Research Article

Journal of PeptideScience

Received: 3 February 2010 Revised: 10 March 2010 Accepted: 25 March 2010 Published online in Wiley Online Library: 28 June 2010

(wileyonlinelibrary.com) DOI 10.1002/psc.1243

Protein *trans*-splicing on an M13 bacteriophage: towards directed evolution of a semisynthetic split intein by phage display‡

Daniel Garbe,[a,b] Ilka V. Thiel[a] and Henning D. Mootz[a]*

Split inteins link their fused peptide or protein sequences with a peptide bond in an autocatalytic reaction called protein *trans*-splicing. This reaction is becoming increasingly important for a variety of applications in protein semisynthesis, polypeptide circularisation, construction of biosensors, or segmental isotopic labelling of proteins. However, split inteins exhibit greatly varying solubility, efficiency and tolerance towards the nature of the fused sequences as well as reaction conditions. We envisioned that phage display as an *in vitro* selection technique would provide a powerful tool for the directed evolution of split inteins with improved properties. As a first step towards this goal, we show that presentation of active split inteins on an M13 bacteriophage is feasible. Two different *C*-terminal intein fragments of the *Ssp* DnaB intein, artificially split at amino acid positions 104 and 11, were encoded in a phagemid vector in fusion to a truncated gpIII protein. For efficient production of hybrid phages, the presence of a soluble domain tag at their *N*-termini was necessary. Immunoblot analysis revealed that the hybrid phages supported protein *trans*-splicing with a protein or a synthetic peptide, respectively, containing the complementary intein fragment. Incorporation of biotin or desthiobiotin by this reaction provides a straightforward strategy for future enrichment of desired mutants from randomised libraries of the *C*-terminal intein fragments on streptavidin beads. Protein semisynthesis on a phage could also be exploited for the selection of chemically modified proteins with unique properties.

Keywords: protein semisynthesis; *Ssp* DnaB intein; *in vitro* selection; phage display

Introduction

There is a great demand for chemically tailored proteins for a variety of applications in biotechnology, biomedicine and basic research. A number of methodologies have been developed which allow the preparation of a protein of interest by linking two or more peptide fragments with a native peptide bond. In these approaches, the peptides can be of either synthetic or recombinant origin. Established chemical ligation strategies are the Native Chemical Ligation (NCL) [1], Expressed Protein Ligation (EPL) [2,3] and the traceless Staudinger ligation [4]. The ligation can also be performed enzymatically by reverse proteolysis techniques [5,6] or the sortase-mediated ligation [7]. Finally, split inteins join two peptide or protein segments (the *N*- and *C*-terminal exteins) in the autocatalytic protein *trans*-splicing reaction [8–10]. The synthetic peptides in these approaches can include one or more chemical moieties and structures of virtually any kind, both in the side chains and backbone. This fact provides one main advantage of the above-mentioned technique over other techniques to introduce synthetic groups into proteins, like bioconjugation reactions on side chain functionalities or manipulation of the protein biosynthesis machinery [11–15].

The split intein approach presents several beneficial properties for the semisynthesis of proteins from a synthetic and a recombinant segment (see Figure 1(A) as example) [16]. Both the *N*- and *C*-terminal intein fragments, Int^N and Int^C, can be obtained either synthetically or by recombinant protein expression. It is neither necessary to introduce any functional groups into these fragments, like thioesters or azido groups, nor to add thiols at high concentrations (as for Native Chemical Ligation (NCL) and Expressed Protein Ligation (EPL)) or additional enzymes that perform the catalysis. The intein fragments usually exhibit high affinity for each other, and the autocatalytic protein splicing reaction can thus be performed at low reactant concentrations and even in complex mixtures [17] or in living cells [9]. As the intein fragments are removed in the course of protein splicing, the reaction can be designed in a traceless way. Several split

* *Correspondence to: Henning D. Mootz, Fakultät Chemie – Chemische Biologie, Technische Universität Dortmund, Otto-Hahn-Strasse 6, D-44227 Dortmund, Germany. E-mail: henning.mootz@tu-dortmund.de*

a *Fakultät Chemie – Chemische Biologie, Technische Universität Dortmund, Otto-Hahn-Strasse 6, D-44227 Dortmund, Germany*

b *Present address: Chair of Chemistry of Biogenic Resources, Technische Universität München, Straubing Center of Science, Schulgasse 16, D-94315 Straubing, Germany*

‡ *Special issue devoted to contributions presented at the Symposium "Probing Protein Function through Chemistry", 20–23 September 2009, Ringberg Castle, Germany.*

Abbreviations used: *aa, amino acid(s); c myc, c-myc epitop (EQKLISEEDL, residues 410–419 of the product of the human oncogene c-myc); DTT, dithiothreitol; Fmoc, 9-fluorenylmethyloxycarbonyl protection group; gpD, headprotein D of λ-phage; HBTU, 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate; His_6, hexahistidine sequence; HRP, horseradish peroxidase; IPTG, isopropyl 1-thio-β D-galactopyranoside; MBP, maltose binding protein; MOI, multiplicity of infection; NTA, nitrilotriacetic acid; OD_{600}, optical density at 600 nm; PelB, signal sequence for pectate lyase B from Erwinia carotovora; PVDF, polyvinylidene fluoride; SPPS, solid phase peptide synthesis; StrepII, sequence of the StrepII tag (WSHPQFEK); Trx, thioredoxin.*

