Electronic Supplementary Material (ESI) for ChemComm.

Supporting information

Non-Enzymatic Protein Templates Amide Bond Formation and Provides Catalytic Turnover

N. Brauckhoff *et al.*

1 Methods

Peptide Synthesis and Characterization

Peptide synthesis was performed on Rink-Amide-MBHA (*p*-methylbenzhydrylamine) resin (Merck Millipore) following standard Fmoc-based solid-phase peptide synthesis (SPPS) protocols.[1] All amino acids were coupled using 4 eq. of Fmoc-protected amino acids according to the initial amine-loading of the resin. The amino acids were mixed with 4 eq. of benzotriazole-1-yl-oxy-tris-pyrrolidino-phosphonium hexafluorophosphate (PyBOP) and 8 eq. of *N*,*N*-diisopropylethylamine (DIPEA) in *N*-methyl-2-pyrrolidinone (NMP). The resin was incubated with this solution twice for 3 h. Exception was the couplings of fluorescent transfer group to the thiol function of the N-terminal cysteine. Due to lower coupling efficiency, 8 eq. building block, 8 eq. PyBOP and 16 eq. DIPEA were used here. Fmoc-deprotection was performed for 15 min in NMP using 25 % piperidine. After each double-coupling step, unreacted amines were capped for 10 min at RT using Ac_2O/DIPEA/NMP (1:1:10). For the final *N*-terminal modification, peptides were deprotected as described above and reacted either with capping solution (NMP/Ac_2O/DIPEA (10:1:1)) for final acetylation or with (2-(2-(Fmoc-amino)ethoxy)ethoxy)acetic acid (Fmoc-NH-PEG-COOH) for spacer coupling. After Fmoc-deprotection, the coupled spacer was *N*-terminally modified with fluorescein isothiocyanate (FITC). Peptides were finally cleaved from the resin using TFA/H_2O/EDT/TIPS (94:2.5:2.5:1) for 4 h and precipitated with Et_2O at −20 °C. After final cleavage, the peptides were dissolved in water/acetonitrile (8:2) and purified by reversed-phase HPLC using a Nucleodur C18 reverse-phase column (10 x 125 mm, 110 Å, particle size 5 µm, Macherey-Nagel; solvent A: water + 0.1 % TFA; solvent B: acetonitrile + 0.1 % TFA; flow rate: 6 mL min^{-1}). Obtained fractions were frozen in liquid nitrogen and lyophilized with a Heto PowerDry® LL1500 freeze drying system (Thermo). Peptide identity and purity were confirmed by HPLC/ESI-MS (high-performance liquid chromatography/ electrospray ionization mass spectrometry) using an HPLC 1200 system (Agilent Technologies) equipped with a Zorbax Eclipse, XDB-C18 reverse-phase column (4.6 x 150 mm, particle size 5 µm, Agilent; solvent A: water + 0.1 % TFA; solvent B: acetonitrile + 0.1 % TFA; flow rate: 1 mL min^{-1}, gradient 2−70 % B in 20 min for peptide A derivatives and 30−70 % B in 20 min for peptide B derivatives). As mass detector a LCQ Advantage Max (Finnigan™) has been used. Analytical data is shown in Supplementary Figures S4 and S5. The peptides were quantified photometrically using a V-550 UV/VIS spectrophotometer (Jasco).

Protein Expression and Purification of Template Protein (T)

The sequence of the template protein **T** (for amino acid sequence see Supporting Figures S1) was cloned into a modified pGEX-4T-1 vector (GE Healthcare) and expressed as GST fusion protein in *E. coli* (strain: BL21Gold(DE3)). Terrific Broth (TB) medium supplemented with 150 µg mL^{-1} ampicillin was inoculated with *E. coli* overnight pre-culture. This culture was maintained at 37 °C until an OD_{600} of 0.8 was reached. Protein expression was initiated by addition of 0.1 mM IPTG and performed overnight at 18 °C. Cells were harvested by centrifugation, resuspended in lysis buffer (50 mM HEPES (pH 7.4), 500 mM NaCl, 2 mM PMSF) and disrupted by microfluidizer. The cell lysate was centrifuged at 70.000 rcf at 4 °C for 1 h to separate cell debris.

GST-**T** fusion protein was isolated from the supernatant via affinity chromatography (Glutathione Sepharose 4 Fast Flow, GE Healthcare). Following elution, the GST tag was cleaved by addition of Tobacco Etch Virus (TEV) protease.[2] The resulting two protein fragments were separated using size exclusion chromatography (SEC) (HiLoad 16/600 Superdex 75 pg, GE Healthcare, buffer: 25 mM HEPES (pH 7.4), 100 mM NaCl). Remaining GST contamination was removed using a GSTrap HP 1 mL column (GE Healthcare). The purified protein **T** domain was concentrated via ultrafiltration and stored at −80 °C.

