Supporting Information

General Remarks

PSCaa **1** was synthesized as described previously.[1] The synthesis of Keto-PSCaa **2** and Cl-PSCaa **3** will be described elsewhere. Primer were obtained from IDT or Valuegene.

Library construction

Overlapping PCRs were performed on the *M. mazei* PylRS gene to randomize the codons for each selected active site residue using oligonucleotides containing NNK, SYT or GBT at each site destined for mutagenic diversification. The amplified full-length PCR product was ligated into the precut pBK5-MmPylRS[2] plasmid to afford the mutant DNA library CRIZ. The library was electroporated into DH10βT1 competent cells harboring pREP selection plasmid. Selection was carried out on GMML plates containing 12.5 μg mL^{-1} tetracycline, 50 μg mL^{-1} kanamycin, 40 μg mL^{-1} chloramphenicol and 1 mM PSCaa as previously described.[2]

Primers used in library construction are the following with mutated resides highlighted in yellow:

302R: 5′-GGGCACGGTCCAGTTTACGMNNATAGTTGTACAGGTTCGGAGDCAGCAT-3′

309R: 5′-GGGCACGGTCCAGTTTACGMNNATAGTTGTACAGGTTCGGAGTCAGCATCGG-3′

322F: 5′-GTAAACTGGACCGTGCCCTGCCGGACCCGATCAAANNKTTCGAGATCGGTCCTTGC-3′

346R: 5′-CAGCATGGTGAATTCTTCCAG-3′

346/348F: 5′-CTGGAAGAATTCACCATGCTGSYTTTCGBTCAGATGGGTAGCGGTTGC-3′

401R: 5′-TGCGGAAGACAGCTCCAG-3′

401F: 5′-CTGGAGCTGTCTTCCGCASYTGTGGGCCCAATCCCGCTGGATCGTGAGTGGGGTATCGAC-3′

417/419F: 5′-GATCGTGAGTGGGGTATCGACAAACCTNNKATCGBTGCGGGTTTTGGTCTGGAGCGTCTG-3′

Protein Expression and Uaa Incorporation in *E. coli*

To translationally incorporate Uaas into myoglobin in *E. coli*, BL21 cells were transformed with plasmids pTak-Myo4TAG and pBK-MmPSCaaRS (or pBK-MmKetoPSCaaRS). For each sample, a colony was picked and grown overnight in 5 mL 2xYT supplemented with 30 μg mL^{-1} chloramphenicol and 50 μg mL^{-1} kanamycin at 37 °C. This starter culture was used to inoculate 50 mL of 2xYT containing antibiotics. When OD_{600} reached 0.5, 1 mM of the corresponding unnatural amino acid was added, and cells were induced for protein expression by adding 0.5 μM IPTG. After 5 h, cells were lysed and sonicated in 5 mL lysis buffer (50 mM TrisHCl, pH 8.0, 500 mM NaCl, 20 mM imidazole pH 8.0, 1% (v/v) Tween 20, 10% (v/v) glycerol and 0.5 mg mL^{-1} lysozyme). Lysed cells were centrifuged for 20 min at 14,000 g, and clarified supernatant was passed through a 0.1 mL column of Ni^{2+}-NTA agarose resin (Qiagen). The column was washed with 10 column volumes of wash buffer (lysis buffer without Tween

20 and lysozyme). Protein was eluted with 400 μL of elution buffer (wash buffer containing 250 mM imidazole, pH 8.0). The sample was concentrated using a Microcon Ultracel YM-10 spin column (Millipore). To incorporate PSCaas into GFP, DH10β cells were transformed with pTak-GFPtag and pBK-MmPSCaaRS. To incorporate Cl-PSCaa **3** into CaM, plasmids pTak-CaM-76TAG and pBK-MmPSCaaRS were transformed into BL-21 cells, and proteins were purified with Ni^{2+}-NTA agarose resin using similar procedures.