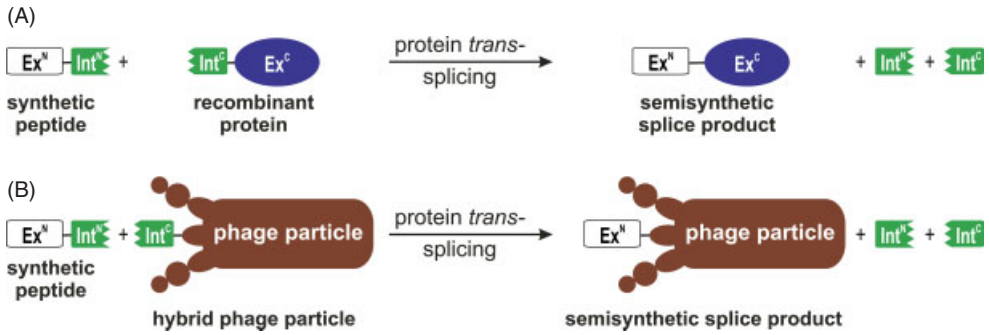

Figure 1. Semisynthetic protein *trans*-splicing for the *N*-terminal chemical modification of proteins. The Int^N fragment of the *Ssp* DnaB intein is only 11 aa in size and can be synthesised by SPPS. (A) Splicing with an Int^C fusion protein. (B) Splicing with Int^C presented on a phage.

inteins have now been reported of which one of the intein fragments is of only 6 to 36 amino acids in size [8,10,18,19], and thus accessible by SPPS, solid phase peptide synthesis. They are spontaneously active under native conditions without the need for a denaturation and a renaturation step. Protein *trans*-splicing can in principle be a highly efficient reaction, as demonstrated by the biochemical analysis of the *Npu* DnaE intein. This natively split intein revealed high splicing efficiencies of up to 80–90% and superior rate constants resulting in a half-life time of the splicing reaction of up to only 60 s [20]. However, most known split inteins are much poorer catalysts. In fact, the most serious potential limitations of split inteins stem from the requirement of the intein fragments to associate and to fold into the correct 3D structure. If folding of the active site is prohibited, for example because of constraints imposed by the fused sequence (in particular the directly flanking amino acids), the splicing efficiency might be reduced or completely abolished. When the ligation of two protein fragments is attempted by protein *trans*-splicing, the protein fragments may cause aggregation or insolubility of the split intein fusion constructs. Furthermore, as many *trans*-splicing split inteins were artificially generated from regular '*cis*' inteins [16,21,22], in which case the two intein fragments are on one continuous polypeptide, the intein fragments themselves can contribute to insolubility. This latter problem appears to be greatly reduced for the few known naturally occurring split inteins, like the DnaE inteins from *Synechocystis* sp. PCC6803 [23] and *Nostoc punctiforme* [20,24]. However, the split inteins most interesting for endeavours in protein semisynthesis due to their small fragment size of 11 aa (Int^N of *Ssp* DnaB intein, [10,25,26]), 6 aa (Int^C of *Ssp* GyrB intein, [18]) and 15 aa (Int^C of *Npu* DnaE intein, [19]) are all artificially split.

A well-studied example for split inteins is the *Ssp* DnaB intein from *Synechocystis* sp. PCC6803 [10,17,25,27,28]. This intein was artificially split at several sites. The version split at the position of the deleted endonuclease domain (aa 104) was exploited for the selective chemical labelling of proteins on a unique cysteine that is appended in prelabelled form to the protein of interest by protein *trans*-splicing [17] and for the generation of cyclic peptides [29]. Of particular interest for protein semisynthesis is the version split after residue 11, because the resulting short *N*-terminal fragment of only 11 aa can be easily prepared by SPPS. As shown in Figure 1(A), the *N*-terminal extein sequence added to this Int^{N11} peptide can then be transferred by protein *trans*-splicing to a protein of interest to incorporate, for example, biophysical probes or other unnatural building blocks [10,26,28].

Our goal is to further improve the splicing efficiency of split inteins for protein semisynthesis using directed protein evolution. We report here the first step towards the selection of split intein mutants by phage display. To this end, we have presented two different Int^C fragments of the *Ssp* DnaB intein on the surface of a filamentous bacteriophage and demonstrated protein *trans*-splicing with the complementary Int^N part as a recombinant protein or synthetic peptide (Figure 1(B)). Phage display as an *in vitro* selection system is a promising approach for the directed evolution of inteins because of the wide range of possible selection conditions that can be applied.

Materials and Methods

General Techniques and Materials

Unless otherwise specified, standard protocols were used. As selection markers the antibiotics ampicillin (100 μg/ml), kanamycin (50 μg/ml) and chloramphenicol (34 μg/ml) were applied. Synthetic oligonucleotides were obtained from Eurofins MWG Operon (Ebersberg, Germany). All plasmids were verified by DNA sequencing. Reagents were purchased from Acros Organics (Nidderau, Germany), Applichem (Darmstadt, Germany), GE Healthcare (Munich, Germany), Novabiochem (Bad Soden, Germany), Roth (Karlsruhe, Germany) or Sigma-Aldrich (Munich, Germany). Restriction enzymes and markers were obtained from Fermentas (St Leon-Rot, Germany). Antibodies were from GE Healthcare (Munich, Germany), MoBiTec (Göttingen, Germany), New England Biolabs (Frankfurt, Germany) and Sigma-Aldrich (Munich, Germany).

Construction of DNA Vectors

The plasmids pTK56 and pCL20 for the expression of proteins **1** and **9**, respectively, were described previously [10,30]. The expression plasmid for protein **1a** was generated by site-directed mutagenesis using the Quikchange protocol from Stratagene and pTK56 as a template. All phagemid vectors were based on a derivate of pHEN1 [31], in which the coding region for the gpIII coat protein was truncated (aa 250–406, ΔgpIII) and the β-lactamase gene was replaced by a gene encoding the chloramphenicol acetyltransferase. Phagemid pDG92 served for the expression of construct **2** with the primary sequence 'PelB-MAQ-HHHHHH-gpD-GSSS-Int^C (aa 107–154)-SIRSCGKLAAALEIKRRAAA-c myc-GTR-ΔgpIII'. Phagemid pDG95 encoded construct **15** with the primary sequence 'PelB-MAQ-HHHHHH-gpD-GSSS-Int^C (aa 12–154; C50S)-SIELEIKRAAA-c myc-GTR-ΔgpIII'. Note that the amino acid sequences of the *Ssp* DnaB intein are numbered according to the published literature [27,32].