Expression and Purification of 2-POI and 2-POI(C1A)

To express POI (for amino acid sequence see Supporting Figures S1), *E. coli* cells (strain: BL21Gold(DE3)) were transformed with plasmid pET28a(+) containing the corresponding genes. Terrific Broth medium containing 50 µg mL^{-1} kanamycin was inoculated with *E. coli* overnight. This culture was maintained at 37 °C until an OD_{600} of 0.8 was reached. Protein expression was initiated by addition of 0.1 mM IPTG and performed overnight at 18 °C. Harvested cells were resuspended in lysis buffer (50 mM HEPES (pH 7.4), 500 mM NaCl, 2 mM TCEP, 2 mM PMSF) and disrupted by microfluidizer. The cell lysate was centrifuged at 70.000 rcf at 4 °C for 1 h to separate cell debris. The His-tagged fusion protein was isolated from the supernatant via affinity chromatography (HisTrap FF crude 5 mL, GE Healthcare) and the eluted protein further purified using size exclusion chromatography (HiLoad 16/600 Superdex 75 pg, GE Healthcare, buffer: 25 mM HEPES pH 7.4, 100 mM NaCl, 1 mM TCEP). The purified proteins were concentrated by ultrafiltration and stored at −80 °C.

HPLC-based detection of template-mediated transfer reactions

Peptides **1*** and **2**$^{\#}$ were dissolved separately in DMSO to obtain stock solutions with a concentration of 4 mM. For each reaction batch, 10 µM peptide **2**$^{\#}$ was incubated with 500 µM TCEP in 20 mM sodium phosphate at pH 7.4 for 30 min at 30 °C (Stock-A). Subsequently, 5 µM peptide **1*** was freshly prepared with and without 5 µM protein template in 20 mM sodium phosphate buffer at pH 7.4 respectively (Stock-A). To start the reaction, 50 µL of Stock-A and 50 µL of Stock-B were combined so that the components were present at final concentrations of 5 µM (**2**$^{\#}$) and 2.5 µM (**1*** and **T**, if applicable) including 250 µM TCEP. The reaction was carried out at 30 °C. The reactions were stopped using 5 µL TFA (5% v/v). To analyze the reaction progress, 100 µL of the reaction solution was used for analytical HPLC (4.6 x 150 mm, particle size 5 µm, Agilent; solvent A: water + 0.1 % TFA; solvent B: acetonitrile + 0.1 % TFA; flow rate: 1 mL min^{-1}, gradient 25−70 % B in 20 min). Transfer products **1** and ***2**$^{\#}$ were detected via ESI-MS.

Fluorescence-based tracking of transfer reactions

Reactions for fluorescence-based readout were performed in 'transfer buffer' (20 mM sodium phosphate, pH 7.4, 3% HEPES, 500 µM TCEP, 0.01% Tween-20). Peptide stock solutions in DMSO (*c*(peptide) = 2 mM) were prepared and then accordingly diluted into transfer buffer. For each reaction, 20 µM **2**$^{\#}$ was incubated with 500 µM TCEP in transfer buffer for 30 min. The concentration of **T** was adjusted to 1 mM in transfer buffer. Finally, 10 µM **1*** (with specified amount of **T**) was prepared in buffer. To start the reaction, 10 µL of the **2**$^{\#}$ solution and 10 µL of the **1*** solution (with or without template **T**) were combined to finally provide 5 µM **1***, 10 µM **2**$^{\#}$ and 250 µM TCEP in a total of 20 µL. The concentrations of **T** corresponded to 0, 0.02,

0.05, 0.1, 0.25 and 1 eq. depending on the sample. For referencing, a sample containing exclusively 5 µM FITC-modified 5-aminovaleric acid in 20 µl buffer was included. A microplate reader (Safire2 plate reader, Tecan) with a set excitation wavelength ex of 470 nm and an emission wavelength of 525 nm was used to read the samples (excitation bandwidth 20 nm, emission bandwidth 5 nm). The amplification (gain) was adjusted by the instrument for each measurement (integration time of 2000 µs). All measurements were performed at 30 °C. For measurements at increased concentrations (Supporting Figure S8), stock solutions were adjusted accordingly.

To determine the fluorescence intensity (I), all obtained raw data $I_{raw}(X)$ were divided by the corresponding value of the internal standard. For calculation of yiels, 0 and 100 % were derived from reference samples and the data accordingly normalized. The turnover number (TON) was calculated based on the transfer yields of template-mediated and -unmediated reactions (Formula 1).[3]

$$TON = [\eta(\text{X eq. } \mathbf{T}) - \eta(\text{without } \mathbf{T})] \cdot X^{-1}$$

Formula 1: Calculation of the exchange number WZ; η = transfer yield, X = used equivalents (0 - 1 eq.) of **T**.

Protein labelling

Initially, a stock solution of **2-POI** and if required of **2-POI(C1A)** was prepared at a concentration of 2 µM in buffer (20 mM sodium phosphate buffer containing 500 mM TCEP, pH 7.4) and incubated for 30 min at RT. In addition, 2 µM of **1*** was added to 20 mM sodium phosphate buffer with 1 eq. **T**, if necessary. To start the reaction, 50 µL of **2-POI** stock was combined with 50 µL of **1*** stock to give final concentrations of 1 µM of **1***, **2-POI**, and, if required, **T**. This reaction mixture was incubated at 30 °C. To stop the reaction and prepare the samples for subsequent SDS-PAGE, 6 µL were combined with 4 µL of SDS sample buffer at given time points and incubated at 95 °C for 5 min. To visualize the labeling of the target protein, SDS-PAGE was performed, loading 10 µL of the respective sample onto the gel. The resulting SDS gel was imaged using a fluorescence reader (Typhoon Trio+ Variable Mode Imager, Amersham) and subsequently stained via Coomassie staining to visualize present protein. For data see Supporting Figure S10 and S11.

