Supporting Information

In Cellulo Protein Semi-Synthesis from Endogenous and Exogenous Fragments Using the Ultra-Fast Split Gp41-1 Intein

Maniraj Bhagawati, Simon Hoffmann, Katharina S. Höffgen, Jacob Piehler, Karin B. Busch, and Henning D. Mootz*

anie_202006822_sm_miscellaneous_information.pdf
anie_202006822_sm_Video_S1.avi
| Table of contents                  | page |
|-----------------------------------|------|
| Figure S1                         | S2   |
| Figure S2                         | S3   |
| Figure S3                         | S4   |
| Figure S4                         | S5   |
| Figure S5                         | S5   |
| Figure S6                         | S6   |
| Figure S7                         | S7   |
| Figure S8                         | S8   |
| Figure S9                         | S9   |
| Figure S10                        | S10  |
| Figure S11                        | S11  |
| Figure S12                        | S12  |
| Figure S13                        | S13  |
| Figure S14                        | S14  |
| Figure S15                        | S15  |
| Table S1                          | S16  |
| Table S2                          | S19  |
| Materials and Methods             | S20  |
| Supporting References             | S24  |
SUPPORTING FIGURES

Figure S1. Protein trans-splicing. (A) Schematic representation of the protein trans-splicing (PTS) reaction. (B) Schematic representation of the strategy for labeling proteins of interest (POI) on the N- or C-terminus. Int^N & Int^C = N- and C-terminal intein fragments; Ex^N & Ex^C N- and C-terminal exteins.
Figure S2. Validation of splice activity for C-terminal fluorescent labeling using the Gp41-1 intein. (A) Schematic representation of the assay. (B) SDS-PAGE analysis of the splicing reaction that transfers the fluorescein-labeled CysTag on 7a to MBP from precursor protein MBP-IntN-H6 to give the splice product MBP-CysTag(fluorescein) (SP). Note that only around 70% of 7 was labeled with fluorescein to give 7a (see second lane on the coomassie stained gel), which resulted in two versions of the splice product: MBP-CysTag(fluorescein) (SP) and MBP-CysTag (SP*). (C) Reaction scheme for labeling of an anti-EGFR nanobody (VHH)[1] using the Gp41-1 intein. (D) SDS-PAGE analysis of the labeling and purification steps for the generation of the fluorescein-labeled nanobody (SP). FT represents the flow through fraction from Streptactin resin. Due to the fact that around 70% of 7 was labeled with fluorescein, two versions of the nanobody splice product were formed: VHH-CysTag(fluorescein) (SP) and VHH-CysTag (SP*). (E) Confocal microscopy images of HeLa cells transiently transfected with HA-EGFR-mCherry (HA= hemagglutinin A tag) that were treated with 5 nM of the fluorescein-labeled nanobody (SP). Scale bars represent 20 µm.
Figure S3. Validation of intracellular splice activity of the Gp41-1 intein. Schematic description (A) and microscopy image (B) of Hela cells transiently transfected with a plasmid encoding Tom20-Int\textsuperscript{N}-mCherry-NLS (8). Schematic description (C) and microscopy image (D) of a Hela cell transiently transfected with a plasmid encoding Int\textsuperscript{C}-eGFP. Schematic description (E) and microscopy image (F) of Hela cells transiently transfected with two plasmids encoding Tom20-Int\textsuperscript{N}-mCherry-NLS (8) and Int\textsuperscript{C}-eGFP. Scale bars represent 20 µm.
Figure S4. Bead loading into HeLa cells. Representative fluorescence and brightfield microscopy images of HeLa cells that were transduced with SBP-Smt3-IntC-CysTag(fluorescein) (7a) via bead loading. Scale bars represent 50 μm.