Journal of PeptideScience

Expression and Purification of Proteins

Protein expression was accomplished in *Escherichia coli* BL21-Gold (DE3) (Agilent Stratagene; La Jolla, California, USA) transformed with the respective plasmids. Cells were grown in Luria-Bertani (LB) medium supplemented with the appropriate antibiotic at 37 °C to an OD_{600} of 0.6–0.8. Before the protein expression was induced by addition of IPTG (final concentration 0.4 mM), the temperature was lowered to 30 °C. After 3–5 h the cells were harvested by centrifugation, resuspended in buffer A (50 mM Tris/HCl, 300 mM NaCl, pH 8.0), and lysed by two passages through a high pressure homogeniser (Avestin EmulsiFlex C5; Mannheim, Germany). Insoluble material was removed by centrifugation at 30 000 *g*. Affinity chromatography was performed using Ni^{2+}-NTA superflow material (Qiagen; Hilden, Germany) and proteins were eluted with buffer B (50 mM Tris/HCl, 100 mM NaCl, 250 mM imidazole, pH 8.0). Elution fractions containing the desired protein were pooled and dialysed against assay buffer (50 mM Tris/HCl, 300 mM NaCl, 2 mM DTT, 2 mM EDTA, pH 7.0) with 10% (v/v) glycerine, shock-frozen in liquid nitrogen and stored at −80 °C. Protein concentrations were determined using the calculated molecular extinction coefficients at 280 nm.

Peptide Synthesis

SPPS was performed partially manually and partially using a Syro XP peptide synthesizer (MultiSynTech; Bochum, Germany) on a Wang-resin according to standard protocols based on the Fmoc protecting group strategy and with HBTU as the activation reagent. The Cys building block was coupled manually as pentafluorophenyl ester without the addition of base. Biotin (Sigma; Munich, Germany) or Desthiobiotin (IBA; Göttingen, Germany) was coupled to the *N*-terminal amino group of the peptide applying a coupling time of at least 8 h. Deprotection/cleavage with reagent K (TFA : H_2O : thioanisole : ethanedithiol : phenol = 85 : 5 : 5 : 2.5 : 2.5) was carried out for 3 h. Crude peptide resulting from a diethylether precipitation of the cleavage mixture was purified by preparative HPLC (Varian; Palo Alto, California, USA) under acidic conditions (buffer A: H_2O with 0.05% TFA; buffer B: acetonitrile with 0.05% TFA) with a Microsorb 300-5 C18 column (Varian) and confirmed by MALDI-TOF MS analysis with an Autoflex II (Bruker Daltonics, Billerica, Massachusetts, USA). The peptides used were of >95% purity as monitored by analytical HPLC (Agilent 1100 series; Böblingen, Germany) using an Eclipse XDB-C18 3.5 µM column (Agilent). Peptide concentrations were determined photometrically at 215 nm and by comparison with another peptide of known concentration bearing the same amino acid sequence and an *N*-terminal carboxyfluorescein modification [28].

Phage Particle Preparation

Amplified phage particles were precipitated from an overnight culture of infected *E. coli* TG1 cells (Agilent Stratagene) [33]. In brief, one colony of *E. coli* cells transformed with the phagemid vector was picked and grown at 37 °C and 250 rpm in 10 ml 2×YT medium (17 g/l tryptone, 10 g/l yeast extract, 5 g/l NaCl, pH 7.0) with chloramphenicol and 2% (w/v) glucose to an OD_{600} of 0.4–0.6. At this stage the bacteria were superinfected by addition of the M13K07 helper phage (GE Healthcare; Munich, Germany) at an MOI of 50. After 30 min incubation at 37 °C in a water bath, the cells were pelleted at 1000 *g* for 10 min. The cell pellet was resuspended in 30 ml fresh 2×YT media with chloramphenicol and kanamycin and then grown overnight at 30 °C and 270 rpm. Cells were removed by two centrifugation steps (3000 *g*, 4 °C, 10 min). The phage particles remaining in solution were precipitated with 5× PEG 8000 solution (20% PEG 8000, 2.5 M NaCl) and the mixture was incubated on ice for at least 1 h. Following centrifugation at 10 000 *g* at 4 °C for 15 min, and removal of the supernatant, the phage/PEG pellet was resuspended in assay buffer and cleared once more from insoluble material by centrifugation (10 000 *g*, 4 °C, 5 min). Phage concentration was determined either photometrically at 270 nm [34] or via titrating infected bacteria cells [35].

Splicing Assay and Western Blot Analysis

For the splicing assays, the phage particles (estimated concentration of hybrid phages 50 nM) were mixed with the corresponding protein (at 0.5 µM) or synthetic peptide (at 5 µM) in assay buffer and incubated at 25 °C. Before the reaction and at various time points during the incubation, samples were removed, quenched by the addition of 4× SDS loading buffer (containing 8% (w/v) SDS and 20% (v/v) β-mercaptoethanol) and boiled before application onto an SDS PAGE gel. For the samples corresponding to the time point at 0 h the respective aliquot was removed immediately after mixing. About 1×10^{11} hybrid phage particles were loaded on one lane of the gel. SDS gels were transferred with a Mini Trans-Blot Cell (Bio-Rad; Hercules, California, USA) onto a PVDF membrane and blocked with 2% (w/v) milk powder or 3% (w/v) BSA dissolved in washing buffer (50 mM Tris/HCl, 137 mM NaCl, 0.1% (v/v) Tween 20, pH 7.6). Treatment with antibodies or a streptavidin/HRP conjugate was performed according to standard protocols and blots were developed using the ECL Western Blot Detection kit (GE Healthcare; Munich, Germany).