Luciferase activity assay

Initially, stock solutions of **2-POI** and **2-POI(C1A)** were prepared at a concentration of 2 µM in buffer (20 mM sodium phosphate buffer containing 500 mM TCEP, pH 7.4) and incubated for 30 min at RT. In addition, 2 µM of **1*** was added to 20 mM sodium phosphate buffer with 1 eq. **T**. To start the reaction, 70 µL of **2-POI** or **2-POI (C1A)** stock was combined with 70 µL of **1*** stock to give final concentrations of 1 µM of **1***, **T**, as well as **2-POI** or **2-POI (C1A)**. For luciferase readout, 3x 10 µL were collected at the beginning of the reaction and after 120 min, respectively. Each of these aliquots was added to 10 µL Dual-Glo luciferase reagent (Promega) in a 96-well plate and then analyzed using a Safire 2 plate reader (measurement mode: luminescence). For data see supporting Figure S12.

2 Supporting Table

Supporting Table S1: Synthesized peptides with characteristics (For sequences and analytical data see Supporting Figure S4 and S5).

Peptide	Molecular formula	*M*	*m/z*	
		g·mol^{-1}	calculated	found
f-1a	$C_{162}H_{254}N_{49}O_{54}P_S$	3815.08	1908.5/1272.7/954.8/764.0	1908.4/1272.7/954.7/764.0
f-1b	$C_{156}H_{240}N_{47}O_{54}PS$	3700.90	1851.5/1234.6/926.2/741.2	1851.0/1234.4/926.2/741.1
f-1	$C_{156}H_{240}N_{47}O_{53}PS$	3684.90	1843.5/1229.3/922.2/738.0	1843.6/1228.9/921.9/738.0
f-2	$C_{108}H_{155}N_{23}O_{31}S_2$	2335.65	1168.8/779.6/584.9	1168.5/779.3/585.0
1*	$C_{166}H_{256}ClN_{49}O_{56}PS_2$	3929.21	1965.6/1310.7/983.3/786.8	1965.7/1310.4/983.0/786.7
2$^{\#}$	$C_{110}H_{169}N_{28}O_{33}S_2$	2486.63	1244.3/829.9/622.7	1245.2/830.4/623.1

3 Supporting Figures

template (**T**)

GAMGVRKGWHEHVTQDLRSHLVHKLVQAIFPTPDPAALKDRRMENLVAYAKKVEGDMYESANSRDEYYHLLAEKIYKIQK
ELEEKRRSR

2-POI

CGNILPSDIMDFVLKNTLEVLFQGPHMEDAKNIKKGPAPFYPLEDGTAGEQLHKAMKRYALVPGTIAFTDAHIEVDITYA
EYFEMSVRLAEAMKRYGLNTNHRIVVCSENSLQFFMPVLGALFIGVAVAPANDIYNERELLNSMGISQPTVVFVSKKGLQ
KILNVQKKLPIIQKIIIMDSKTDYQGFQSMYTFVTSHLPPGFNEYDFVPESFDRDKTIALIMNSSGSTGLPKGVALPHRT
ACVRFSHARDPIFGNQIIPDTAILSVVPFHHGFGMFTTLGYLICGFRVVLMYRFEEELFLRSLQDYKIQSALLVPTLFSF
FAKSTLIDKYDLSNLHEIASGGAPLSKEVGEAVAKRFHLPGIRQGYGLTETTSAILITPEGDDKPGAVGKVVPFFEAKVV
DLDTGKTLGVNQRGELCVRGPMIMSGYVNNPEATNALIDKDGWLHSGDIAYWDEDEHFFIVDRLKSLIKYKGYQVAPAEL
ESILLQHPNIFDAGVAGLPDDDAGELPAAVVVLEHGKTMTEKEIVDYVASQVTTAKKLRGGVVFVDEVPKGLTGKLDARK
IREILIKAKKGGKIAV

2-POI(C1A)

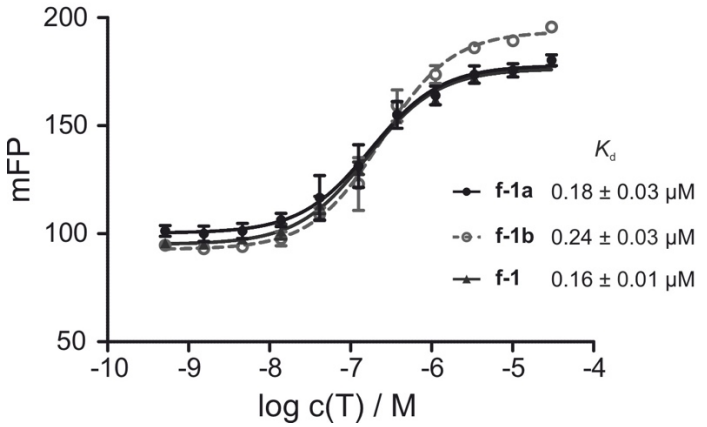
AGNILPSDIMDFVLKNTLEVLFQGPHMEDAKNIKKGPAPFYPLEDGTAGEQLHKAMKRYALVPGTIAFTDAHIEVDITYA
EYFEMSVRLAEAMKRYGLNTNHRIVVCSENSLQFFMPVLGALFIGVAVAPANDIYNERELLNSMGISQPTVVFVSKKGLQ
KILNVQKKLPIIQKIIIMDSKTDYQGFQSMYTFVTSHLPPGFNEYDFVPESFDRDKTIALIMNSSGSTGLPKGVALPHRT
ACVRFSHARDPIFGNQIIPDTAILSVVPFHHGFGMFTTLGYLICGFRVVLMYRFEEELFLRSLQDYKIQSALLVPTLFSF
FAKSTLIDKYDLSNLHEIASGGAPLSKEVGEAVAKRFHLPGIRQGYGLTETTSAILITPEGDDKPGAVGKVVPFFEAKVV
DLDTGKTLGVNQRGELCVRGPMIMSGYVNNPEATNALIDKDGWLHSGDIAYWDEDEHFFIVDRLKSLIKYKGYQVAPAEL
ESILLQHPNIFDAGVAGLPDDDAGELPAAVVVLEHGKTMTEKEIVDYVASQVTTAKKLRGGVVFVDEVPKGLTGKLDARK
IREILIKAKKGGKIAV

