no_packstat_calculation
-no_optH true
-use_input_sc
-beta
-score:weights beta_nov16_cst
```

Fast Relax algorithm

```
<ROSETTASCRIPTS>
<SCOREFXNS>
        <ScoreFunction name="beta" weights="beta_nov16_cst.wts" />
        <ScoreFunction name="beta_weighted" weights="beta_nov16_cst.wts" >
                <Reweight scoretype="atom_pair_constraint"       weight="%%wt%%" />
                <Reweight scoretype="coordinate_constraint"      weight="%%wt%%" />
                <Reweight scoretype="angle_constraint"           weight="%%wt%%" />
                <Reweight scoretype="dihedral_constraint"        weight="%%wt%%" />
                <Reweight scoretype="res_type_constraint"        weight="%%wt%%" />
        </ScoreFunction>
</SCOREFXNS>

<RESIDUE_SELECTORS>
</RESIDUE_SELECTORS>

<TASKOPERATIONS>
</TASKOPERATIONS>

<FILTERS>
        <Rmsd name="ca_rmsd" chains="A" threshold="10000000" superimpose="1"/>
        <EnzScore name="cstE" scorefxn="beta" whole_pose="1" score_type="cstE" energy_cutoff="9999.0"/>
        <PoseInfo name="pinfo"/>
        <Rmsd name="ca_rmsd_moveable_loop1" threshold="10000000" superimpose="0" >
                <span begin_res_num="%%start1%%" end_res_num="%%end1%%" />
        </Rmsd>
                <Rmsd name="ca_rmsd_moveable_loop2" threshold="1000000" superimpose="0" >
                <span begin_res_num="%%start2%%" end_res_num="%%end2%%" />
        </Rmsd>
</FILTERS>


<MOVERS>
        <AddOrRemoveMatchCsts name="add_enzdes_csts" cst_instruction="add_new"/>

        <FastRelax name="fastrelax" repeats="8" scorefxn="beta_weighted">
                <MoveMap name="mm">
                        <Span begin="1" end="%%pre_start1%%" chi="0" bb="0" />
                        <Span begin="%%start1%%" end="%%end1%%" chi="1" bb="1" />
                        <Span begin="%%post_end1%%" end="%%pre_start2%%" chi="0" bb="0" />
                        <Span begin="%%start2%%" end="%%end2%%" chi="1" bb="1" />
                        <Span begin="%%post_end2%%" end="%%end2%%" chi="0" bb="0" />
                        <Jump number="1" setting="0" />
                </MoveMap>
        </FastRelax>
</MOVERS>

<APPLY_TO_POSE>
</APPLY_TO_POSE>

<PROTOCOLS>
        <Add mover="add_enzdes_csts" />
        <Add mover="fastrelax" />
        <Add filter="ca_rmsd" />
        <Add filter="ca_rmsd_moveable_loop1" />
        <Add filter="ca_rmsd_moveable_loop2" />
        <Add filter="cstE" />
</PROTOCOLS>

</ROSETTASCRIPTS>
```

Isomerization cis to trans azobenzene

```
<ROSETTASCRIPTS>
        <SCOREFXNS>
                <ScoreFunction name="beta_weighted" weights="beta_nov16_cst.wts" >
                        <Reweight scoretype="atom_pair_constraint" weight="%%wt%%" />
                        <Reweight scoretype="coordinate_constraint" weight="%%wt%%" />
                        <Reweight scoretype="angle_constraint" weight="%%wt%%" />
                        <Reweight scoretype="dihedral_constraint" weight="%%wt%%" />
                        <Reweight scoretype="res_type_constraint" weight="%%wt%%" />
                </ScoreFunction>
                <ScoreFunction name="beta" weights="beta_nov16_cst.wts" />
        </SCOREFXNS>

        <RESIDUE_SELECTORS>
                <ResidueName name="azo" residue_name3="%%AZO%%,%%TRANS_AZO%%"/>
                <Not name="not_azo" selector="azo"/>
        </RESIDUE_SELECTORS>

        <TASKOPERATIONS>
                <ExtraRotamersGeneric name="extra_rot" ex1="1" ex2="1" extrachi_cutoff="0"/>
                <IncludeCurrent name="include_current" />