Results and Discussion

Strategy for Phage Display of Intact Int^C Fragments

Protein *trans*-splicing generates a covalent linkage between the two extein sequences and, therefore, is in principle well suited for selection procedures based on a binding event, such as phage display developed by Smith in 1985 [36]. We chose the monovalent phage display for larger proteins [37], a modification of Smith's original method. Here, the DNA encoding for the protein of interest, i.e. the Int^C fragment, is fused *N*-terminally to the gene for the truncated coat protein III (ΔgpIII, aa 250–406) into a phagemid plasmid. Transformation of a Gram-negative, F-pilus bearing host bacterium with the phagemid plasmid followed by superinfection with a helper phage programmes the cell to produce the hybrid particles of the filamentous bacteriophage M13. Addition of a biotinylated Int^N-fusion protein would link the biotin moiety to the phage coat and provide for the possible enrichment of phages active in protein *trans*-splicing using streptavidin beads.

GpIII and its fusion proteins are secreted by the Sec pathway in an unfolded state through the inner membrane and have to fold in the periplasm. It was unclear at the beginning of this work if an artificially split intein fragment, which is probably mostly unfolded, could efficiently be processed along this pathway and would remain stable towards potential periplasmic proteolytic digest. Initial attempts to present an Int^C fragment were performed with the *Ssp* DnaB intein split at the endonuclease position. The fusion construct StrepII-Int^{C107}-c myc-ΔgpIII (Int^{C107} = aa 107–154 of the *Ssp* DnaB mini-intein) was encoded in the phagemid vector preceded by a PelB signal sequence. However, it could not be

detected by Western blot analysis of the phage preparations (data not shown). This finding indicated problems during secretion, folding or due to proteolysis. We therefore decided to add a small and stably folding protein domain to the *N*-terminus of the Int^{C107} fragment. The head protein of the λ-phage (gpD) fulfils these criteria. It contains no cysteines, was shown by Plückthun and co-workers to exhibit superior secretion properties for phage displayed proteins [38], and was reported to increase the solubility of fusion proteins [20,39]. Indeed, Western blot analysis of phages prepared from *E. coli* cells producing PelB-His_6-gpD-Int^{C107}-c myc-ΔgpIII (construct **2**) from the phagemid vector showed the incorporation of the intact fusion protein (see Figure 2(B), left lane, and compare Table 1 for calculated molecular weights of all proteins in this study). In contrast, no hybrid phages could be detected in reasonable amount in case of the fusion construct PelB-StrepII-gpD-Int^{C107}-c myc-ΔgpIII (data not shown). The reasons for this dramatic difference as a function of an alternative short tag sequence are not known. Efficient production of hybrid phages was achieved in the same manner for the *Ssp* DnaB intein split at position 11 (see below, last chapter of the results section).

As the *Ssp* DnaB intein contains a catalytic serine instead of a cysteine at the *C*-terminal splice junction (Ser+1), no harm by the oxidative environment in the periplasm was expected. The single internal cysteine residue was mutated to serine to avoid potential scrambling of disulfide bridges. The mutant intein (C50S) was shown to be active in protein *trans*-splicing using purified model proteins ([28] and data not shown). This mutant was used for the phages prepared with the Int^{C12} (*Ssp* DnaB mini-intein split at position 11, see below, last chapter of the results section).

Table 1. Overview of primary sequence composition of protein and peptide constructs as well as splice products

Construct	Sequence	Calculated MW (kDa)
1	MBP-Int^{N104}-His_6	56.4
1a	MBP-Int^{N104} (C1A)-His_6	56.3
2	His_6-gpD-Int^{C107}-c myc-ΔgpIII	38.3*
3	MBP-c myc-ΔgpIII	63.6
6	c myc-ΔgpIII	20.3
7	Biotin-KKESG-Int^{N11}	1.9
8	Desthiobiotin-KKESG-Int^{N11}	1.8
9	Int^{C12}-Trx-His_6	30.2
10	Biotin-KKESG-Trx-His_6	14.5
11	Desthiobiotin-KKESG-Trx-His_6	14.4
13	Int^{C12}	16.5
14	Trx-His_6	13.7
15	His_6-gpD-Int^{C12} (C50S)-c myc-ΔgpIII	48.5*
16	Biotin-KKESG-c myc-ΔgpIII	20.4
17	Desthiobiotin-KKESG-c myc-ΔgpIII	20.3
19	c myc-ΔgpIII	19.6

* Note that the gpIII and its fusions migrate at an apparently higher molecular weight in SDS gels [40].

Protein *Trans*-Splicing on a Phage Displaying Int^C Split at the Endonuclease Position

To test for functionality of the Int^{C107} fragment, displayed as fusion protein His_6-gpD-Int^{C107}-c myc-ΔgpIII (construct **2**) on the phage coat, purified protein MBP-Int^{N104}-His_6 (construct **1**) containing the complementary intein fragment was incubated with hybrid phage particles. Figure 2(A) shows the reaction scheme of the expected products resulting from either protein *trans*-splicing or *C*-terminal cleavage. The latter reaction is a known side reaction of this intein [10,17,30,32]. The Western blot analysis shown in Figure 2(B) revealed after 2 h the formation of a new band recognised by the anti-c myc antibody that corresponded to the molecular weight of the expected splice product **3** (MBP-c myc-ΔgpIII; see Table 1). The reaction also yielded the *C*-terminal cleavage product **6** (c myc-ΔgpIII) as identified by its molecular weight. A control reaction with a splice-inactive mutant confirmed these designations. Incubation of protein **1a** (MBP-Int^{N104} (C1A)-