Supporting Figure S1: Protein sequences.

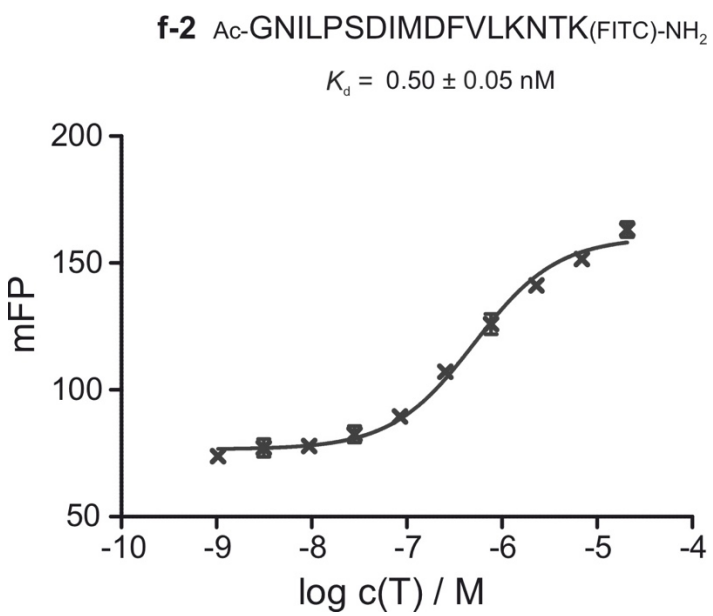
f-1a FITC-PEG-DSQKRREILSRRPpSYRKILNDLSSDAP-NH_2

f-1b FITC-PEG-DSQ**A**RREILSRRPpSYR**A**ILNDLSSDAP-NH_2

f-1 FITC-PEG-D**A**Q**A**RREILSRRPpSYR**A**ILNDLSSDAP-NH_2

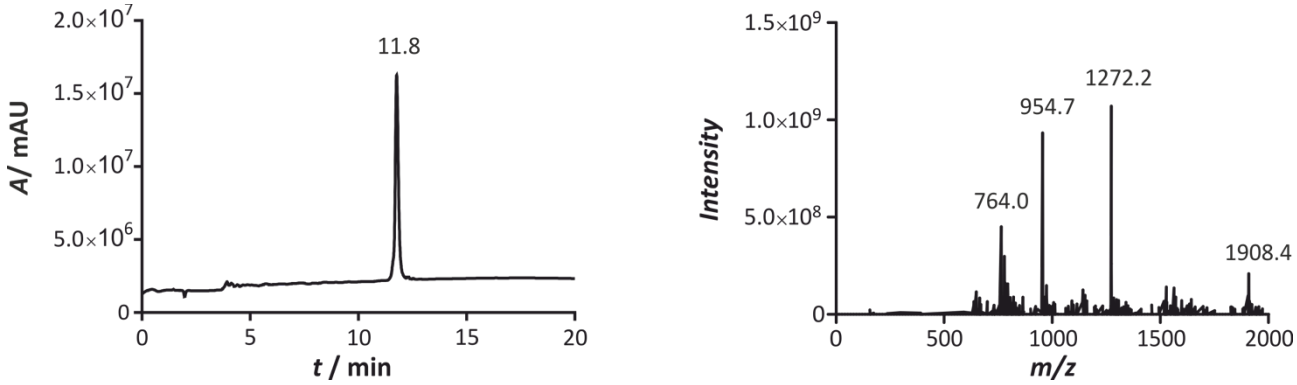


Supporting Figure S2: Fluorescently labelled peptides **f-1**, **f-1a**, **f-1b** derived from the cAMP response element-binding protein (CREB) with variations at three positions. Peptide **f-1a** contains the original CREB-derived sequence. Fluorescence polarization assay of these peptides with protein **T** including derived dissociation constants (K_d).