                <RestrictToRepacking name="repack_only" />
                <LimitAromaChi2 name="limchi2"/> #recommended rosetta wiki 11-2019
                <DetectProteinLigandInterface name="dpli_nodesign" design="0" cut1="6.0" cut2="8.0"
                cut3="10.0" cut4="12.0"/>
                <OperateOnResidueSubset name="rep_azo" selector="azo" >
                        <RestrictToRepackingRLT/>
                </OperateOnResidueSubset>
                <OperateOnResidueSubset name="fix_else" selector="not_azo" >
                        <PreventRepackingRLT/>
                </OperateOnResidueSubset>
        </TASKOPERATIONS>

        <FILTERS>
                <Rmsd name="ca_rmsd" chains="AB" threshold="1000000000" superimpose="1" />
                        <EnzScore name="cstE" scorefxn="beta" whole_pose="1" score_type="cstE"
                        energy_cutoff="9999.0"/>
                        <PoseInfo name="pinfo" />
                        <Rmsd name="ca_rmsd_moveable_loop1" threshold="10000000" superimpose="0" >
                                <span begin_res_num="%%start1%%" end_res_num="%%end1%%" />
                        </Rmsd>
                        <Rmsd name="ca_rmsd_moveable_loop2" threshold="1000000" superimpose="0" >
                                <span begin_res_num="%%start2%%" end_res_num="%%end2%%" />
                        </Rmsd>
        </FILTERS>

        <MOVERS>
                <AtomTree name="foldtree" fold_tree_file="foldtree"/>
                <MutateResidue name="mutate" target="1X" new_res="%%TRANS_AZO%%"
                preserve_atom_coords="true" />
                <PackRotamersMover name="repack1" scorefxn="beta_weighted"
                task_operations="rep_azo,fix_else"/>
                <AddOrRemoveMatchCsts name="add_enzdes_csts" cst_instruction="add_new" />
                        <TaskAwareMinMover name="min" bb="0" chi="1" jump="1" scorefxn="beta_weighted"
                        tolerance="0.001" max_iter="400" task_operations="extra_rot,include_current"/> #
                        tolerance std = 0.01., max_iter=200
                <PackRotamersMover name="repack"
                task_operations="extra_rot,include_current,dpli_nodesign,limchi2"
                scorefxn="beta_weighted"/>

       <ParsedProtocol name="min_repack_min">
          <Add mover="min"/>
          <Add mover="repack"/>
          <Add mover="min"/>
       </ParsedProtocol>
```

```
                <FastRelax name="fastrelax" repeats="8" scorefxn="beta_weighted"
                movemap_disables_packing_of_fixed_chi_positions="true"
                task_operations="limchi2,include_current,extra_rot">
                <MoveMap name="mm" >
                        <Span begin="1" end="%%pre_start1%%" chi="0" bb="0" />
                        <Span begin="%%start1%%" end="%%end1%%" chi="1" bb="1" />
                        <Span begin="%%post_end1%%" end="%%pre_start2%%" chi="0" bb="0" />
                        <Span begin="%%start2%%" end="%%end2%%" chi="1" bb="1" />
                        <Span begin="%%post_end2%%" end="252" chi="0" bb="0" />
                        <Jump number="1" setting="1" />
                       <ResidueSelector selector="azo" chi="1" bb="1" />
                </MoveMap>
               </FastRelax>