Mass spectrometry

The protein samples of myoglobin were desalted on a ZIP tip (Millipore) and analysed by high-resolution Fourier-transform MS on a Thermo LTQ-Orbitrap XL mass spectrometer (ThermoFisher, San Jose, CA). The sample was loaded onto a capillary column (75 mm diameter) with an integrated spray tip, which was filled with reversed phase material (Zorbax SB C-18, particle size 5 mm, bed length 5 cm). Protein was eluted with a gradient of 0.1% formic acid and an increasing proportion of acetonitrile at a flow rate of 300 nL min^{-1}. The eluate was electrosprayed directly into the mass spectrometer. Fourier-transform mass spectra were recorded at a resolution of 60000 for a scan range of m/z 400–1800 followed by a select ion scan of the most intense ion at resolution 60000. Data were charge-deconvoluted using the Thermo Qualbrowser 2.0 Xtract program. Intact proteins of CaM were analyzed by ESI-TOF using an Agilent 6210 mass spectrometer coupled to an Agilent 1100 HPLC system. Two micrograms of protein samples were injected by an auto-sampler and separated on an Agilent Zorbax SB-C8 column (2.1 mm ID × 10 cm length) by a reverse-phase gradient of 0–80% acetonitrile for 15 min. Mass calibration was performed right before the analysis. Protein spectra were averaged and the charge states were deconvoluted using Agilent MassHunter software.

Uaa incorporation in mammalian cells

Genes for MmPylRS and MmPSCaaRS were subcloned into plasmid pMPcua-OmeRS between the *Xho* I and *Nhe* I sites using primers MmPylRS XhoI and MmPylRS NheI.[2] The HeLa-GFP(182TAG) stable cells were transfected with the resultant pMPcua-MmPylRS and pMPcua-MmPSCaaRS. Cells were grown in DMEM supplemented with 10% FBS; 0.1 mM of PSCaa was added to or withheld from the media. Cells were imaged 48 h after transfection with an Olympus IX8I microscope equipped with a Hamamatsu EM-CCD under the same image capture conditions for all samples (λ_{ex} = 480 ± 20 nm, λ_{em} = 535 ± 40 nm). Transfected HeLa cells were trypsinized and washed with PBS twice. After suspending in 1.0 mL PBS and 5 μL of propidium iodide, cells were analysed with a FACScan (Becton & Dickinson, excitation 488 nm, emission 530 ± 30 nm). For each sample the total fluorescence intensity of 10,000 cells was recorded and normalized to cells transfected with pMPcua-MmPylRS with Boc-Lys added.