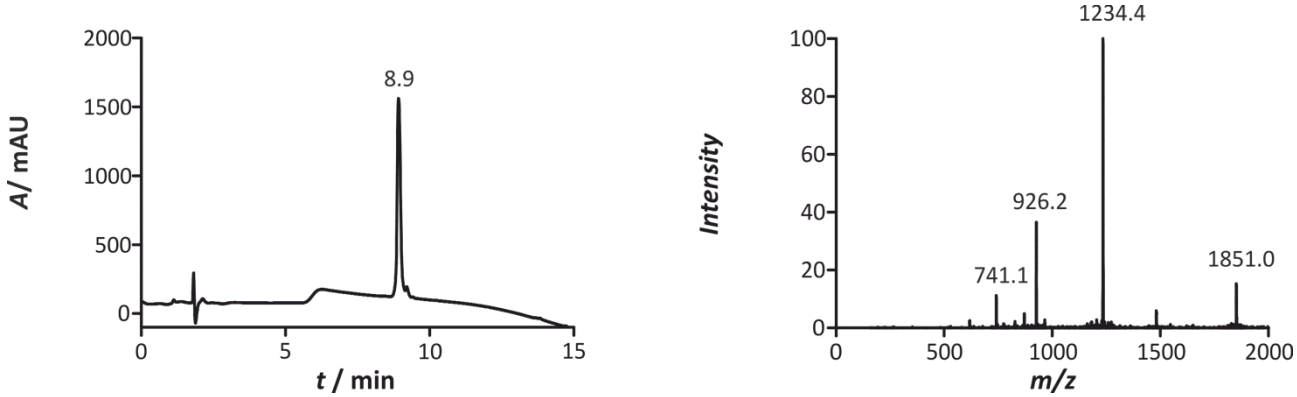


Supporting Figure S3: Fluorescently labelled peptide **f-2** derived from the mixed-lineage leukemia (MLL) protein. Fluorescence polarization assay of the peptide with protein **T** including derived dissociation constant (K_d).

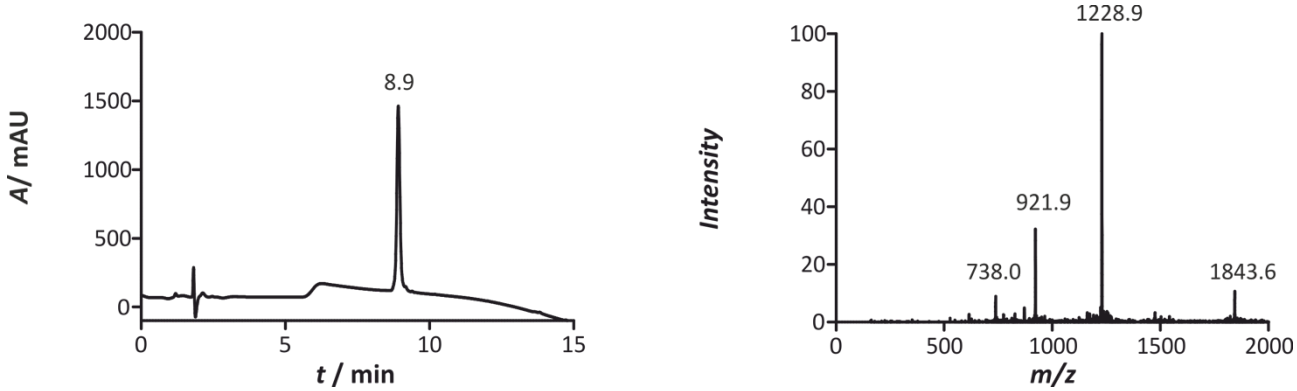
f-1a FITC-PEG-DSQKRREILSRRPpSYRKILNDLSSDAP-NH_2 (M: 3815.08 g·mol^{-1})



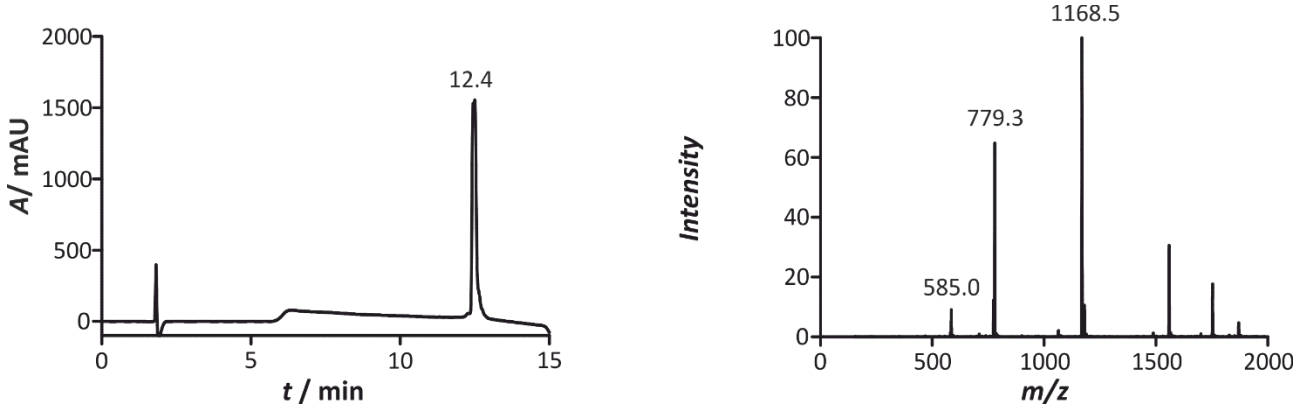
f-1b FITC-PEG-DSQARREILSRRPpSYRAILNDLSSDAP-NH_2 (M: 3700.90 g·mol^{-1})



f-1 FITC-PEG-DAQARREILSRRPpSYRAILNDLSSDAP-NH_2 (M: 3684.90 g·mol^{-1})

