Electronic Supplementary Material (ESI) for ChemComm.

Electronic Supplementary Information

Light-control of the ultra-fast Gp41-1 split intein with preserved stability of a genetically encoded photo-caged amino acid in bacterial cells

Jana K. Böcker, Wolfgang Dörner, and Henning D. Mootz*

Institute of Biochemistry

University of Muenster

Wilhelm-Klemm-Str. 2

D-48149 Münster, Germany

*Email: Henning.Mootz@uni-muenster.de

Supplemental Figures

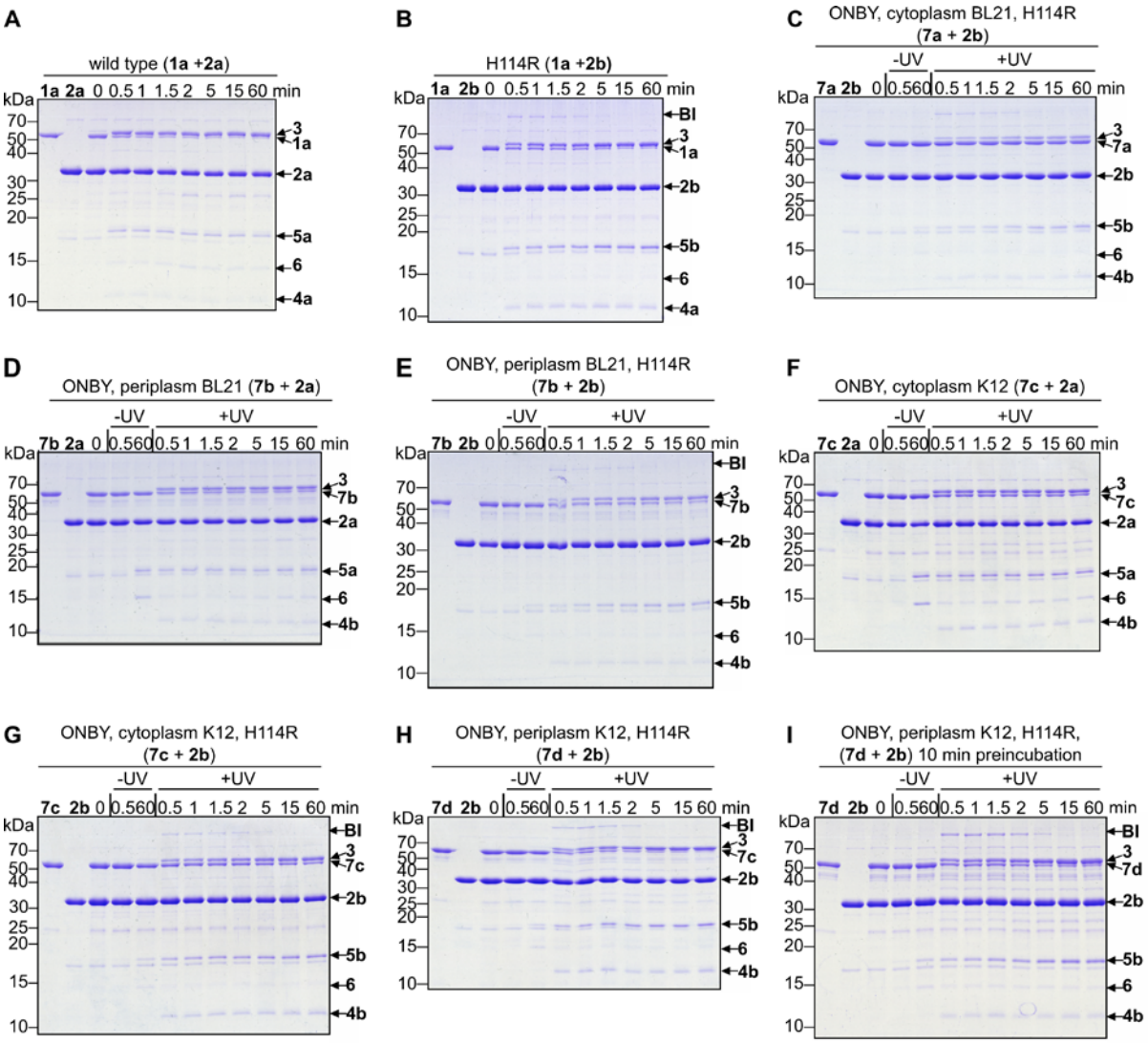

Figure S1 Protein *trans*-splicing assays of the Gp41-1 intein fragments. Shown are Coomassie-stained SDS-PAGE gels. Splice assays were carried out with Int^N constructs MBP-Int^N-H_6 (**1a**, **7a**, **7b**, **7c** and **7d**) at 10 µM concentrations and Int^C constructs H_6-smt3-Int^C-Trx-H_6 (**2a** and **2b**) at 30 µM concentration at 37°C in the presence of 2 mM TCEP. Samples were either