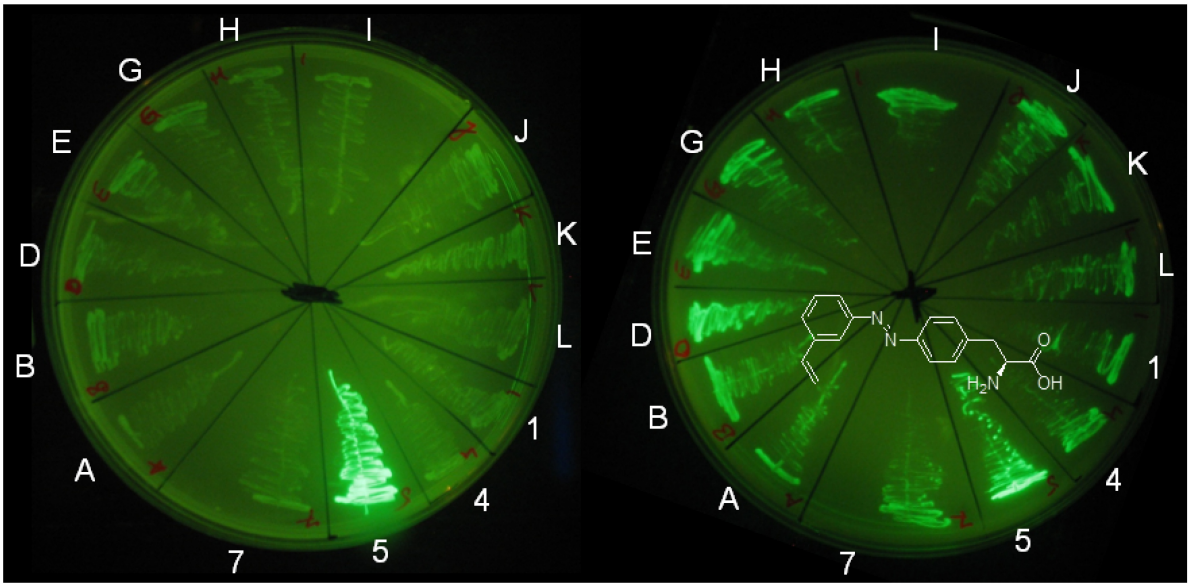
Molecular docking

For MmPSCaaRS and MmKetoPSCaaRS, protein models bearing the selected amino-acid substitutions were generated from the 3QTC template. Initially, the replacement side-chains were introduced with a suitable preferred conformation in the program Coot[3] and the target Uaa ligand was positioned roughly as guided by the Ome in 3QTC, and these atomic groups were then manually adjusted to minimize steric conflicts. The initial synthetase-Uaa complexes were subsequently subjected to one round of simple dynamics with pdbtools in the program Phenix.[4]

Photoisomerization

UV/visible and circular dichroism (CD) spectra of the photoswitchable CaM were measured at a protein concentration of 7 μM on a Bio-Logic MOS450 spectrometer in Na-phosphate (10 mM) buffer containing NaF (100 mM) and at pH 7.5 using a 0.1 cm quartz cuvette (Helma) at 25 °C. The UV and CD spectrum of the pure *trans*-isomer was recorded upon storage in the dark at room temperature. The UV/visible and CD spectrum of the photostationary state of the *cis* state was recorded after illumination at 365 nm for 2 min at 25 °C.

A)



B)

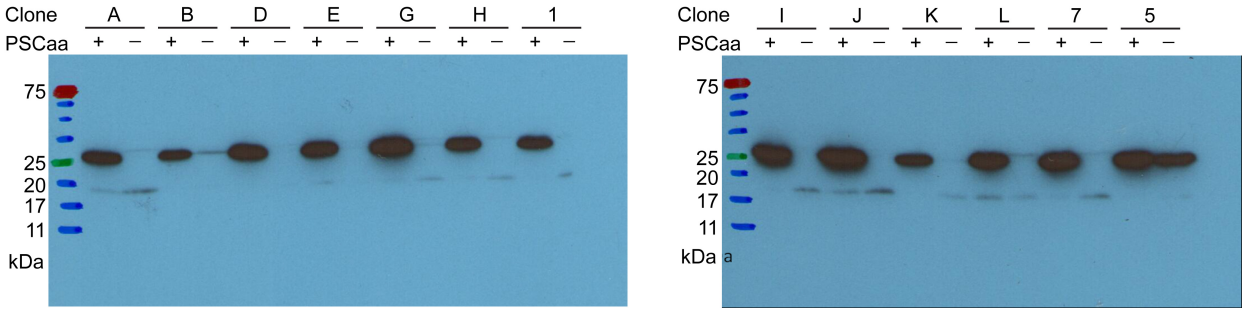


Figure S1. a) Fluorescence imaging of clones streaked on LB plates without PSCaa (left plate) and with 1 mM PSCaa (right plate). b) Western blot analysis of GFP expression by identified clones. Cell lysates were separated by SDS-PAGE and probed with an anti-His antibody (Samples were normalized for constant cell numbers for each lane). Clone 5 lacks exclusive specificity for the PSCaa, as determined by fluorescence and Western blot. Other clones showed PSCaa-dependent GFP fluorescence and expression. Clone D was used for further studies and named MmPSCaaRS.