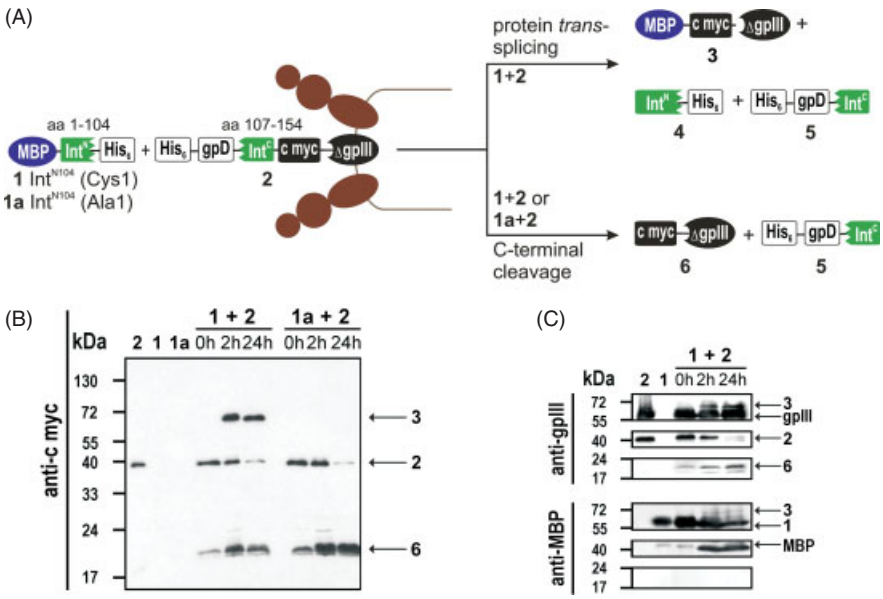

Figure 2. Splicing on a phage using the *Ssp* DnaB intein split at the endonuclease position. (A) Scheme of the reactions. (B) Western blot analysis of the reaction at different time points with an anti-c myc antibody. (C) Western blot analysis using anti-gpIII and anti-MBP antibodies. The hybrid phage was at a concentration of about 50 nM, while proteins **1** and **1a** were added at 0.5 μM.

Journal of PeptideScience

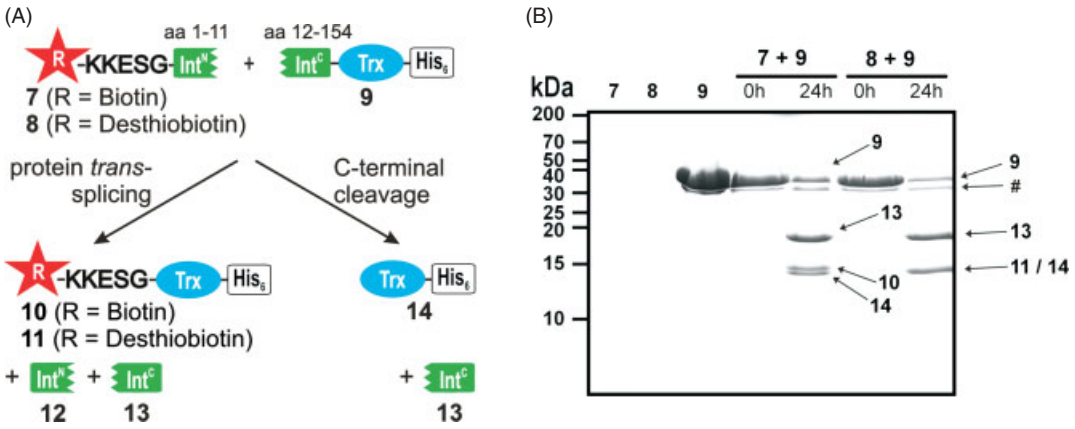

Figure 3. Semisynthetic protein *trans*-splicing with biotinylated peptides. (A) Scheme of the reactions. (B) Coomassie-stained SDS-PAGE gel. All reactants were incubated at a concentration of 20 μM. (# denotes a contamination).

His_6) with phage particles presenting construct **2** still resulted in *C*-terminal cleavage to give protein **6** but no formation of splice product **3** was observed (Figure 2(B)). The splice and cleavage products from the reaction of **1** and **2** presented on phage particles could also be detected by anti-gpIII and anti-MBP antibodies (Figure 2(C)). Note that the anti-gpIII antibodies also stained the native gpIII protein still present in the phage preparations. A protein band at about 40 kDa recognised by the anti-MBP antibody indicated the formation of MBP, which could result either from cleavage at the *N*-terminal splice junction or from proteolytic degradation of protein **1**.

Protein *Trans*-Splicing on a Phage Using a Semisynthetic Intein

After having shown protein *trans*-splicing on the phage for the *Ssp* DnaB intein split at the endonuclease position (aa 104), we tested the split position 11. As the small Int^{N11} fragment (aa 1–11) consists of only 11 amino acids, Ex^N–Int^{N11} peptides can be easily synthesised by SPPS for the incorporation of synthetic building blocks into proteins by protein *trans*-splicing (Figure 1). For a later selection procedure in the phage biopanning protocol we envisaged the utility of the biotin–streptavidin interaction to enrich phages active in protein *trans*-splicing. For this purpose, we synthesised peptides **7** and **8** with biotin and desthiobiotin modifications, respectively (Figure 3(A)). To confirm that both supported protein *trans*-splicing, we first carried out control reactions with the purified protein Int^{C12}-Trx-His_6 **9** (20 μM) [10,28]. Interestingly, in case of biotin–peptide **7** a double band on the Coomassie-stained SDS-PAGE gel corresponded to splice and *C*-terminal cleavage products, while these two products could not be resolved under the same conditions in case of the desthiobiotin–peptide **8** (Figure 3(B)). MS-analysis of the excised bands from the reaction with peptide **7** revealed the upper one as the splice and the lower one as the *C*-terminal cleavage product, while in case of the reaction with peptide **8** the presence of both products in the single band was confirmed (trypsin fragment of the splice product: ESGSIEGSGGGSDK; calc. 1266.544; obs. 1266.041 $[M+H]^+$; trypsin fragment of the *C*-terminal cleavage product: SIEGSGGGSDK; calc. 993.448; obs. 993.046 $[M+H]^+$).