Figure S5. Quantification of intracellular fluorescence intensity in a HeLa cell undergoing splicing (A) Overlay of the fluorescein and mCherry fluorescence signals of a cell expressing Tom20-IntN-mCherry-NLS (8) that was transduced with SBP-Smt3-IntC-CysTag(fluorescein) (7a). This is the same cell that is marked as cell 3 in Fig. 4B. (B) Quantification of the fluorescence intensity along the white line shown in A. The light gray box denotes the cytosolic region, while the darker gray box denotes the nuclear region. The regions with high fluorescein fluorescence intensity that are not within these boxes are the mitochondria. The scale bar represents 10 μm.
Figure S6. C-terminal labeling of Tom20 with fluorescein in HeLa cells using 7a. Additional examples for representative microscopy images of cells expressing 8 that were transduced with 7a. These cells show the same distribution of fluorophores as of cell type 3 in Fig. 4 A-C. Insignificant levels of residual mCherry fluorescence (red channel) can be seen in all cells, except for the last one. This observation indicates that close to complete protein trans-splicing occurred in most of the cells. The scale bars represent 10 µm.
**Figure S7. Quantification of cytosolic fluorescence background with SBP-Smt3-Int\textsuperscript{C}-CysTag\textsuperscript{fluorescein} (7a)** (A, C, E) Fluorescence microscopy images (fluorescein channel) of HeLa cells not expressing 8 that were transduced with 7a. The three cells are the same as shown in Fig. 4C (i-iii). (B, D, F) Quantification of the fluorescence intensity along the white line shown in A, C, E respectively. The light gray box denotes the cytosolic regions, while the darker gray box denotes the nuclear regions. The highest fluorescence intensity is observed in the nucleus and is set to a value of 1 via normalization. (G, I, K) Fluorescence microscopy images (fluorescein channel) of HeLa cells expressing Tom20-Int\textsuperscript{N}-mCherry-NLS (8) that were transduced with 7a. The three cells are the same as shown in Fig. 4C (iv-vi). (H, J, L) Quantification of the fluorescence intensity along the white line shown in G, I, K respectively. The light gray box denotes the cytosolic regions, while the darker gray box denotes the nuclear regions. The regions with high fluorescence intensity that are not within these boxes are the mitochondria. The highest fluorescence intensity is observed on the mitochondria and is set to a value of 1 via normalization. The relative background fluorescence intensity in the cytosol varies from approximately 0.3 to approximately 0.1 as the relative abundance of 7a over 8 decreases (evident from the fluorescence intensity in the nucleus). Scale bars represent 20 µm.

**Figure S8. C-terminal labeling of Tom20 with fluorescein using the Gp41-1 intein.** SDS-PAGE analysis followed by fluorescence imaging and western blotting of lysates from HeLa cells expressing 8 that were transduced or not transduced with 7a. See Figure 4 for description of the constructs.
Figure S9: C-terminal labeling of CytERM using the Gp41-1 intein. CytERM localizes in the outer leaflet of the endoplasmic reticular membrane. We created stable HeLa cell lines expressing CytERM-IntN-mCherry-NLS. (A) Schematic representation of the protocol for labeling CytERM with fluorescein. (B) Representative microscopy images showing labeling of CytERM with fluorescein via intein trans-splicing in HeLa cells. The overlay panel indicates the two types of cells present in this image field: Cell 1, expressing CytERM-IntN-mCherry-NLS but not transduced with 7a; cell 2, expressing CytERM-IntN-mCherry-NLS and transduced with 7a. Scale bars represent 20 µm.
Figure S10. Validation of splice activity for N-terminal fluorescent labeling using the Gp41-1 intein. (A) Schematic representation of the assay. p-azidophenylalanine (AzF) was incorporated via amber stop codon suppression in E. coli into a short tag sequence (ANN(AzF)TKGSTSRY) to give the protein ClickTag-IntN-H6 and coupled via copper catalyzed alkyne-azide cycloaddition (CuAAC) to an Alexa647-alkyne dye. (B) SDS-PAGE analysis of the protein trans-splicing reaction that transfers the Alexa647-labeled ClickTag to Trx-H6 (Trx = thioredoxin) to give the splice product AzFTag (Alexa647)-Trx-H6 (SP).
Figure S11. N-terminal labeling of Tom22 using the Gp41-1 intein. (A) Schematic representation of the protocol for labeling Tom22 with Alexa647. Tom22 is another protein of the mitochondrial TOM complex and has its N terminus facing to the cytosol. (B) Representative microscopy images showing labeling of Tom22 with Alexa647 via protein trans-splicing in HeLa cells. The overlay panel indicates the two types of cells present in this image field: Cell 1, expressing NLS-mCherry-IntC-Tom22 but not transduced with ClickTag(Alexa647)-IntN-H6; and cell 2, expressing NLS-mCherry-IntC-Tom22 and transduced with ClickTag(Alexa647)-IntN-H6. Scale bars represent 20 µm.
Figure S12. Quantification of cytosolic fluorescence background with NLS-SBP-Smt3-IntC-CysTag(fluorescein) (9a) (A, C, E) Fluorescence microscopy images (fluorescein channel) of HeLa cells not expressing 8 that were transduced with 9a. The three cells are the same as shown in Fig. 4F (vii-ix). (B, D, F) Quantification of the fluorescence intensity along the white line shown in A, C, E respectively. The light gray box denotes the cytosolic regions, while the darker gray box denotes the nuclear regions. The highest fluorescence intensity is observed in the nucleus and is set to a value of 1 via normalization. (G, I, K) Fluorescence microscopy images (fluorescein channel) of HeLa cells expressing Tom20-IntN-mCherry-NLS (8) that were transduced with 9a. The three cells are the same as shown in Fig. 4F (x-xii). (H, J, L) Quantification of the fluorescence intensity along the white line shown in G, I, K respectively. The light gray box denotes the cytosolic regions, while the darker gray box denotes the nuclear regions. The regions with high fluorescence intensity that are not within these boxes are the mitochondria. The highest fluorescence intensity is observed on the mitochondria and is set to a value of 1 via normalization. In comparison to our observations when labeling with 7a, which lack the NLS sequence (Fig. S6), the relative background fluorescence intensity in the cytosol when using 9a does not exceed approximately 0.1 despite changes in the relative abundance of 9a over 8 (evident from the fluorescence intensity in the nucleus). Scale bars represent 20 µm.