irradiated for 30 sec using a 365 nm LED lamp (+ UV) or left without irradiation (- UV). See Figures 1 and 2 in the main text for designation of protein constructs and products. BI = branched intermediate.

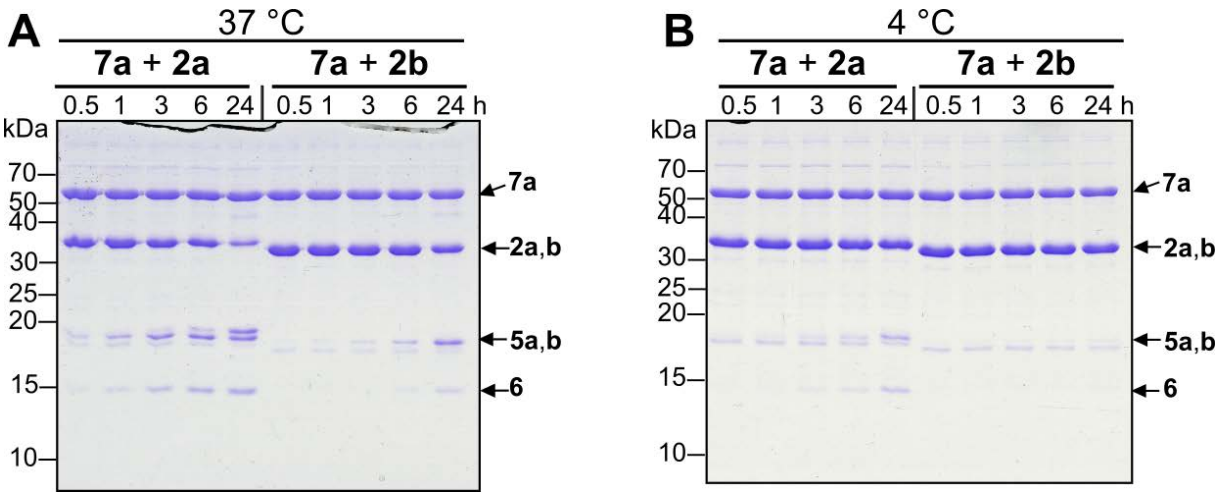

Figure S2 Estimating degree of unintended side reactions. The Int^N construct MBP-Int^N(F65ONBY)-H_6 (**7a**; 10 µM) was incubated with the Int^C construct H_6-smt3-Int^C-Trx-H_6 (**2a**) or H_6-smt3-Int^C(H114R)-Trx-H_6 (**2b**) (each at 30 µM). (A) Reactions with non-irradiated samples were performed at 37°C. Direct comparison of the bands corresponding to the product of C-terminal cleavage Trx-H_6 (**6**) revealed a 7-fold reduction of C-terminal cleavage after 3 h incubation time. (B) Reactions with non-irradiated samples were performed at 4°C. See Figures 1 and 2 in the main text for designation of protein constructs and products.