        </MOVERS>

        <APPLY_TO_POSE>
        </APPLY_TO_POSE>

        <PROTOCOLS>
                <Add mover="foldtree"/>
                <Add mover="mutate" />
                <Add mover="add_enzdes_csts" />
               <Add mover="min_repack_min"/>
                <Add mover="fastrelax" />
                <Add filter="ca_rmsd" />
                <Add filter="ca_rmsd_moveable_loop1" />
                <Add filter="ca_rmsd_moveable_loop2" />
                <Add filter="cstE" />
        </PROTOCOLS>
</ROSETTASCRIPTS>
```

Supporting Figures

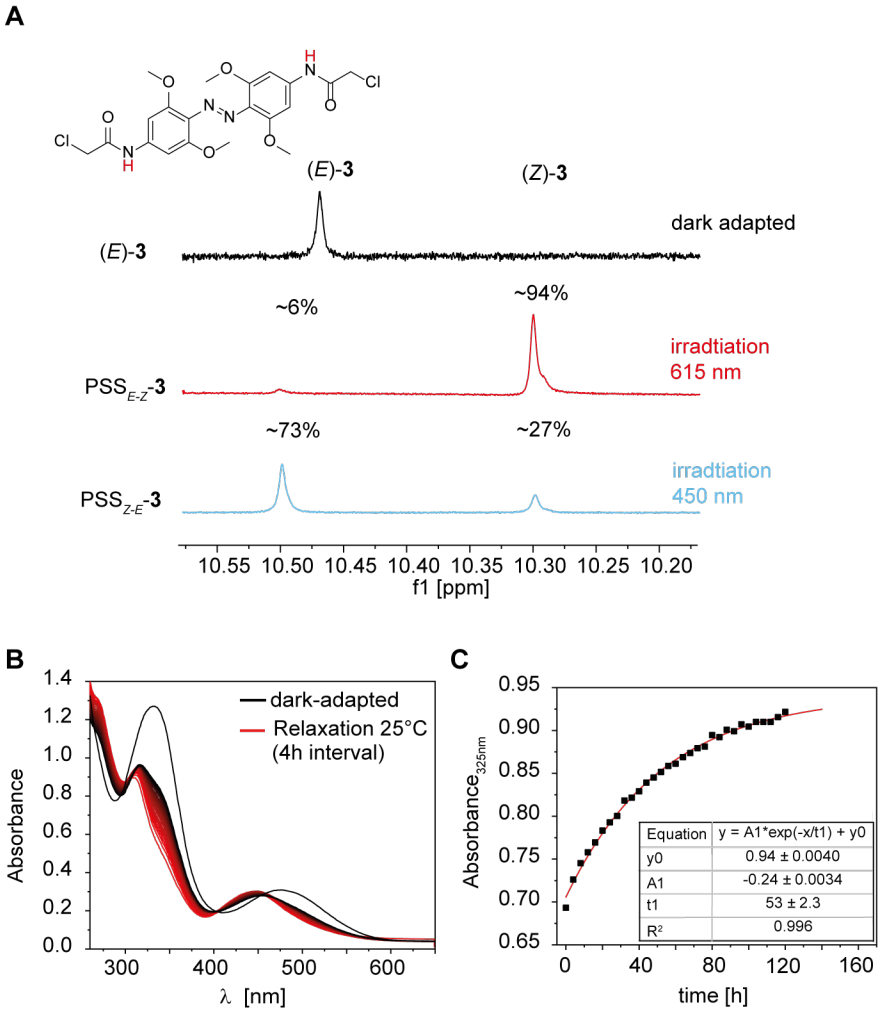


Figure S1: Characterization of azobenzene derivative (3). A) ^{1}H NMR spectra of the *E*-isomer (dark-adapted), $PSS_{E\rightarrow Z}$, and $PSS_{Z\rightarrow E}$ of **3** recorded in DMSO-*d*6. Individual amide protons are marked in red. B) The thermal stability of 1 mM **3** was determined at 25 °C. The first spectrum of thermally relaxed azobenzene (dark-adapted) was recorded. After irradiation with λ= 615 nm for 30 min the absorbance spectrum was monitored over 120 h and recorded every 4 h. C) Plot showing the normalized absorbance of **3** at 325 nm against relaxation time was used to determine the half-life of the *cis* conformer.

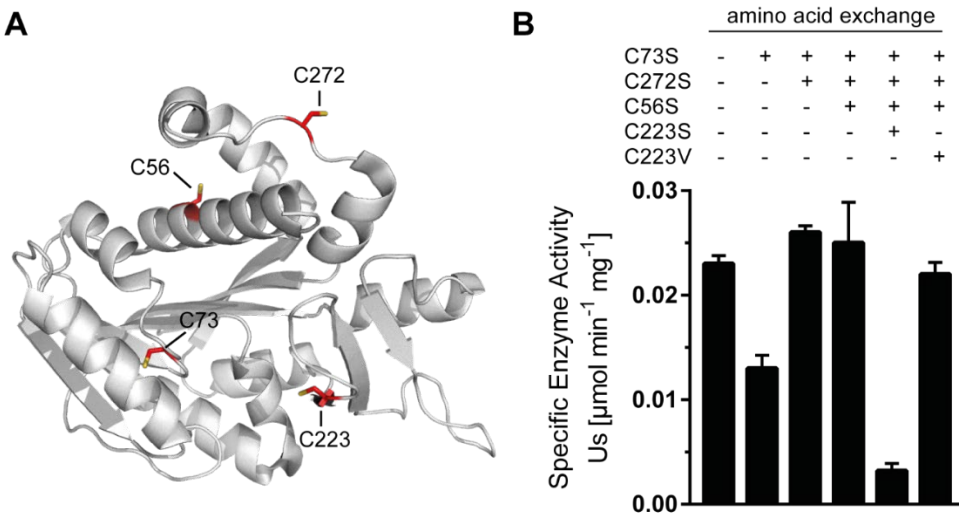