Figure S13. Quantification of intracellular fluorescence intensity in a HeLa cell undergoing splicing (A) Overlay of the fluorescein and mCherry fluorescence signals of a cell expressing Tom20-IntN-mCherry-NLS (8) that was transduced with NLS-SBP-Smt3-IntC-CysTag(fluorescein) (9a). This is the same cell that is marked as cell 3 in Fig. 4E. (B) Quantification of the fluorescence intensity along the white line shown in A. The light gray box denotes the cytosolic region, while the darker gray box denotes the nuclear region. The regions with high fluorescein fluorescence intensity that are not within these boxes are the mitochondria. The scale bar represents 10 µm.
Figure S14. C-terminal labeling of Tom20 with fluorescein in HeLa cells using 9a. Additional examples of representative microscopy images of cells expressing 8 that were transduced with 9a. These cells show the same distribution of fluorophores as of cell type 3 in Fig. 4 D-F. Insignificant levels of residual mCherry fluorescence (red channel) can be seen in all cells, except for the last one. This observation indicates that close to complete protein trans-splicing occurred in most of the cells. The scale bars represent 10 µm.
Figure S15: Analysis of single particle tracking of Alexa647-labeled Tom20. Jump distance distributions of Tom20 with different time delays ($dt = 32$ ms). The black curves show the fits to the data using a two-state model.
### SUPPORTING TABLES