Next, we turned to protein *trans*-splicing experiments with hybrid phages displaying the complementary intein fragment in the form of construct **15** (His_6-gpD-Int^{C12} (C50S)-c myc-ΔgpIII). As shown in Figure 4, splice products were indeed formed for both peptides **7** and **8**. Similar to the previous finding, we observed the same characteristic shift for the biotinylated product **16**, but not for the desthiobiotinylated product **17**, when compared to the *C*-terminal cleavage product. Both splice products were unequivocally identified in the blot using horseradish peroxidase-labelled streptavidin as a probe. A densitometric analysis indicated that the Int^{C12}-protein **15** on the phage was consumed to more than 50%. Figure 4(C) also indicates that partial proteolysis of the hybrid phages could be observed depending on the particular preparation (similarly for construct **2**; data not shown). This resulted in a band of similar molecular weight as the *C*-terminal cleavage product **19** (c myc-ΔgpIII) that was recognised by the anti-c myc antibody (marked by an asterisk in Figure 4(B) and (C)).

Conclusions

In summary, we have shown successful protein *trans*-splicing on an M13 filamentous phage. Two different Int^C fragments of artificially split *Ssp* DnaB inteins were fused in monovalent display to a truncated phage gpIII protein. For presentation on the phage, it was necessary to *N*-terminally tag the Int^C fragments with a His_6-gpD protein, probably to improve expression and secretion levels. In this format, the Int^C fragments were remarkably stable and of sufficient resistance towards proteolysis. The complementary Int^N fragments were added to the phage preparations either as recombinant protein or as a synthetic peptide. Interestingly, we could perform the splicing reactions with the low concentrations of the Int^C-phage particles of only about 50 nM. The proteins and peptides containing the Int^{N104} and Int^{N11} fragments were used at concentrations of 0.5 μM and 5 μM, respectively. Given that a K_d value of 1.1 ± 0.2 μM [28] was determined for the Int^{N11} and Int^{C12} fragments of the semisynthetic intein, the yield of the reaction is quite remarkable. However, one should keep in mind that once the active intein complex is formed the reaction is irreversible. Furthermore, the K_d value was measured with mutant constructs inactive for protein splicing that might have a lower affinity than the active split intein pieces used here. The K_d for the Int^{N104}/Int^{C107} pair split at the endonuclease position is not known, but probably smaller. For the natively split *Ssp* DnaE intein a K_d in the low nanomolar range was reported [41].

With these results in hand, the way is paved to evolve the catalytic activity of the presented intein by random mutagenesis of the Int^C fragments and enrichment of the improved variants through *in vitro* selection procedures, e.g. based on binding of

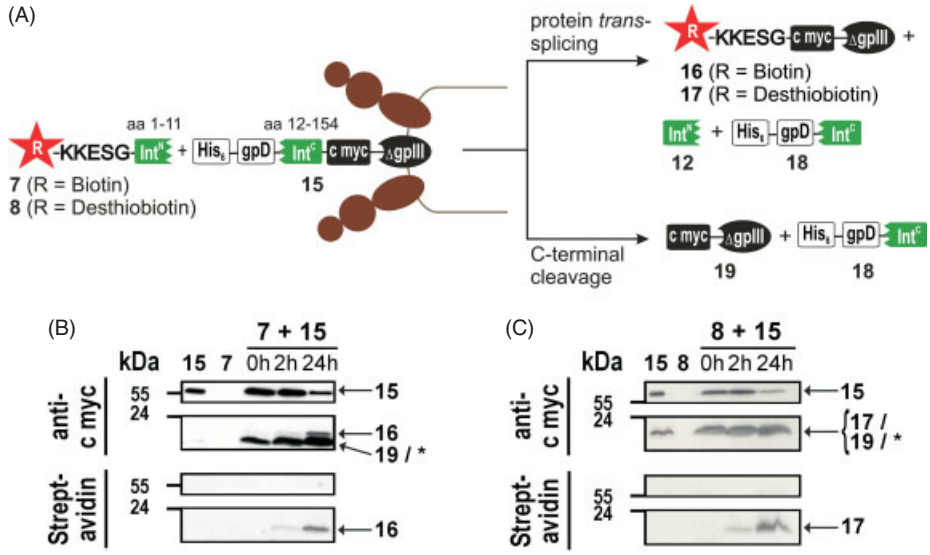

Figure 4. Semisynthetic protein *trans*-splicing on a phage. (A) Scheme of the reactions. (B) and (C) Western blot analysis of the reactions with biotin–peptide **7** (5 μM) and desthiobiotin–peptide **8** (5 μM), respectively, using anti-c myc antibody or streptavin/HRP conjugate. The hybrid phage was at a concentration of about 50 nM. (* denotes protein band corresponding to proteolytic degradation and/or *C*-terminal cleavage).