Figure S 2: Native cysteine residues in wild-type Ecm1. A) Crystal structure of Ecm1 (PDB 1RI5) showing native cysteines as sticks in red. B) Native cysteine residues were substituted by serine (S) or valine (V) using side-directed mutagenesis. Respective cysteine-free variants were subjected to activity assays. Methylation of the substrate GpppA (**1a**) in the presence of SAM (**2a**) at 37°C was monitored, and the specific enzyme activity was compared to the wild-type enzyme. Data and error bars show average and standard deviation of three independent experiments (n=3).

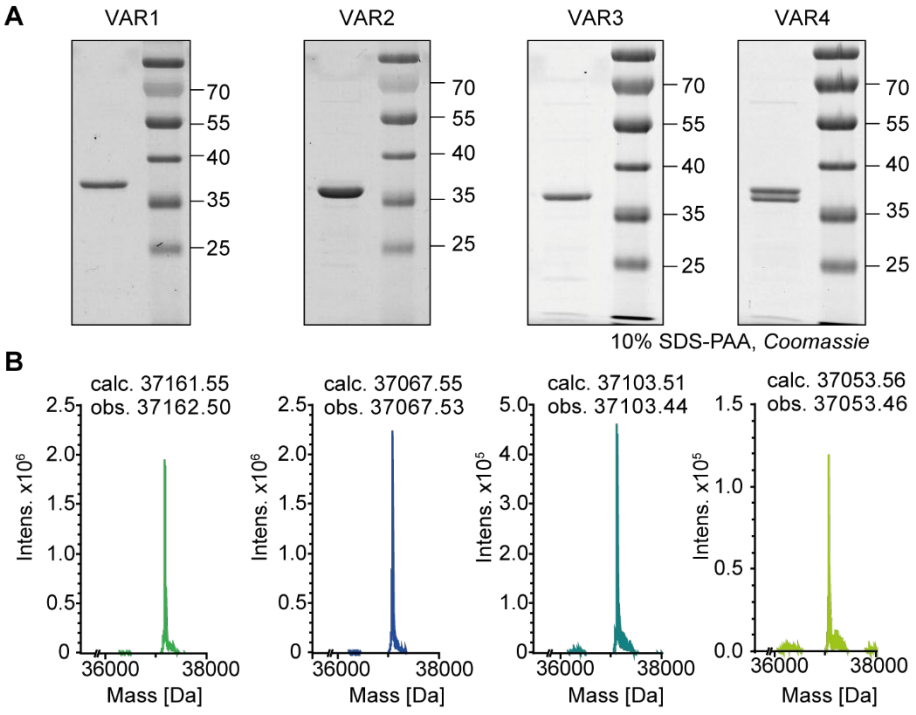


Figure S3: Characterization of purified Ecm1 variants. A) SDS-PAA gel of Ecm1 variants VAR1–4 after NiNTA purification. B) Monoisotopic mass of purified Ecm1 variants VAR1–4 confirmed the presence of desired cysteine mutations.

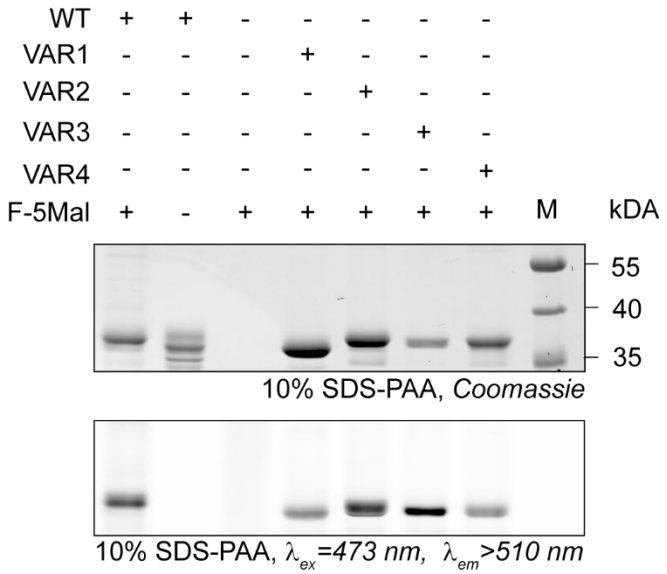


Figure S4: Protein labeling with Fluorescein-5-maleimide (F-5Mal). 10 µM of the respective protein were labeled with a 25-fold excess of the fluorophore. Samples were separated on a 10 % SDS-PAA gel. The fluorescence signal was recorded on a Typhoon gel imaging system with λ_{ex} =473 nm, λ_{em} ≥ 510 nm.

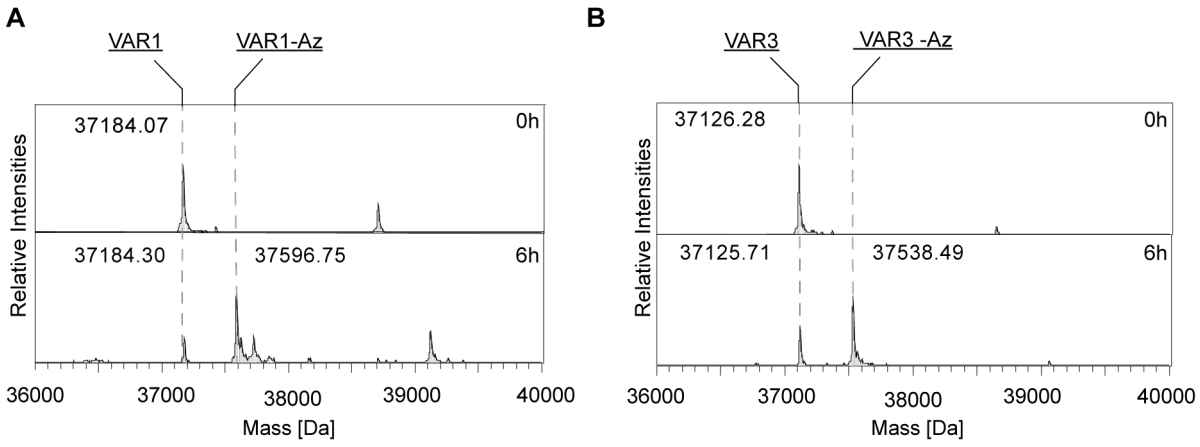