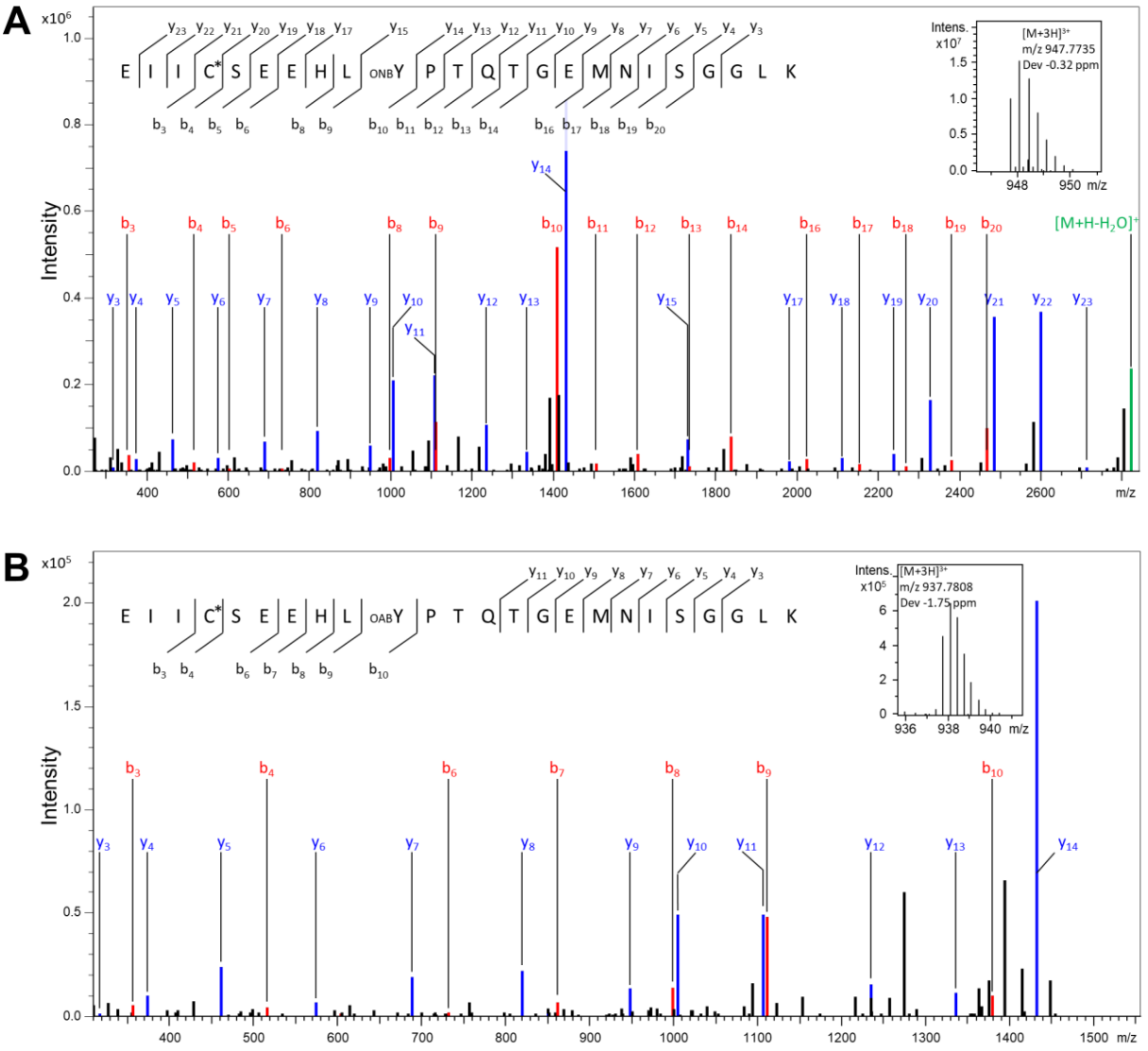

Figure S3 Verification of reduction of ONBY at position 65 to the amine form OABY. A tryptic digest was performed with the band corresponding to Int^N construct **7a** (without irradiation) excised from a Coomassie-stained SDS-PAGE gel. Shown are the fragmentation results of the LC-MS/MS analysis with (A) ONBY at position 65 of the Int^N fragment and (B) OABY at this position. The observed precursor fragment masses of 2840.2993 Da (ONBY peptide) and 2810.3210 Da (OABY peptide) match the calculated masses of 2840.2997 Da (ONBY peptide) and 2810.3255 Da (OABY peptide), respectively. The inserts show the mass analysis of the precursor ions.

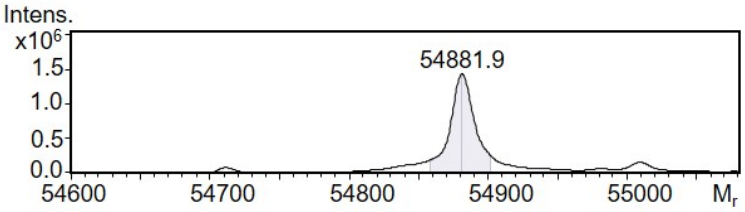

Figure S4 MS analysis of Gp41-1^{N} protein MBP-Int^{N}(F65ONBY)-His$_6$ (**7d**) after incubation with 50 mM DTT at 37 °C for 3 h. The protein was prepared by periplasmatic expression in *E. coli* K12 UT5600(DE3) cells. M_r calculated with ONBY: 54882.5 Da.

Materials and methods

General techniques and materials

E. coli Dh5α cells and standard protocols were used for plasmid construction. pMAL-c2X (NEB) and pMAL-p5X (NEB) were used as backbones for cytoplasmatic and periplasmatic expression of constructs containing Int^N fragments respectively and pBAD (Thermo Fisher Scientific) was used as backbone to encode constructs containing Int^C fragments. *Ortho*-nitrobenzyl-tyrosine (ONBY) was synthesized as previously reported.[1] All protein splicing assays were performed at least in triplicate and error bars represent standard deviations.

Table S1 Plasmids used in this study

No.	Protein	Parent vector	Expression plasmid	Protein sequence