Position	#A		#B		#D		#E		#G		#H	
	codon	AA	codon	AA	codon	AA	codon	AA	codon	AA	codon	AA
302	TCT	S	ACT	T	ACT	T	ACT	T	ACT	T	TCT	S
309	GCG	A	GCT	A	TCG	S	TCG	S	TCT	S	GGG	G
322	GTT	V	ACT	T	ATT	I	CTT	L	TTG	L	CTG	L
346	GCT	A	GCT	A	GTT	V	GTT	V	GTT	V	GCT	A
348	GGT	G	GGT	G	GGT	G	GGT	G	GGT	G	GGT	G
401	GTT	V	GTT	V	GTT	V	GTT	V	GTT	V	GGT	V
417	TGG	W	TGG	W	TGG	W	TGG	W	TGG	W	TGG	W
419	GGT	G	GGT	G	GGT	G	GGT	G	GGT	G	GGT	G

Position	#I		#J		#K		#L		#1		#7	
	codon	AA	codon	AA	codon	AA	codon	AA	codon	AA	codon	AA
302	TCT	S	ACT	T	ACT	T	TCT	S	ACT	T	TCT	S
309	GCG	A	AGT	S	TCT	S	TCG	S	GCT	A	GCG	A
322	GTT	V	ATT	I	TTG	L	CTG	L	TTG	L	GTT	V
346	GCT	A	GTT	V	GTT	V	GCT	A	GCT	A	GCT	A
348	GGT	G	GGT	G	GGT	G	GGT	G	GGT	G	GGT	G
401	GTT	V	GTT	V	GTT	V	GTT	V	GTT	V	GGT	V
417	TGG	W	TGG	W	TGG	W	TGG	W	TCG	S	TGG	W
419	GGT	G	GGT	G	GGT	G	GCT	A	GGT	G	GGT	G

Table S1. Sequencing of clones obtained after selection for PSCaa. Converged clones are boxed with the same color. Identical substitutions at four sites in multiple hits are highlighted in yellow. Clone #D was used for incorporation of PSCaa (MmPSCaaRS). Clone #B was used to incorporate the Keto-PSCaa **2** (MmKetoPSCaaRS) .

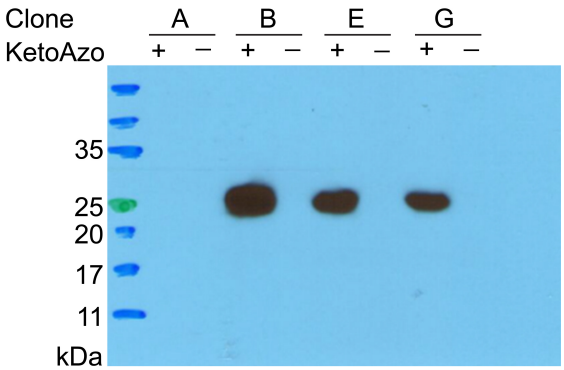


Figure S2. Western blot screen for clones incorporating Uaa **2**. GFP expression in the presence or absence of of **2** (1 mM) was detected by probing the cell lysates separated by SDS-PAGE with an anti-His antibody (Samples were normalized for constant cell numbers for each lane). Clone B was named MmKetoPSCaaRS for further studies.