Figure S5: LC/MS analysis of monitoring conjugation of Ecm1 variants with 3. Deconvoluted mass spectrum showing the average mass of Ecm1 VAR1 (obs. 37184.07) and VAR1-Az (obs. 37596.75) in A) and Ecm1 VAR3 (obs. 37126.28) and VAR3-Az (obs. 37538.49) in B). The samples were taken at the indicated time-points after the azobenzene derivative **3** was added to the protein solution. Subsequently, samples were immediately quenched adding 5% formic acid and analyzed via LC/MS.

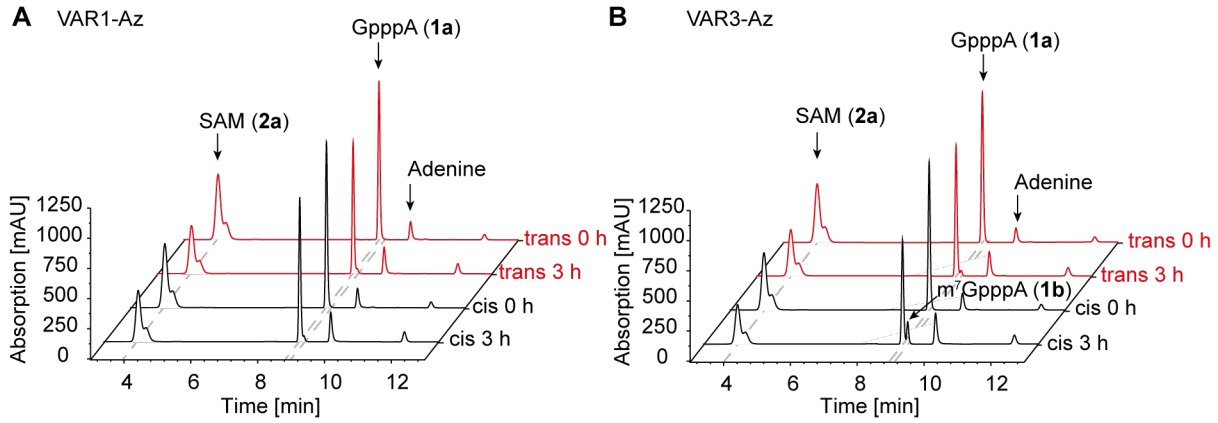


Figure S6: HPLC chromatogram of reaction of GpppA (**1a)** with SAM (**2a**) after biotransformations with Ecm1 VAR1-Az (left) and VAR3-Az (right). Respective samples were irradiated with 615 nm (*cis*) or 450 nm (*trans*).

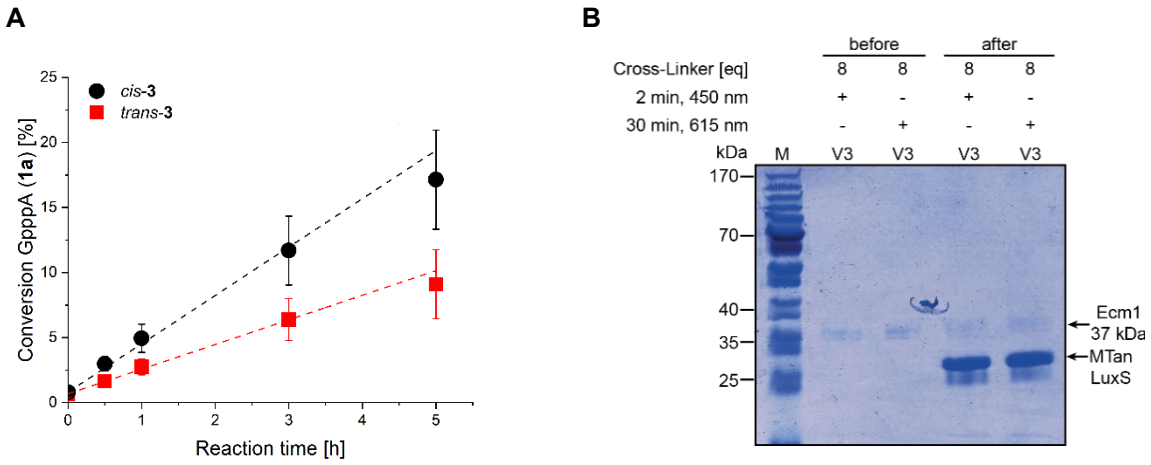


Figure S7: A) Biotransformation of GpppA (**1a**) with Ecm1 VAR3-Az with azobenzene **3** in *cis*-state (black) and *trans*-state (red). B) SDS-PAGE of modified Ecm1 VAR3-Az irradiated for 2 min at 450 nm and 30 min at 615 nm, before and after addition to the enzyme activity assay.

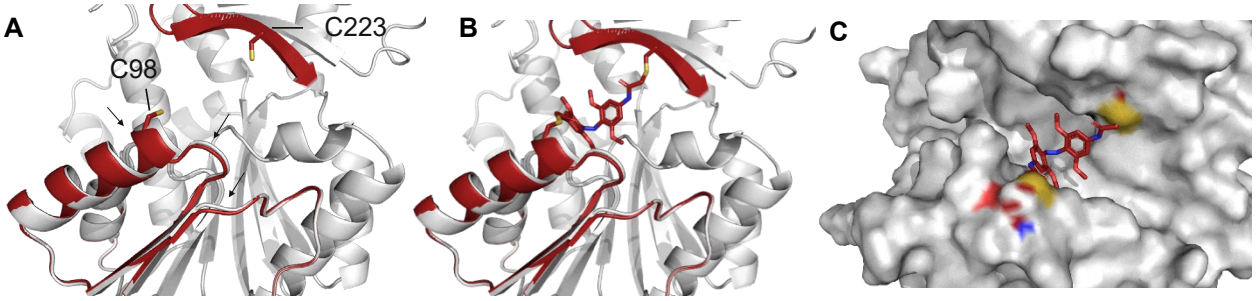