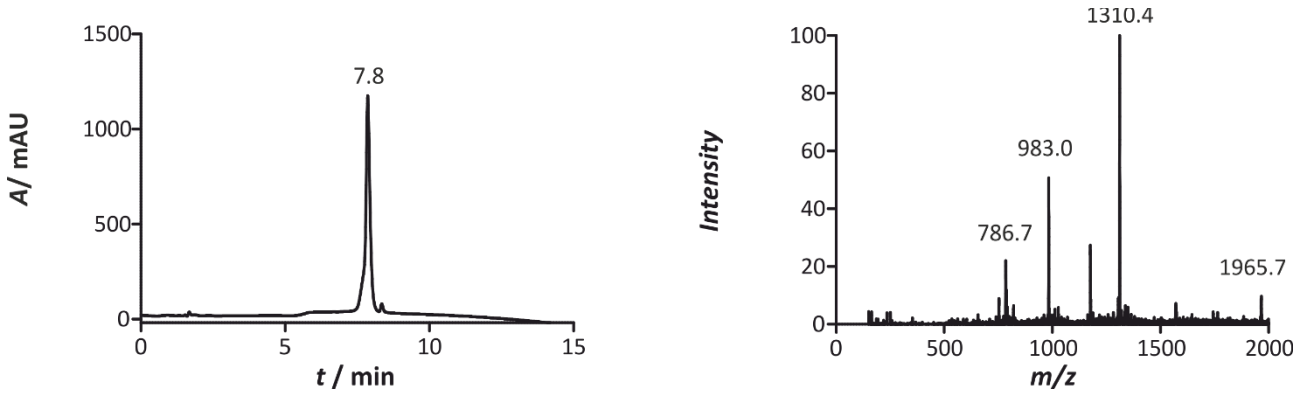


f-2 Ac-GNILPSDIMDFVLKNTK(FITC)-NH_2 (M: 2335.65 g·mol^{-1})

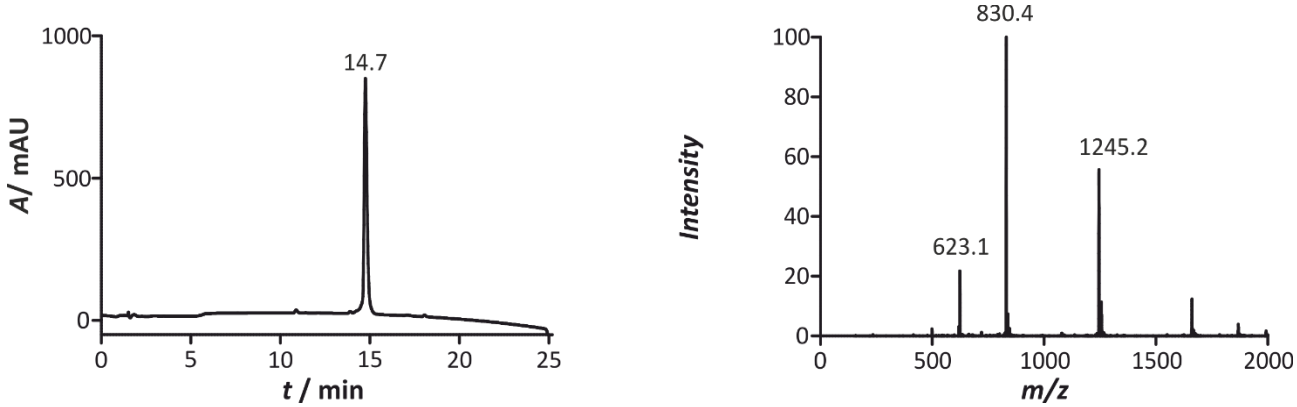


Supporting Figure S4: Characterization of peptides **f-1a**, **f-1b**, **f-1** and **f-2** (for calculated m/z-values see Supporting Table S1).

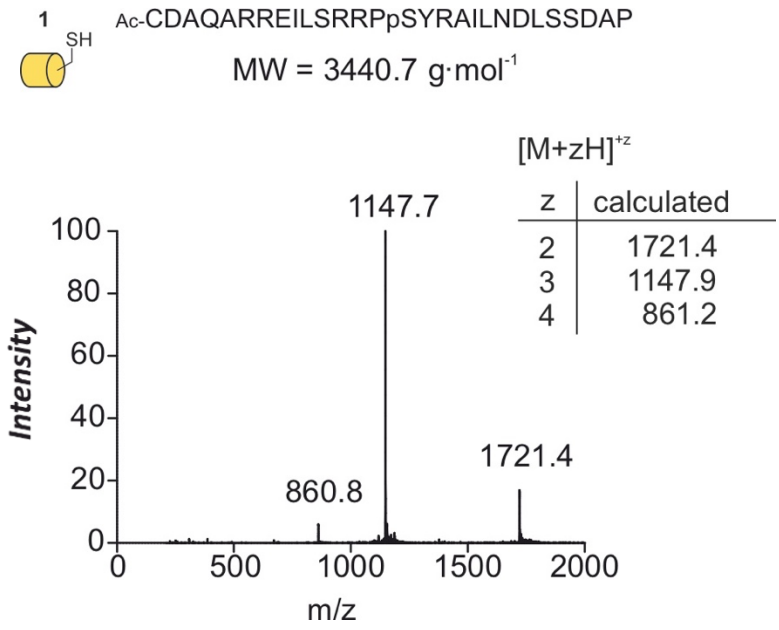
1* Ac-C[FITC-NH(CH_2)$_4$CO]DAQARREILSRRPpSYRAILNDLSSDAP-NH_2 (M: 3929.21 g·mol^{-1})