1a	MBP-Int^N-H_6	pMAL-c2X	pAB64	MKTEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEH PDKLEEKFPQVAATGDGPDIIFWAHDRFGGYAQSGLLAEIT PDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNKDL LPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFTWP LIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVDLI KNKHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTS KVNYGVTVLPTFKGQPSKPFVGVLSAGINAASPNKELAKEF LENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIAA TMENAQKGEIMPNIPQMSAFWYAVRTAVINAASGRQTVD EALKDAQTNSSSNNNNNNNNNNLGIEGRISEFLVPRGSTRS GYCLDLKTQVQTPQGMKEISNIQVGDLVLSNTGYNEVLNV FPKSKKKSYKITLEDGKEIICSEEHLFPTQTGEMNISGGL KEGMCLYVKEGGHHHHHH
1b	MBP-Int^N(F65Y)-H_6	pMAL-c2X	pJKB53	See above, with F65Y mutation
2a	H_6-smt3-Int^C-Trx-H_6	pBAD	pAB74	MGSSHHHHHHGSGLVPRGSASMSDSEVNQEAKPEVKPEV KPETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFAKRQGKEM DSLRFLYDGIRIQADQTPEDLDMEDNDIIEAHREQIGGMM LKKILKIEELDERELIDIEVSGNHLFYANDILTHNSSSDVCGTG SDKIIHLTDDSFDTDVLKADGAILVDFWAHWCGPCKMIAPI LDEIADEYQGKLTVAKLNIDHNPGTAPKYGIRGIPTLLLFKNG EVAATKVGALSKGQLKEFLDANLAGSEFRSHHHHHH
2b	H_6-smt3-Int^C(H114R)-Trx-H_6	pBAD	pJKB61	See above, with H114R mutation
7a and **7c**	MBP-Int^N(F65ONBY)-H_6	pMAL-p5X	pJKB43	See **1a**, with F65ONBY mutation
7b and **7d**	SP-MBP-Int^N(F65ONBY)-H_6	pMAL-p5X	pJKB154	See **7a**, with SP-MBP = MKIKTGARILALSALTTMMFSASA LAKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHP DKLEEKFPQVAATGDGPDIIFWAHDRFGGYAQSGLLAEITP DKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNKDLLP NPPKTWEEIPALDKELKAKGKSALMFNLQEPYFTWPLIAAD GGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVDLIKNKHM NADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYG VTVLPTFKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLL TDEGLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAQ KGEIMPNIPQMSAFWYAVRTAVINAASGRQTVDEALKDA QTNSSSNNNNNNNNNNLGIEGRISEFLVPRGSTRS