Figure S 8: Computational model of Ecm1 VAR1-Az (C98/C223) cross-linked with azobenzene **3** in the *cis*-state (red) vs. the starting structure (gray). A) The overlay of the two Ecm1 models indicates which loops have been moved to accommodate the azobenzene in the *cis*-state (arrows). B) Rosetta model after FastRelax script (red) shown in A) with azobenzene ligand. C) The surface representation of the computational model indicates how the cross-linked azobenzene is oriented in the binding pocket.

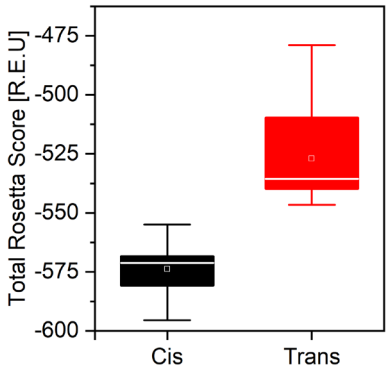


Figure S9: Box plot showing the distributions of Total Rosetta Score in Rosetta energy units (R.E.U) for the *cis* (black) and *trans* (red) states. Pooled results for 50 FastRelax simulations for *cis* and 100 simulations for the *trans*-state.

Supporting Tables

Table S1: Ecm1 variants produced in this work

Variant	Amino acid exchange
ΔC1	C73S
ΔC2	C73S, C272S
ΔC3	C56S, C73S, C272S
ΔC4	C56S, C73S, C223V C272S
Ecm1 Var1	ΔC3 V98C
Ecm1 Var2	ΔC4 S152C Y226C
Ecm1 Var3	ΔC4 T149C I155C
Ecm1 Var4	ΔC4 T149C Y226C

Table S2: Constraint Score for cross linked models after EnzDes algorithm

Variant	Models	Rosetta Score	Constraint Energy
C98 / C223	281	-392.036	1.152
C152 / C226	216	-390.283	1.377
C149 / C226	111	-389.041	1.724
C149 / C155	109	-387.752	1.349
C149/ C223	82	-386.971	8.832
C152 / C223	22	-385.417	1.593
C149 / C152	11	-379.567	5.481

Aminoacid sequence of the expressed Ecm1 variants

Ecm1 VAR1