#### Table S1: List of recombinantly produced proteins used in this study

| Name of construct (protein number) | Encoding plasmid | Vector backbone | Sequence |
|-----------------------------------|------------------|-----------------|---------|
| **H$_6$-Smt3-Int$^C$(N125A)-CysTag-Trx-H$_6$ (1)** | pJKB122 | pBAD | MGSSHHHHHHHHSGGLVPRGSASMSDSENVQEAKEV<RFKVPVPEPHILNLKVSQDSSEIFKKTTPRLRMLAEAFKQGKEMDSLRLFYDGIQR ADQTPELDMEEDNDIEAHREIQGGMLKKILKIIELEDRELVIDGEV SGNLFYANDILTHASSSDCOTGSDKIIHLTDDSDTDVLKAGDAIL LVDFNAWCGPCKMIAPILDEIADEYQKLLLLTAKLNIHINHPGTAPKY GIRQIPULLFLKNGEVAATKVQALSKQLKELFDANSLAGSEQHHHHHH |
| **MBP-Int$^N$(C1A)-H$_6$ (2)** | pAB101 | pMal-c2X | MKTEEGKLV1WINGDGKYNLAEVGGKFEKOTGIKVTIVHPDKLEEK FFQVAATGDGPIDDWADHROPFGYQSGLLEAITPDKAFQDKLYFYPT WDAVRNYKLIAYPIAEVSLAYNDLLNPKTVTTIELDALKELK AKGKSAIMFNLQEYFTWPLIAADGGYAFKENGYDKDYGVDVNAAG AKAGLTFVLVDIKKKNNADTDYIAAENKGETATMINGPKNAWSN IRTSQVNYGTVLPTFQGQSSPFGVQVSAGIAAASPHELKEFLE NYLTDTEGALVENKDLKPLGAVALNYELAATMEMAQKGE IMPN1PQMSAFWYAVRTAVINAAAGQTYDEALKDAQTINSSNINNN NNNNNLGIEGRKELFVQRSTGRSGLALKTQTQPGMEISNIQ VGDVLVSLNTYNEVNLNVFPKXKKKSYKILEDGEKIECSEEHLFFTQ TGEMNISGGLKEGMCYVEGGHHHHHH |
| **H$_6$-Int$^N$(C1A)-mCherry (3)** | pMBH81 | pET28 | MGHIHIHIHGGSGSSKLILSRGSTSGYALDKTVQTPQGMEISNIQ VQGLDVLSNTGYNELVNVFSSKKSKYKIDLEDGKEICSEEINLFPT QTQEMNHSGGLKEGMCYVEGGAAASSATMSSKGEEDNMMAIKEFMP RFKVKHMESVNGHEFIEEIEGEGFYQYATKLVTKGGPLFPFAWD ILSPFMYGSKAYVKHPADIYYLKLSPFEGKWNFRENFDGGGVT VTQDISSQDQGFIYVKLRGTNPSPGVQMGKSTGMWASSEMRYPE DGALKIEIKQLKDGDGYDAEVTYAXKMKFPVLPGAYNVIKLID ITSHEDETYTIVEQYKAEHNGSGMELYKSSRLSVPITPPFRSLSGC |
| **SBP-Int$^C$(N125A)-eGFP (4)** | pMBH82 | pBAD | MGGMMLKIKILKIIELEDRELVGVSQNHLYANDILTHASSSDVVG SHYQAMVSKEELFTGVVPLVELGDGNSKFSVSYSEGEDYAYG KLTLLKFICCTTGKLFVPWPTLVTTLGVTQCFSRYDPNMQHDFKKSA MFEGYVQERTIFKKDDGYKTRAERVFGEDETNLVRIELKIGDFKEGD NLGKHLELYNHNVYIMADQKNDIKVNNFVHIINIIEGDSVQADH YOQNTP1GDGPVLLPDNHYLSTQALSIDKDEMRKIVILEEVTATAG ITLGDMELYKRRRYTMDEETGWRGGHVVEAGLEQLARLEHH PQQKREPP |
| **H<sub>6</sub>-Int<sup>N</sup>-mCherry** (5) | pMBH77 A | pET28 | $pMBH77 A$ | $pET28$ |
|---|---|---|---|---|
| **SBP-Int<sup>C</sup>-eGFP** (6) | pMBH78 _1 | pBAD | $SBP-Int<sup>C</sup>-eGFP$ (6) | $pMBH78 _1$ | pBAD |
| **SBP-Smt3-Int<sup>C</sup>-CysTag** (7) | pMBH56 | pET16b | $SBP-Smt3-Int<sup>C</sup>-CysTag$ (7) | $pMBH56$ | pET16b |
| **NLS-SBP-Smt3-Int<sup>C</sup>-CysTag** (9) | pMBH71 | pET16b | $NLS-SBP-Smt3-Int<sup>C</sup>-CysTag$ (9) | $pMBH71$ | pET16b |
| **MBP-Int<sup>N</sup>-H<sub>6</sub>** | pAB64 | pMal-c2X | $MBP-Int<sup>N</sup>-H<sub>6</sub>$ | $pAB64$ | pMal-c2X |
| **VHH-Int<sup>N</sup>-SBP** | pSH80 | pGDR11 | $VHH-Int<sup>N</sup>-SBP$ | $pSH80$ | pGDR11 |
| ClickTag-Int\(^N\)-H\(_6\) | pKSH04 | pET28a | MANNNXTRGSTRSGYCLDLRTQVQTPQGMKEISNIQVGDVLVSNTGYNEVLNVFFKSKKSYKITYLEDGKEICSEEHLPFTQTGEMNISGGGLKEMCLYVKEGGSGSDYKDDDDOKHHHHHHHH |
| H\(_6\)-Smt3-Int\(^C\)-Trx-H\(_6\) | pAB74 | pBAD | MGSSHHHHHHHHGSLVPRGSASMSDSEVNVSEQKEAKPEVKPEVFKPETHINLKVSDGSSEIFFKIKKTTPLRRMLEAFAKRQGKEMDSLRFYLYDGIRIQADQTEOLMDNIDIEAHREQIGGMMLKIKIEELDERELIDIEVSGNHLFYANDILTHNSSSDVCGDSDKITILHLDSSFDTDVLKADGAILVDFWAWHCPCMIAPILDEIADEYQGKLTVAKLNIDHNPGTAPKYGIRGIPTLLLFKNGEVAATKVGLSKQQLKEFLDANLAGSEFRSHHHHH |
| Name of construct (number) | Encoding plasmid | Vector backbone | Sequence |
|---------------------------|-----------------|-----------------|----------|
| **Tom20-Int\(^N\)-mCherry-NLS (8)** | pMBH57 | pSems | MVGRNSIAAIAGVCALPGYCIYFDRKRRSDPNFKNRKRERKKQLAK ERAGLSKLDPDKDAEAVQFKLLREILQGEELLAQGEYEKGVHDHTNAI A VCCQQQLQLQVLQTLPPVTQMLTTLQPTISQRHVSWSQALAEDEDFEF TRSGLCLDLKTQVTPQGKEISNIVQGDLVSNTGNYEVMVFPSKK KSYKITLEGDIEITCSEEHLFPTQGEMNISGLGEGMCILVYKEGGA ST MVSKGEEDNMAIKFMRKVIHMEGSVNGHEFEIEEGEGREETQTA KLKVTGKGPFLFAWDSQPQMQSGKAYHVPADPYYLKLSSFGFNMK RNENFEDGTVTQDSSLQOQFYIKLUKRNFTSDGFMQKKTMTG WEASESRSVFEDGALGPEIKRKLDKQHYYADEVKTVKPKVQVPG AYNVNIKDLITSHNEDYTIVEQYERAERGRSTGMDELYKSGLRSRGSL EASAGSAGSAGSAASAGGIPAAKVRKLD |
| **Int\(^C\)-eGFP** | pMBH22 | pcDNA4 B | MMLKIKILELDREDLIDIEVSQGHDLYANIDILTNNSSSDVGSHYQLA MVSKGEELPTGVVIPWEQDGVDVNGKFSVSGEGRDATGYKTLFIC TTGKLVPWNPLTTLTGQCFSRDPHKQDDFKSAEMEGYQERT IFKKoddGYNKTRAEVKFEGLTVNRELKIDFREDGNIHGKHLEYYN SHNYYMADQKNGKIMVRNHDGSDLASHYQMTPQDGPVV LDDYLSQDALSIPKPKRVDHMLLEFVTGAAGTILGMDELYKMQVYDV PDLNSADIQHSSGRSLMGPRQKLMEEEDLMHHTGHHHHH |
| **CytERM-Int\(^N\)-mCherry-NLS** | pMBH70 | pSems | MDPPVVLGCLSLCLLLSLWKKQSGKKLETRGSYCLBLKTQVQTPQG MEISNIQDVGDLVSLNTGNYEVMVFPSKKKSYKITLEGDIEITCSEE HLFTPTQGEMNISGLGEGMCILVYKEGGA STMVSKGEEDNMAIKFMR FKVHMESGSGHEFEIEEGEGREETQTAKLKVTGKGPFLFAWDSQPQ MQGYSGKAYKHPADIPDLKLSSFGFNMKRNENFEDGTVTQDSSLQOQFYIKLRLQKRLDKQHYYADEVKTVKPKVQVPG AYNVNIKDLITSHNEDYTIVEQYERAERGRSTGMDELYKSGLRSRGSL EASAGSAGSAGSAASAGGIPAAKVRKLD |
| **NLS-mCherry-Int\(^C\)-Tom22** | pKSH03 | pSems | MPAKVRKVDVTYRASSMVKGEEDNMAIKFMRKVIHMEGSGVNGHEFE IEGEGEGREETFQTAKLKVTGKGPFLFAWDSQPQMQSGKAYKHPA DIPDLKLSSFGFNMKRNENFEDGTVTQDSSLQOQFYIKLRLQKRLDKQHYYADEVKTVKPKVQVPG AYNVNIKDLITSHNEDYTIVEQYERAERGRSTGMDELYKSGLRSRGSL EASAGSAGSAGSAASAGGIPAAKVRKLD |
MATERIALS AND METHODS