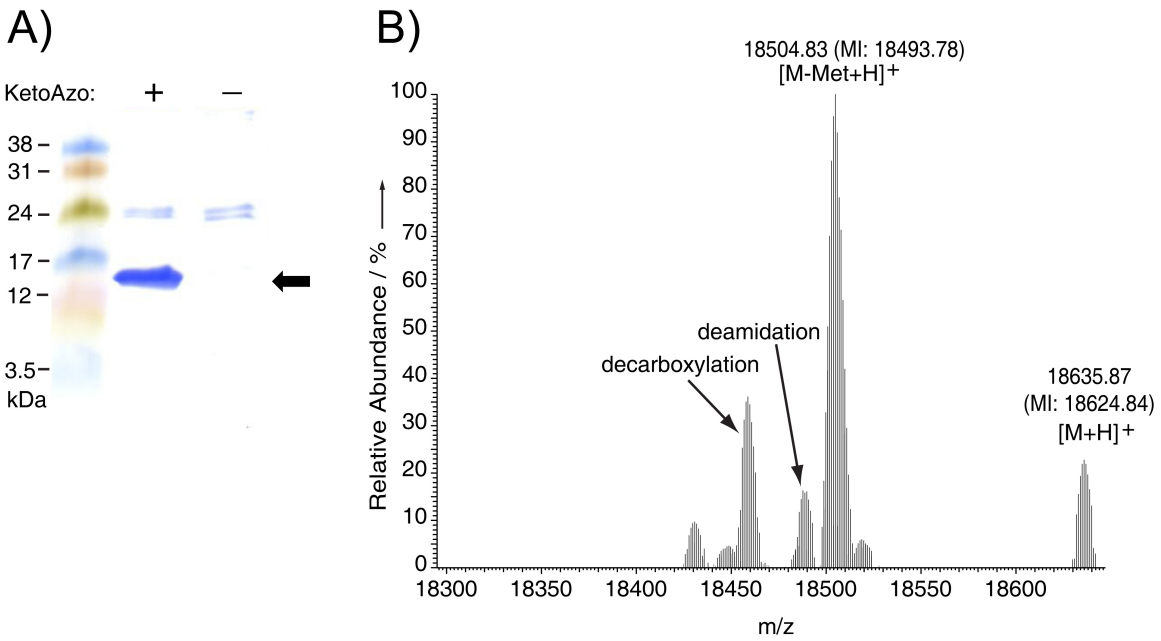


Figure S3. Specific incorporation of Keto-PSCaa **2** into myoglobin by the $tRNA^{Pyl}_{CUA}$-MmKetoPSCaaRS pair. A) SDS-PAGE analysis of myoglobin expression in the presence or absence of 1 mM Keto-PSCaa **2** in the growth medium. Samples were normalized for constant cell numbers for each lane. Arrow indicates the full-length myoglobin protein. B) High resolution ESI-FT MS analysis of of the expressed myoglobin. Average and monoisotopic (indicated by MI) masses are labelled. A peak with a monoisotopic mass of 18624.84 Da was observed, which corresponds to intact myoglobin having a single Keto-PSCaa residue at position 4 (expected $[M+H]^+$ = 18624.81 Da). The second peak observed corresponds to the Keto-PSCaa-containing myoglobin lacking the beginning Met (expected $[M\text{-}Met+H]^+$ = 18493.81 Da, measured 18493.78 Da). Deamidation and decarboxylation peaks for the $[M\text{-}Met+H]^+$ species were also detected and labelled. Notably, no peaks corresponding to myoglobin proteins containing any other amino acids at the amber codon position were observed.