the splice product to a solid support like streptavidin-coated beads. Although in the present scheme only the Int^C fragment is encoded by the phagemid vector and therefore accessible to the coupling of genotype and phenotype, we expect that this does not impose a severe restriction in the case of the semisynthetic intein. Because this intein is split at position 11, 143 of the 154 aa (or 93%) of the intein are phagemid-encoded and should provide a critical sequence space for beneficial mutations. It is in particular this intein that is of high interest for the semisynthesis of proteins. Importantly, the stability of the phage particle and the *in vitro* nature of the protein *trans*-splicing reaction and phage display selection procedure will allow the application of a variety of useful conditions to improve the intein as a ligation tool. For example, one would expect to select for more robust intein mutants by enforcing improved folding properties in the absence of cytoplasmic chaperones. All selections for splicing should favour mutants with decreased cleavage activity. Intein mutants with higher splicing efficiency caused by higher affinity between the intein fragments and increased reaction rates are likely to have a selectable advantage when the splicing reaction on the phage is performed at low reactant concentrations and for short periods of time. Likewise, parameters could be controlled for a higher tolerance towards a range of temperature, pH and buffer conditions, as well as the presence of additives like denaturants and detergents. In contrast, most of these parameters are difficult or impossible to control or change in previously described *in vivo* selection systems of *cis*- and *trans*-splicing inteins [42–47] although *in vivo* selected mutants might also display improved properties under different *in vitro* conditions [42].

Apart from improving split inteins, we also note that our approach for the semisynthesis of proteins on a phage would provide a new way to enable the directed evolution of peptide-based biomolecules that consist of a chemical component and a genetically encoded part [48,49].

Acknowledgements

We are grateful to Christina Ludwig for her assistance during the peptide synthesis and Hans-Dieter Arndt for the access to the peptide synthesiser. We thank Christian Heinis for the kind gift of the pHEN1 phagemid vector. Funding for this work was provided by the DFG and the Fonds der Chemischen Industrie.

References

1 Dawson PE, Muir TW, Clark-Lewis I, Kent SB. Synthesis of proteins by native chemical ligation. *Science* 1994; **266**: 776–779.
2 Muir TW, Sondhi D, Cole PA. Expressed protein ligation: a general method for protein engineering. *Proc. Natl. Acad. Sci. USA* 1998; **95**: 6705–6710.
3 Evans TC, Benner J, Xu MQ Jr. Semisynthesis of cytotoxic proteins using a modified protein splicing element. *Protein Sci.* 1998; **7**: 2256–2264.
4 Saxon E, Armstrong JI, Bertozzi CR. A "traceless" Staudinger ligation for the chemoselective synthesis of amide bonds. *Org. Lett.* 2000; **2**: 2141–2143.
5 Jackson DY, Burnier J, Quan C, Stanley M, Tom J, Wells JA. A designed peptide ligase for total synthesis of ribonuclease A with unnatural catalytic residues. *Science* 1994; **266**: 243–247.
6 Bordusa F. Proteases in organic synthesis. *Chem. Rev.* 2002; **102**: 4817–4868.
7 Mao H, Hart SA, Schink A, Pollok BA. Sortase-mediated protein ligation: a new method for protein engineering. *J. Am. Chem. Soc.* 2004; **126**: 2670–2671.
8 Lew BM, Mills KV, Paulus H. Characteristics of protein splicing in trans mediated by a semisynthetic split intein. *Biopolymers* 1999; **51**: 355–362.
9 Giriat I, Muir TW. Protein semi-synthesis in living cells. *J. Am. Chem. Soc.* 2003; **125**: 7180–7181.
10 Ludwig C, Pfeiff M, Linne U, Mootz HD. Ligation of a synthetic peptide to the N terminus of a recombinant protein using semisynthetic protein trans-splicing. *Angew. Chem. Int. Ed. Engl.* 2006; **45**: 5218–5221.
11 Hermanson GT. *Bioconjugate Techniques*. Academic Press: San Diego, 2008.
12 Wang L, Brock A, Herberich B, Schultz PG. Expanding the genetic code of *Escherichia coli*. *Science* 2001; **292**: 498–500.
13 Kiick KL, Saxon E, Tirrell DA, Bertozzi CR. Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation. *Proc. Natl. Acad. Sci. USA* 2002; **99**: 19–24.
14 Budisa N. Prolegomena to future experimental efforts on genetic code engineering by expanding its amino acid repertoire. *Angew. Chem. Int. Ed. Engl.* 2004; **43**: 6426–6463.