General

Restriction enzymes for cloning were purchased from Thermo Scientific and oligonucleotides from Biolegio. Cloned plasmids were verified by DNA sequencing by Seqlab. If not further specified, buffer reagents, antibiotics, media components and other compounds used for cell culture or biochemical assays were purchased from Carl Roth, Applichem, Sigma Aldrich, Fluka or Jena Bioscience.

Protein production and purification

All proteins were produced in *E. coli* BL21(DE3) Gold cells. Cells were cultured at 37 °C in LB-medium with the corresponding antibiotic until an OD_{600} of 0.6 – 0.8 was reached. Protein expression was carried out at 18 °C for 16 h by adding IPTG (0.4 mM) for expression under a T7 promoter, or arabinose (0.2%) for expression under a pBAD promoter. Cell pellets were collected by centrifugation, resuspended in the respective purification buffer and stored at -20 °C till further use. Resuspended cells were ruptured either using an emulsifier or sonicator. Insoluble fractions were removed by centrifugation and the supernatant fractions were used to purify the proteins.

For purification via Ni-NTA affinity chromatography of His-tagged proteins, cell pellets were resuspended in Ni-NTA buffer (50 mM Tris/HCl, 300 mM NaCl, pH 8.0). Purification was performed at 4 °C using flow gravity flow columns with a bed volume of 2 mL of Ni-NTA resin (Cube Biotech). For washing, two steps with Ni-NTA buffer + 20 mM imidazole and one step with Ni-NTA buffer + 40 mM imidazole were performed followed by elution with Ni-NTA buffer + 250 mM imidazole.

For purification via Strep-tactin affinity chromatography of SBP-fused proteins, cells were resuspended in buffer W (100 mM Tris/HCl, 150 mM NaCl, 1 mM EDTA, pH 8.0). Purification was performed at at 4 °C using flow gravity flow columns with a bed volume of 2 mL of Strep-tactin resin (IBA Lifesciences). Before elution the column was washed three times with buffer W. Proteins were eluted with buffer W + 2.5 mM desthiobiotin.