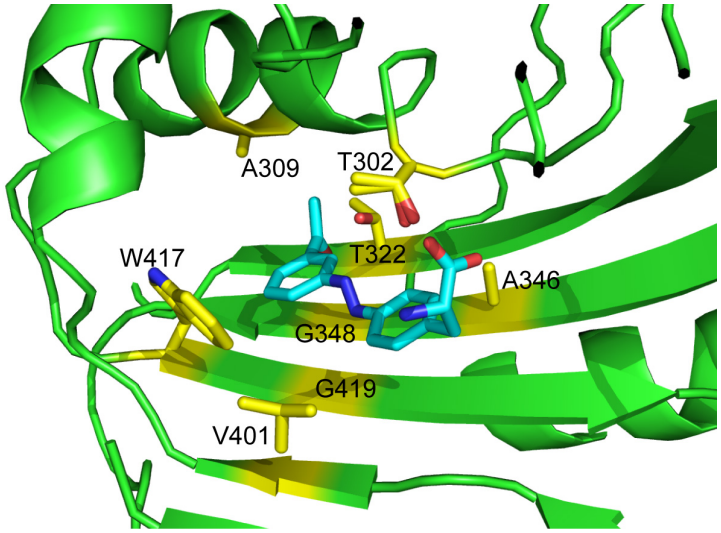


Figure S4. Molecular docking of Keto-PSCaa **2** into a MmKetoPSCaaRS model built from 3QTC bearing the identified substitutions. Recognition of the main chain and the azobenzene moiety of **2** is similar to that of PSCaa **1**, as described in Figure 4 and in the text. Notably, T322 in MmKetoPSCaaRS may hydrogen bond with the keto oxygen of Keto-PSCaa **2**. In mutant synthetases specific for PSCaa **1** and Keto-PSCaa **2**, residue 309 and 322 are substituted by various distinct amino acids (Ala, Gly, or Ser309; and Ile, Leu, Thr, or Val322), whereas in PylRS mutants specific for Phe analogues bearing shorter side chains (Ome, Bpa, and Nap),[2][5] these two residues aptly are not substituted and retain the residue identities (Leu309 and Ile322) of wild-type PylRS. This difference is consistent with the location of these two residues deeper in the binding pocket and in contact with the second phenyl ring and the meta-substituent of PSCaas.