15 Wang L, Schultz PG. Expanding the genetic code. *Angew. Chem. Int. Ed. Engl.* 2004; **44**: 34–66.
16 Mootz HD. Split inteins as versatile tools for protein semisynthesis. *ChemBioChem* 2009; **10**: 2579–2589.
17 Kurpiers T, Mootz HD. Regioselective cysteine bioconjugation by appending a labeled cystein tag to a protein by using protein splicing in trans. *Angew. Chem. Int. Ed. Engl.* 2007; **46**: 5234–5237.
18 Appleby JH, Zhou K, Volkmann G, Liu XQ. Novel split intein for trans-splicing synthetic peptide onto C terminus of protein. *J. Biol. Chem.* 2009; **284**: 6194–6199.
19 Aranko AS, Zuger S, Buchinger E, Iwai H. In vivo and in vitro protein ligation by naturally occurring and engineered split DnaE inteins. *PLoS One* 2009; **4**: e5185.
20 Zettler J, Schutz V, Mootz HD. The naturally split Npu DnaE intein exhibits an extraordinarily high rate in the protein trans-splicing reaction. *FEBS Lett.* 2009; **583**: 909–914.
21 Noren CJ, Wang J, Perler FB. Dissecting the chemistry of protein splicing and its applications. *Angew. Chem. Int. Ed. Engl.* 2000; **39**: 450–466.
22 Evans TJT, Xu MQ. Mechanistic and kinetic considerations of protein splicing. *Chem. Rev.* 2002; **102**: 4869–4884.
23 Wu H, Hu Z, Liu XQ. Protein trans-splicing by a split intein encoded in a split DnaE gene of Synechocystis sp. PCC6803. *Proc. Natl. Acad. Sci. USA* 1998; **95**: 9226–9231.
24 Iwai H, Zuger S, Jin J, Tam PH. Highly efficient protein trans-splicing by a naturally split DnaE intein from Nostoc punctiforme. *FEBS Lett.* 2006; **580**: 1853–1858.
25 Sun W, Yang J, Liu XQ. Synthetic two-piece and three-piece split inteins for protein trans-splicing. *J. Biol. Chem.* 2004; **279**: 35281–35286.
26 Ando T, Tsukiji S, Tanaka T, Nagamune T. Construction of a small-molecule-integrated semisynthetic split intein for in vivo protein ligation. *Chem. Commun. (Camb.)* 2007: 4995–4997.
27 Wu H, Xu MQ, Liu XQ. Protein trans-splicing and functional mini-inteins of a cyanobacterial dnaB intein. *Biochim. Biophys. Acta* 1998; **1387**: 422–432.
28 Ludwig C, Schwarzer D, Mootz HD. Interaction studies and alanine scanning analysis of a semi-synthetic split intein reveal thiazoline ring formation from an intermediate of the protein splicing reaction. *J. Biol. Chem.* 2008; **283**: 25264–25272.
29 Kinsella TM, Ohashi CT, Harder AG, Yam GC, Li W, Peelle B, Pali ES, Bennett MK, Molineaux SM, Anderson DA, Masuda ES, Payan DG. Retrovirally delivered random cyclic Peptide libraries yield inhibitors of interleukin-4 signaling in human B cells. *J. Biol. Chem.* 2002; **277**: 37512–37518.
30 Brenzel S, Kurpiers T, Mootz HD. Engineering artificially split inteins for applications in protein chemistry: biochemical characterization of the split Ssp DnaB intein and comparison to the split Sce VMA intein. *Biochemistry* 2006; **45**: 1571–1578.
31 Hoogenboom HR, Griffiths AD, Johnson KS, Chiswell DJ, Hudson P, Winter G. Multi-subunit proteins on the surface of filamentous phage: methodologies for displaying antibody (Fab) heavy and light chains. *Nucleic Acids Res.* 1991; **19**: 4133–4137.
32 Mathys S, Evans TC, Chute IC, Wu H, Chong S, Benner J, Liu XQ, Xu MQ. Characterization of a self-splicing mini-intein and its conversion into autocatalytic N- and C-terminal cleavage elements: facile production of protein building blocks for protein ligation. *Gene* 1999; **231**: 1–13.
33 Yamamoto KR, Alberts BM, Benzinger R, Lawhorne L, Treiber G. Rapid bacteriophage sedimentation in the presence of polyethylene glycol and its application to large-scale virus purification. *Virology* 1970; **40**: 734–744.
34 Lowman HB. In *Methods in Molecular Biology*, Cabilly S (ed.). Humana Press: Totowa, 1998; 249.
35 Koch J, Breitling F, Dubel S. Rapid titration of multiple samples of filamentous bacteriophage (M13) on nitrocellulose filters. *Biotechniques* 2000; **29**: 1196–1198, 2002.
36 Smith GP. Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science* 1985; **228**: 1315–1317.
37 Bass S, Greene R, Wells JA. Hormone phage: an enrichment method for variant proteins with altered binding properties. *Proteins* 1990; **8**: 309–314.
38 Steiner D, Forrer P, Stumpp MT, Pluckthun A. Signal sequences directing cotranslational translocation expand the range of proteins amenable to phage display. *Nat. Biotechnol.* 2006; **24**: 823–831.
39 Forrer P, Jaussi R. High-level expression of soluble heterologous proteins in the cytoplasm of *Escherichia coli* by fusion to the bacteriophage lambda head protein D. *Gene* 1998; **224**: 45–52.
40 van Wezenbeek P, Schoenmakers JG. Nucleotide sequence of the genes III, VI and I of bacteriophage M13. *Nucleic Acids Res.* 1979; **6**: 2799–2818.
41 Shi J, Muir TW. Development of a tandem protein trans-splicing system based on native and engineered split inteins. *J. Am. Chem. Soc.* 2005; **127**: 6198–6206.
42 Wood DW, Wu W, Belfort G, Derbyshire V, Belfort M. A genetic system yields self-cleaving inteins for bioseparations. *Nat. Biotechnol.* 1999; **17**: 889–892.
43 Adam E, Perler FB. Development of a positive genetic selection system for inhibition of protein splicing using mycobacterial inteins in *Escherichia coli* DNA gyrase subunit A. *J. Mol. Microbiol. Biotechnol.* 2002; **4**: 479–487.
44 Lew BM, Paulus H. An in vivo screening system against protein splicing useful for the isolation of non-splicing mutants or inhibitors of the RecA intein of Mycobacterium tuberculosis. *Gene* 2002; **282**: 169–177.
45 Buskirk AR, Ong YC, Gartner ZJ, Liu DR. Directed evolution of ligand dependence: small-molecule-activated protein splicing. *Proc. Natl. Acad. Sci. USA* 2004; **101**: 10505–10510.
46 Lockless SW, Muir TW. Traceless protein splicing utilizing evolved split inteins. *Proc. Natl. Acad. Sci. USA* 2009; **106**: 10999–11004.
47 Zeidler MP, Tan C, Bellaiche Y, Cherry S, Hader S, Gayko U, Perrimon N. Temperature-sensitive control of protein activity by conditionally splicing inteins. *Nat. Biotechnol.* 2004; **22**: 871–876.
48 Dwyer MA, Lu W, Dwyer JJ, Kossiakoff AA. Biosynthetic phage display: a novel protein engineering tool combining chemical and genetic diversity. *Chem. Biol.* 2000; **7**: 263–274.
49 Tian F, Tsao ML, Schultz PG. A phage display system with unnatural amino acids. *J. Am. Chem. Soc.* 2004; **126**: 15962–15963.