Purified proteins were dialyzed against a splice buffer (50 mM Tris/HCl, 300 mM NaCl, 1 mM EDTA, pH 7.0), and finally dialyzed against splice buffer + 10 % glycerol before flash freezing in liquid nitrogen and storage at -80 °C. Protein concentrations were determined using the calculated extinction coefficient at 280 nm.

Splice assays

Splice assays were performed in splice buffer containing 1 mM DTT at 37 °C. The fluorescently labelled species was used at a concentration of 1 µM and the unlabelled species at a concentration of 5 µM. The reaction was stopped at the described time points by taking an aliquot of the reaction mixture and boiling (5 min, 96 °C) the aliquots in 4x SDS sample buffer (500 mM Tris/HCl, 8 % (v/v) SDS, 40 % (v/v) glycerol, 20 % (v/v) 2-mercaptoethanol, 5 mg mL^{-1} bromophenol blue, pH 6.8).

Microscale thermophoresis

Microscale thermophoresis (MST) was performed using the Monolith NT.115 (NanoTemper, Munich) with splice-inactive intein mutants. The construct MBP-IntN^{C1A}.H6 (2) was used in a dilution series from 500 nM down to 0.015 nM. Alexa647-C2® labeled H6-Smt3-Int^{C[N125A]}.
CysTag-Trx-H6 (1a) was added to each sample \((c_{\text{final}}=0.25 \text{ nM})\) together with BSA \((0.05\% \text{ (w/v))}\) and Tween-20 \((0.1\% \text{ (v/v))}\). The samples were incubated at room temperature for 5 to 10 minutes before loading into standard treated capillaries. Measurements were performed at 25 °C with 80% laser intensity and 20% MST power \((\text{IR laser } \lambda=1480 \text{ nm})\). Binding was quantified using the ratio of \(F_{\text{hot}} \) to \(F_{\text{cold}}\) resulting in \(F_{\text{norm}}\).[2] The change in fluorescence was plotted against the concentration of the concomitant intein fragment on a logarithmic scale and fitted to a sigmoidal curve using Origin2019b (OriginLab, Northhampton) to determine the \(K_D\)-value.

**Fluorescent labeling of CysTag-fusion proteins**

Purified proteins containing CysTags were diluted in splice buffer to a concentration of 25 µM and incubated with EDTA (5 mM) and TCEP (100 µM) for 20 min at 37 °C. The required fluorophore-maleimide conjugate \((\text{Alexa647-C2 Maleimide or Fluorescein-5 maleimide, Thermo Fisher Scientific})\) was added at a final concentration of 200 µM and the reaction mixture was incubated for 2 hours at 25 °C. The reaction was quenched with DTT (2 mM) and the unreacted fluorophore was removed by affinity chromatography on Ni-NTA resin.

**Fluorescent labeling of ClickTag-fusion proteins**

Purified proteins carrying the ClickTag were diluted in labeling buffer \((20 \text{ mM HEPES}, 150 \text{ mM NaCl}, \text{pH 8.0})\) and were incubated with 150 µM Alexa647 alkyne, 0.5 mM CuSO4, 0.1 mM TBTA \((\text{Tris}((1\text{-benzyl-4-triazolyl})methyl)amine), \text{and 0.5 mM TCEP for 30 min at 25 °C})\). The reaction was quenched by addition of EDTA to a final concentration of 10 mM, copper ions and unreacted fluorophore were removed by dialysis against dialysis buffer \((20 \text{ mM HEPES}, 150 \text{ mM NaCl}, \text{pH 8.0})\) and subsequent affinity chromatography on Ni-NTA resin.

**Fluorescent C-terminal labeling of nanobodies using protein trans-splicing of CysTag**

The fluorescently labeled CysTag-fusion protein \((7a) \,(15 \mu\text{M})\) and the nanobody-construct \(\text{VHH-IntN-His}_6\) \((10 \mu\text{M})\) of the anti-EGFR nanobody\([1]\) were mixed in splice buffer and incubated at 37 °C for 1 hour. The reaction mix was loaded on a Ni-NTA-column and the labeled nanobody was collected in the flow through. The bound components were eluted with Ni-NTA-buffer + 180 mM imidazole. The eluted nanobody was dialyzed against splice buffer and stored at 4 °C until further usage.

**Cell culture and transfection**

HeLa cells were cultured in MEM supplemented with Earle’s salts, that was further supplemented with 10% fetal calf serum, 1% non-essential amino acids, 1% L-glutamine and 1% penicillin/streptomycin at 37 °C under 5% CO2. Transfection was carried out using Lipofectamine 2000 (Thermo Fisher Scientific) using the manufacturer’s instructions. Transient transfection was used when specifically indicated.