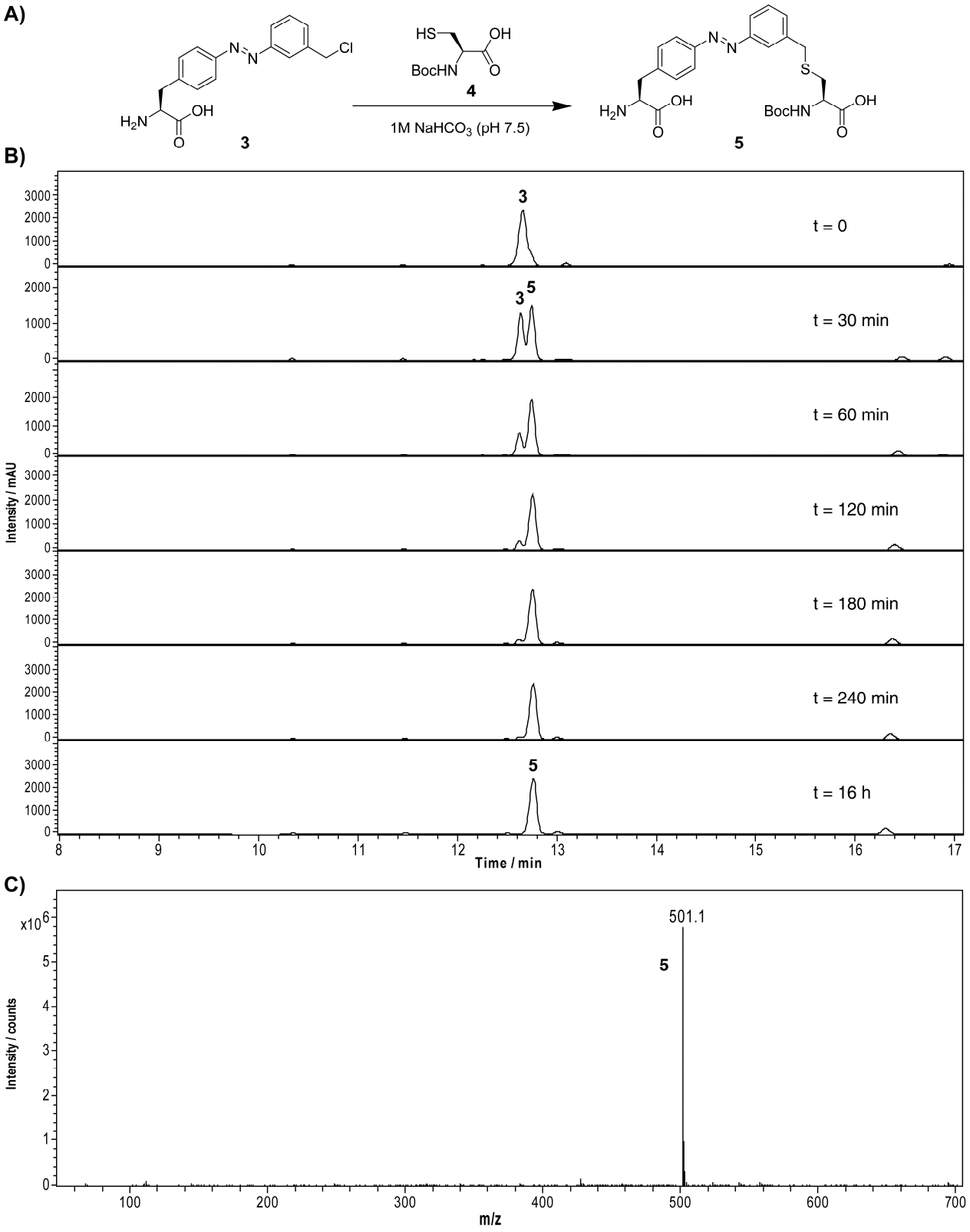


Figure S5. LC-MS monitoring of the crosslinking reaction between Cl-PSCaa **3** and Boc-protected cysteine **4**. A) Scheme showing the crosslinking reaction carried out in 1 M $NaHCO_3$ at pH 7.5. B) UV chromatogram at different time points of the reaction. C) m/z measured for the peak containing product **5**.

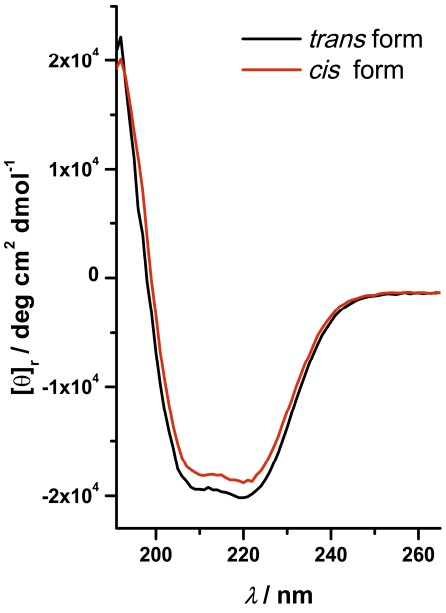


Figure S6. Influence of bridge photoisomerization on the protein conformation of CaM analyzed by CD spectroscopy. When the bridge was photoisomerized from the *trans* form (black line) to the *cis* form (red line), the negative ellipticity for CaM decreased, indicating a decrease in helix content of CaM.

Supporting References:

[1] C. Hoppmann, P. Schmieder, N. Heinrich, M. Beyermann, *ChemBioChem* **2011**, *12*, 2555–2559.

[2] J. K. Takimoto, N. Dellas, J. P. Noel, L. Wang, *ACS Chem. Biol.* **2011**, *6*, 733-743.

[3] P. Emsley, K. Cowtan, *Acta Crystallogr. D. Biol. Crystallogr.* **2004**, *60*, 2126-2132.

[4] P. D. Adams, P. V. Afonine, G. Bunkoczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, *Acta Crystallogr. D. Biol. Crystallogr.* **2010**, *66*, 213-221.

[5] V. K. Lacey, G. V. Louie, J. P. Noel, L. Wang, *ChemBioChem* **2013**, *14*, 2100-2105.