SP = signal peptide from native MBP; all constructs were verified by DNA sequencing.

Protein expression and purification

E. coli BL21 (DE3)-Gold or K12 UT5600 (DE3) cells were used for protein production. Typically, cells were grown in 600 mL LB medium with appropriate antibiotics until an OD600 of 0.6 was reached. ONBY at a final concentration of 1 mM dissolved in 1 M NaOH solution was added and the cells were incubated for 15 min before induction of expression using 400 µM IPTG and/or 0.2 % arabinose. Expression was carried out for 3 h at 37 °C. After expression, cells were harvested by centrifugation. For purification of whole cell extracts, cell pellets were resuspended in 10-15 mL Ni-NTA buffer (300 mM NaCl, 50 mM Tris/HCl, pH 8) and cells were disrupted using an Emulsiflex C5 (Avestin). After centrifugation, 20 mM imidazole was added to the cleared supernatant. For purification of periplasmatic protein fractions, cold osmotic shock was used for disruption of the outer cell membrane.[2] Here, cells pellets were resuspended in 150 mL ice cold cell fractioning buffer 1 (0.2 M Tris/HCl, 200 g/L sucrose, 0.1 M EDTA, pH 8) and shaken at 4 °C for 20 min. Following a centrifugation step, the pellets were resuspended in 30 mL cell fractioning buffer 2 (0.01 M Tris/HCl, 0.005 M $MgSO_4xH_2O$, 0.2 % SDS (v/v), 1 % Triton X-100 (v/v), pH 8) and shaken for 15 min at 4 °C. After centrifugation, the supernatant was filtered and dialyzed against Ni-NTA-buffer containing 20 mM imidazole. In each case, Ni-NTA chromatography was applied for protein purification. The solutions were loaded on a gravity flow column with approximately 1 mL bead volume. After washing, proteins were eluted with Ni-NTA buffer containing 250 mM imidazole. Protein fractions were pooled and dialyzed three times against splice buffer (300 mM NaCl, 50 mM Tris/HCl, 1 mM EDTA, pH 7). 2 mM DTT was added during the first dialysis step and 10 % Glycerol (v/v) was added during the third step. Protein aliquots were frozen in liquid nitrogen and stored at -80°C until further use.

Splicing assays and UV irradiation

Splicing reactions were carried out in splice buffer at 37 °C containing 2 mM TCEP. 30 µM Int^C constructs and 10 µM Int^N constructs were used. 30 sec UV irradiation (365 nm) was performed using an LED UV-lamp (1400 mA M365LP1 lamp and DC2200 LED Driver, Thorlabs, NJ, USA). The Int^N(ONBY) construct either was irradiated before the Int^C construct was added to the assay mixture or both proteins were mixed fix and then irradiated. Both intein fragments were incubated at 37 °C for 10 min before mixing. The reactions were quenched at the indicated time points by mixing aliquots with 4X SDS sample buffer containing β-mercaptoethanol and heating for 15 min at 98 °C.

Densitometric analysis

The software GelAnalyzer2010a (gelanalyzer.com) was used for calculating intensities of protein bands in Coomassie-stained SDS gels. Splice product formation was calculated following normalization according to its molecular weight and taking the decrease of Int^N precursor protein into account.