**Creation of stable cell lines**

HeLa cells for the expression of 8 and CytERM-IntN-mCherry-NLS were stably transfected and used for protein transduction via bead-loading. Cells were split at different concentrations and were grown under selection pressure \((2 \mu\text{g/mL puromycin or 400 µg/mL geneticin G418})\) for two weeks without changing the culture dish. The cells were further cultivated in antibiotic-free medium at 37 °C under 5% CO2 until cell colonies were formed. Colonies were transferred to a 96-well plate by pipetting and split to larger wells when the cell culture reached confluence.
The selection pressure was set back on for three days every two weeks. This process created a mixture of cells with some cells not expressing the intended constructs.

**Bead-loading**

The purified protein was diluted in PBS to a final concentration of 25 µM and was incubated with 0.1% Pluronic® F-68 (Thermo Fisher) for 5 min at 25 °C. The cell culture medium was removed, then the protein-Pluronic® F-68 solution was spread over the cells and glass beads (< 106 µm, Sigma Aldrich) were added. The culture dish was hit on the bench with eight sharp hits before the cells were incubated at 37 °C under 5% CO₂ for 2 min. Glass beads were removed by washing with PBS (5x 1 mL).

**Confocal laser scanning microscopy**

Cells that were transduced via bead loading and labeled by trans-splicing were washed with PBS and fixed with 4% paraformaldehyde for 20 minutes at room temperature. Cells were washed three more times with PBS and the coverslips on which the cells were plated were mounted on slides using Aqua/Poly-Mount mounting solution (Polysciences). Confocal microscopy was carried out using a 63X water-immersion objective lens on a Leica DMI8 system. Intracellular trans-splicing experiments for each pair of complementary intein fragment fusion proteins were carried out in at least 5 independent repeats.

**dSTORM**

Cells that were transduced via bead loading and labeled by trans-splicing were washed with PBS and fixed with 4% paraformaldehyde for 20 minutes at room temperature. Cells were washed three more times with PBS and placed onto coverslip holders. As a redox system to facilitate the switching of the fluorophore between a dark and bright state 0.5 mg/ml glucose oxidase, 40 mg/ml catalase, 5% w/v glucose, 50 mM β-mercaptoethylamine in 1 ml PBS was added to the coverslip. Imaging was carried out in a HiLO mode with an inverted Olympus IX83 microscope equipped with a Quad-line TIRF condenser, a back-illuminated EMCCD camera (Andor iXon Ultra 897), and a UAPON 150× 1.45 NA TIRFM objective. A 642 nm fiber laser (MPB Communications) was used for fluorescence excitation, while a 405 nm diode-pumped solid state laser (CrystaLaser) was used to facilitate the switching of the fluorophores to the dark state. Data was analyzed using the Matlab-based SLIMfast software package developed at the University of Osnabrück.

**Single particle tracking**

Cells that were transduced via bead loading and labeled by trans-splicing were washed with PBS and placed into coverslip holders. Cell culture medium containing 0.5 mg/ml glucose oxidase, 40 mg/ml catalase, 5% w/v glucose, 50 mM β-mercaptoethylamine was added to the coverslip. Single fluorophore tracking data was collected using an inverse microscope (IX71, Olympus) equipped with a TIRF condensor and a CMOS-ORCA camera (Hamamatsu). Alexa647 was excited in the HiLo mode with a 642 nm laser (Omicron LuxX 642-140) using a 60X TIRF objective lens (Olympus). Data was collected with a frame rate of 32 ms/frame, and usually 5000 frames were recorded. Single particle tracking data was analyzed using the Trackmate plugin of FIJI, followed by kinetic modeling using SpotOn. In order to account for the blinking of Alexa647, a maximum frame gap of 5 frames was set in the LAP tracker used within Trackmate.
Surface modification

Surface modification for protein immobilization on a tris-NTA-functionalized PEG polymer brush was carried out as described in detail previously [3]. Transducer glass slides coated with a thin silica layer for reflectance interference (RIf) detection were cleaned with a plasma oxidizer, rinsed with MilliQ water and dried in a nitrogen stream. The surfaces were then functionalized with glycidyloxypropyltrimethoxysilane. Subsequently, melted diamino-PEG (2000 g/mol) was reacted with the epoxy groups on the surface at 75 °C for 4 hours. For functionalization with tris-NTAcarboxylic acid (tris-NTA), 5 µL of a solution of tris-NTA in chloroform (100 mg/ml) was applied on the surface. Coupling of the carboxy group with the amine modified surface was started by addition of 5 µL of N,N'- diisopropylcarbodiimide and the reaction was carried out by incubation for 1 h at 75 °C. The excess reaction mixture was washed off with chloroform after the reaction. The substrates were then incubated in trifluoroacetic acid overnight to cleave of the tert-butyl esters and free the NTA head groups.