2$^{\#}$ H_2N-CDDGNILPSDIMDFVLKNTK(DABCYL)-NH_2 (M: 2486.63 g·mol^{-1})

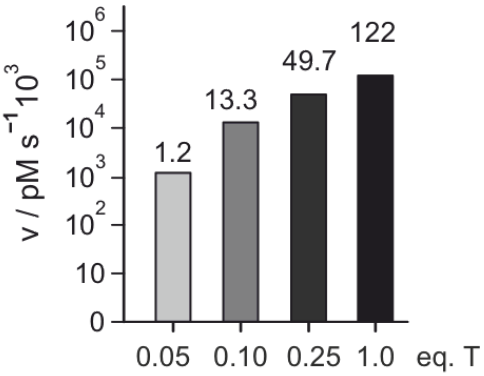


Supporting Figure S5: Characterization of peptides **1*** and **2**$^{\#}$ (for calculated m/z-values see Supporting Table S1).



Supporting Figure S6: Sequence and MS spectrum of reaction product **1** including calculated and found *m*/*z* values.

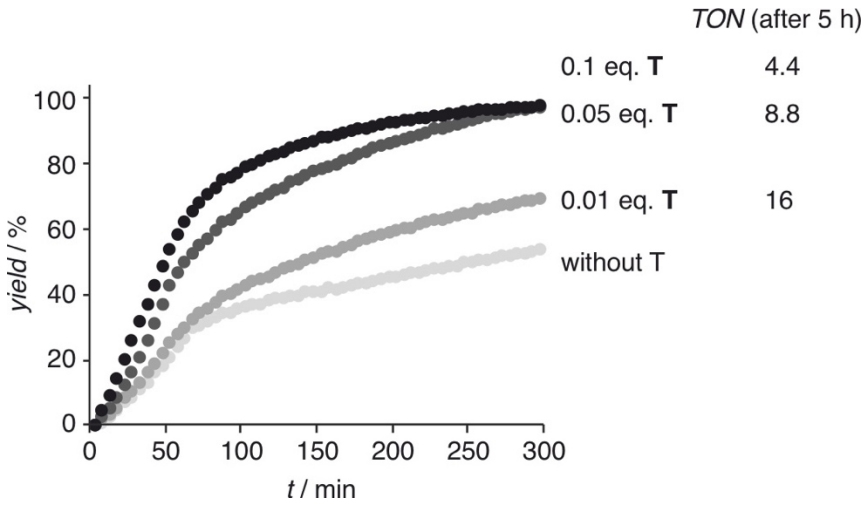
A



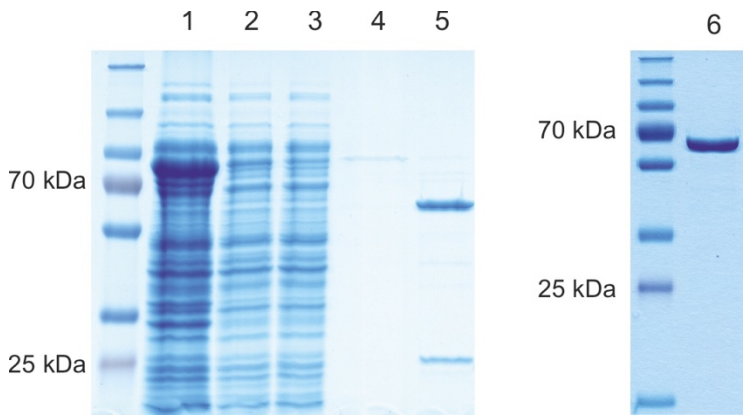
B

eq. **T**	2 h	5 h	(*t*)
0.05	2.0	3.8	(*TON)*
0.10	3.1	5.0	(*TON*)
0.25	2.6	3.6	(*TON*)

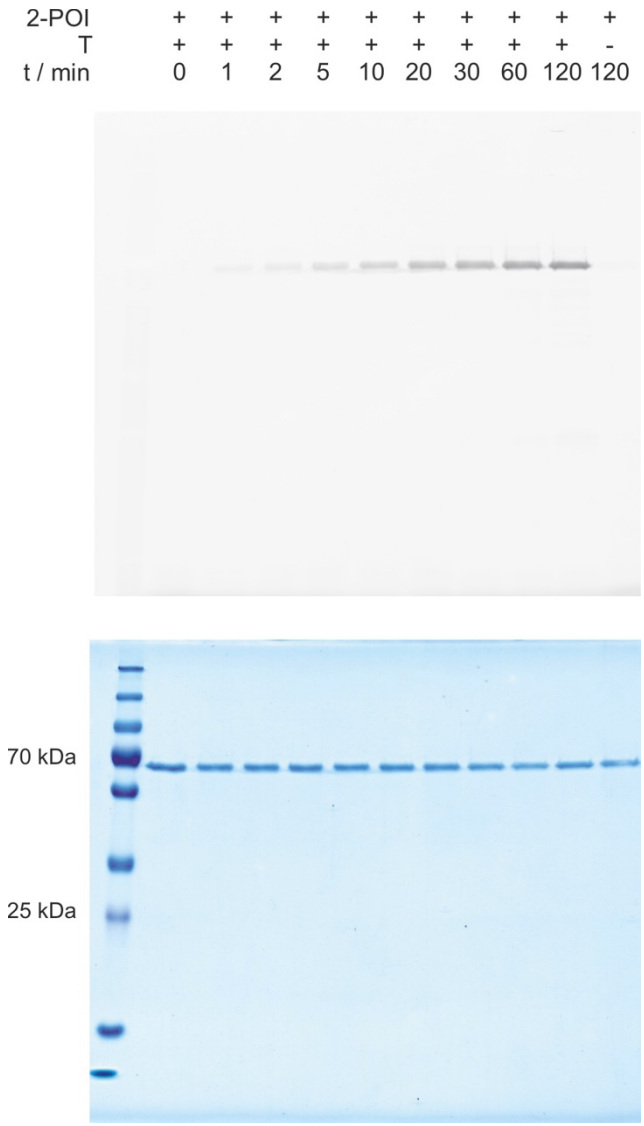
Supporting Figure S7: A) Initial rates for fluorescence-based readout (manuscript Figure 3C). B) Corresponding turnover numbers (*TON*) after 2 h and 5 h.