```
        10         20         30         40         50         60
MRGSHHHHHH GMASMTGGQQ MGRDLYDDDD KDRWGSMEGK KEEIREHYNS IRERGRESRQ

        70         80         90        100        110        120
RSKTINIRNA NNFIKASLIR LYTKRGDSVL DLGSGKGGDL LKYERAGIGE YYGVDIAECL

       130        140        150        160        170        180
INDARVRARN MKRRFKVFFR AQDSYGRHMD LGKEFDVISS QFSFHYAFST SESLDIAQRN

       190        200        210        220        230        240
IARHLRPGGY FIMTVPSRDV ILERYKQGRM SNDFYKIELE KMEDVPMESV REYRFTLLDS

       250        260        270        280        290        300
VNNCIEYFVD FTRMVDGFKR LGLSLVERKG FIDFYEDEGR RNPELSKKMG LGSLTREESE

       310
VVGIYEVVVF RKLVPESDA
```

ECM1 VAR2

```
        10         20         30         40         50         60
MRGSHHHHHH GMASMTGGQQ MGRDLYDDDD KDRWGSMEGK KEEIREHYNS IRERGRESRQ

        70         80         90        100        110        120
RSKTINIRNA NNFIKASLIR LYTKRGDSVL DLGSGKGGDL LKYERAGIGE YYGVDIAEVS