Real-time solid phase detection

Association of intein fragments and splicing were monitored using simultaneous real-time detection by reflectance interference (RIf) and total internal reflection fluorescence spectroscopy (TIRFS). The measurements were performed under continuous flow-through conditions using home-built set-ups described earlier [4]. mCherry was excited with a 532 nm laser. The excitation laser power was ~100µW illumination a 1 mm spot, excluding photobleaching during binding experiments. All binding experiments were carried out in HBS (20 mM HEPES pH 7.5 and 150 mM sodium-chloride) supplemented with 0.01% Triton X-100. Prior to binding experiments, tris-NTA functionalized surfaces were sequentially treated with 20 mM EDTA, 10 mM NiCl₂ and 250 mM imidazole. This was followed by an injection of 250 nM of the described His-tagged IntN fusion protein (state I in Figures 2 C,D and 3 A,B,C). This was followed by an injection of the described IntC fusion protein at the shown concentrations for a fixed period (state II), and a subsequent period of wash with the running buffer (state III). After each wash phase, 250 mM imidazole was injected to remove all bound proteins, and then the entire cycle was repeated for further measurements.

Derivation of equation for fitting on-surface splice kinetics

\[
[m\text{Cherry}]_t = [A]_t + [A :: B]_t
\]

where \([A]_t\) is the concentration of the precursor H₆-Intⁿ⁻mCh and \([A :: B]_t\) is the concentration of the complex between H₆-Intⁿ⁻mCh and SBP-Intᶜ⁻eGFP at a time point \(t\) after introduction of SBP-Intᶜ⁻eGFP into the flow cell.

Furthermore,
\[ [A]_t = [A]_0 - ([A \leftrightarrow B]_t + [D]_t) \] (2)

where \([A]_0\) is the total concentration of H6-IntN-mCh immobilized on the surface before the introduction of SBP-IntC-eGFP at time point \(t = 0\), and \([D]_t\) is the concentration of IntN-mCh that is produced during the splice reaction at a time point \(t\) after introduction of SBP-IntC-eGFP into the flow cell.

From equations (1) and (2), we get
\[ [mCherry]_t = [A]_0 - [D]_t \] (3)

During the wash phase of the experiment no further association of the precursors to form the complex occurs. However, dissociation of the complex, as well as splicing continues to occur.

Therefore, the rate of change of the complex,
\[ \frac{d[A \leftrightarrow B]_t}{dt} = -k_d[A \leftrightarrow B]_t - k_s[A \leftrightarrow B]_t \]

where \(k_d\) and \(k_s\) are the dissociation and splicing rate constants respectively.

This solves to
\[ [A \leftrightarrow B]_t = [A \leftrightarrow B]_0 e^{-(k_d+k_s)t} \] (4)

where \([A \leftrightarrow B]_0\) is the concentration of the complex before the start of the wash phase.

Rate of formation of IntN-mCh that is produced during the splice reaction occurring during the wash phase of the experiment,
\[ \frac{d[D]_t}{dt} = k_s[A \leftrightarrow B]_t \]

Using equation (3) and solving,
\[ [D]_t = -[A \leftrightarrow B]_0 k_s (k_d + k_s) e^{-(k_d+k_s)t} + C \] (5)

where \(C\) is a constant of integration.

Using equations (3) and (5),
\[ [mCherry]_t = [A]_0 + [A \leftrightarrow B]_0 k_s (k_d + k_s) e^{-(k_d+k_s)t} - C \]

Therefore, fitting the decay of the mCherry derived fluorescence signal during the wash phase of the experiment using an exponential function provides the quantity \(k_d + k_s\). Knowing \(k_d\) from the experiments done using splice inactive mutants, the splicing rate constant \(k_s\) can be calculated.

**SUPPORTING REFERENCES**

[1] K. R. Schmitz, A. Bagchi, R. C. Roovers, P. M. van Bergen en Henegouwen, K. M. Ferguson, *Structure* 2013, 21, 1214-1224.

[2] S. A. Seidel, P. M. Dijkman, W. A. Lea, G. van den Bogaart, M. Jerabek-Willemsen, A. Lazic, J. S. Joseph, P. Srinivasan, P. Baaske, A. Simeonov, I. Katritch, F. A. Melo, J. E. Ladbury, G. Schreiber, A. Watts, D. Braun, S. Duhr, *Methods* 2013, 59, 301-315.
[3]  a) J. Piehler, A. Brecht, R. Valiokas, B. Liedberg, G. Gauglitz, *Biosensors & Bioelectronics* **2000**, *15*, 473-481; b) S. Lata, J. Piehler, *Anal Chem* **2005**, *77*, 1096 -1105.

[4]  a) M. Gavutis, S. Lata, J. Piehler, *Nature Protocols* **2006**, *1*, 2091-2103; b) M. Gavutis, S. Lata, P. Lamken, P. Muller, J. Piehler, *Biophys J* **2005**, *88*, 4289-4302.