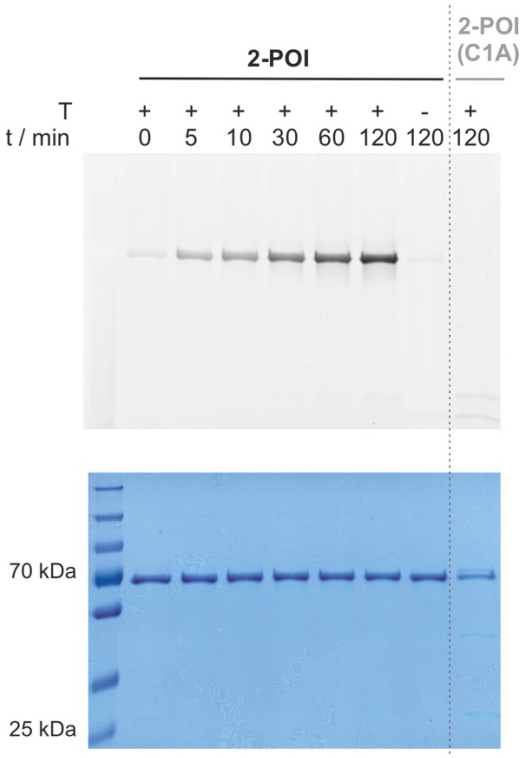
Supporting Figure S8: Reaction time-course determined based on fluorescence-based read out for different equivalents of **T** (conditions: *T* = 35 °C, phosphate buffer, pH 7.4, *c* = 50 µM **1***, 75 µM **2**$^{\#}$). Turnover numbers (*TON*) after 5 h are provided.



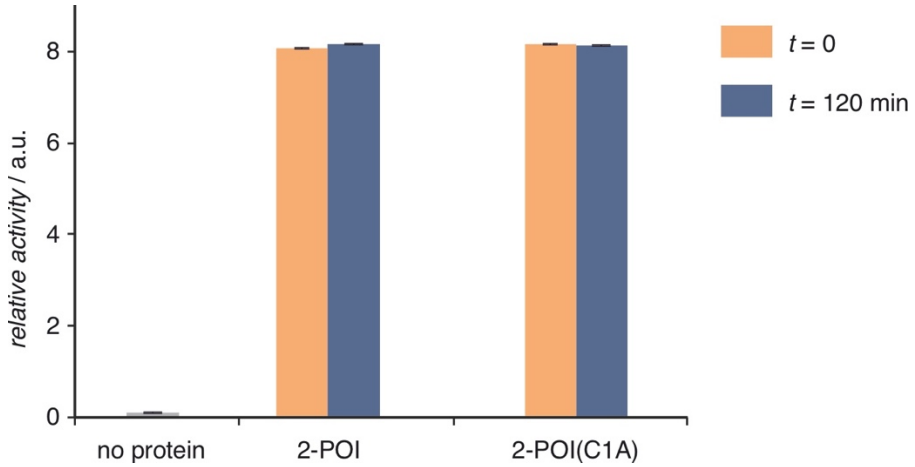
Supporting Figure S9: SDS-PAGE gel of **2-POI** expression. Lane 1 = digested cell lysate, lane 2 = supernatant of cell lysate after centrifugation, lane 3 = flow-through of Ni-NTA column purification, lane 4 = wash step of Ni-NTA column purification, lane 5 = sample after digestion using TEV protease and lane 6 = purified **2-POI** after size exclusion chromatography.



Supporting Figure S10: Full SDS-PAGE gel of manuscript Figure 4B.



Supporting Figure S11: SDS-PAGE analysis of time-dependence of labelling reaction with **1***, **2-POI** or **2-POI(C1A)** and **T** (each *c* = 1µM, buffer: 20 mM sodium phosphate, pH 7.4, 500 µM TCEP, *T* = 30 °C). Top: fluorescence imaging of gel indicating labelled protein ***2-POI**, Bottom: Coomassie-stained gel indicative of total protein content.



Supporting Figure S12: Luciferase activity of 2-POI and 2-POI(C1A) before and after the labelling reaction (*t* = 120 min). Errors account for standard deviation derived from three technical replicates.

References

1 W. C. Chan and P. D. White, *Fmoc solid phase peptide synthesis: a practical approach*, Oxford University Press, New York, 2000.

2 D. S. Waugh, *Protein Expression Purif.*, 2011, **80**, 283–293.

3 M. Boudart, *Chem. Rev*, 1995, **95**, 661–666.