LC-MS analysis of intact proteins

MS data of intact proteins were obtained using an UltiMate™ 3000 RS system (Thermo Fisher Scientific Inc., MA, USA) connected to a maXis II UHR-TOF LC-MS system (Bruker Corp., MA, USA) with a standard ESI source (Apollo, Bruker Corp., MA, USA). 8 µL of the acidified protein solution were injected onto an AdvanceBio RP-mAb column (ID 2.1 mm x L 50 mm, 3.5 µm, C4, Agilent Technologies, CA, USA), desalted for 7 min with 95 % mobile phase A (0.1% formic acid in water) and 5 % mobile phase B (0.1 % formic acid in acetonitrile) and subjected to a fast segmented gradient (5 % to 60 % B in 2 min, 2 min at 60 % and then 60 % to 100 %B in 1 min). MS parameters: positive mode, end plate offset 500 V, capillary voltage 4.5 kV, nebulizer pressure 5.0 bar, dry gas 9.0 l min^{-1}, drying temperature 200°C, mass range m/z 300 to m/z 3000. The resulting spectra were analyzed with DataAnalysis 4.4 (Bruker Daltonik GmbH) and deconvoluted with the implemented MaxEnt algorithm.

In-gel tryptic digestion and LC-MS/MS analysis of proteins

For in-gel tryptic digestion, protein bands were cut from Coomassie-stained SDS gels and a protocol including ProteaseMAX™ Surfactant (Promega, WI, USA) was used as previously described.[3] The employed LC-MS consisted of an UltiMate™ 3000 RSLCnano system (Thermo Fisher Scientific Inc., MA, USA) connected to a maXis II UHR-TOF LC-MS system with a nano-ESI source (CaptiveSpray with nanoBooster, Bruker Corp., MA, USA). After acidification, tryptic peptides were loaded on a C18 trapping column (Acclaim PepMap100, 5 µm, 100 Å, ID 300 µm x L 5 mm, Thermo Scientific) at a flow rate of 20 µL/min in 98 % mobile phase A (0.1 % formic acid in water) and 2 % mobile phase B (0.1 % formic acid in acetonitrile) and washed for 5 min. A segmented gradient was applied (2 % B from 0 to 10 min, 2 % to 60 % B in 50 min) at a flow rate of 300 nL min^{-1} for the separation on a C18 nano column (Acclaim™ PepMap™ 100 C18, 2 µm, 100 Å, ID 0.075 mm x L 250 mm, Thermo Scientific). MS settings: capillary voltage 1.2 kV, nanoBooster with acetonitrile and 0.2 bar N_2, dry gas 3.0 l min^{-1}, dry temperature 150°C and mass range m/z 150 to m/z 2200. MS spectra were recorded every 3 s and CID was employed for data-dependent MS/MS acquisition for the ten most abundant precursor ions. DataAnalysis 4.4 (Bruker Daltonik GmbH) was used for chromatogram processing and peak list generation. ProteinScape 4.0 (Bruker Daltonik GmbH) served for post-run calibration with the ScoreBooster tool, for sequence verification and for the identification of modifications by submission to an in-house Mascot server (Matrix Science) and search against a focused database.

References

1. a) A. Deiters, D. Groff, Y. Ryu, J. Xie and P. G. Schultz, *Angew Chem Int Ed Engl*, 2006, **45**, 2728-2731; b) J. K. Böcker, K. Friedel, J. C. Matern, A. L. Bachmann and H. D. Mootz, *Angew Chem Int Ed Engl*, 2015, **54**, 2116-2120.
2. H. C. Neu and L. A. Heppel, *J Biol Chem*, 1965, **240**, 3685-3692.
3. S. V. Saveliev, C. C. Woodroofe, G. Sabat, C. M. Adams, D. Klaubert, K. Wood and M. Urh, *Anal Chem*, 2013, **85**, 907-914.