       130        140        150        160        170        180
INDARVRARN MKRRFKVFFR AQDSYGRHMD LGKEFDVISS QFSFHYAFAT SECLDIAQRN

       190        200        210        220        230        240
IARHLRPGGY FIMTVPSRDV ILERYKQGRM SNDFYKIELE KMEDVPMESV REYRFTLLDS

       250        260        270        280        290        300
VNNVIECFVD FTRMVDGFKR LGLSLVERKG FIDFYEDEGR RNPELSKKMG LGSLTREESE

       310
VVGIYEVVVF RKLVPESDA
```

ECM1 VAR3

```
        10         20         30         40         50         60
MRGSHHHHHH GMASMTGGQQ MGRDLYDDDD KDRWGSMEGK KEEIREHYNS IRERGRESRQ

        70         80         90        100        110        120
RSKTINIRNA NNFIKASLIR LYTKRGDSVL DLGSGKGGDL LKYERAGIGE YYGVDIAEVS

       130        140        150        160        170        180
INDARVRARN MKRRFKVFFR AQDSYGRHMD LGKEFDVISS QFSFHYAFAC SESLDCAQRN

       190        200        210        220        230        240
IARHLRPGGY FIMTVPSRDV ILERYKQGRM SNDFYKIELE KMEDVPMESV REYRFTLLDS

       250        260        270        280        290        300
VNNVIEYFVD FTRMVDGFKR LGLSLVERKG FIDFYEDEGR RNPELSKKMG LGSLTREESE

       310
VVGIYEVVVF RKLVPESDA
```

ECM1 VAR4

```
        10         20         30         40         50         60
MRGSHHHHHH GMASMTGGQQ MGRDLYDDDD KDRWGSMEGK KEEIREHYNS IRERGRESRQ

        70         80         90        100        110        120
RSKTINIRNA NNFIKASLIR LYTKRGDSVL DLGSGKGGDL LKYERAGIGE YYGVDIAEVS

       130        140        150        160        170        180
INDARVRARN MKRRFKVFFR AQDSYGRHMD LGKEFDVISS QFSFHYAFAC SESLDIAQRN

       190        200        210        220        230        240
IARHLRPGGY FIMTVPSRDV ILERYKQGRM SNDFYKIELE KMEDVPMESV REYRFTLLDS

       250        260        270        280        290        300
VNNVIECFVD FTRMVDGFKR LGLSLVERKG FIDFYEDEGR RNPELSKKMG LGSLTREESE

       310
VVGIYEVVVF RKLVPESDA
```

References

1. J. M. Holstein, L. Anhauser and A. Rentmeister, *Angew. Chem. Int. Ed.*, 2016, **55**, 10899-10903.
2. K. M. Blacklock, B. J. Yachnin, G. A. Woolley and S. D. Khare, *J. Am. Chem. Soc.*, 2018, **140**, 14-17.
3. N. Yasuike, K. M. Blacklock, H. X. Lu, A. S. I. Jaikaran, S. McDonald, M. Uppalapati, S. D. Khare and G. A. Woolley, *ChemPhotoChem*, 2019, **3**, 431-440.
4. M. D. Hanwell, D. E. Curtis, D. C. Lonie, T. Vandermeersch, E. Zurek and G. R. Hutchison, *J Cheminform*, 2012, **4**, 17.
5. C. R. Groom, I. J. Bruno, M. P. Lightfoot and S. C. Ward, *Acta Crystallogr., Sect. B: Struct. Sci., Cryst. Eng. Mater.*, 2016, **72**, 171-179.
6. P. Conway, M. D. Tyka, F. DiMaio, D. E. Konerding and D. Baker, *Protein Sci.*, 2014, **23**, 47-55.
7. F. Richter, A. Leaver-Fay, S. D. Khare, S. Bjelic and D. Baker, *PLoS One*, 2011, **6**, e19230.
8. S. J. Fleishman, A. Leaver-Fay, J. E. Corn, E. M. Strauch, S. D. Khare, N. Koga, J. Ashworth, P. Murphy, F. Richter, G. Lemmon, J. Meiler and D. Baker, *PLoS One*, 2011, **6**, e20161.