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†

Steffen Brenzel, Thomas Kurpiers, and Henning D. Mootz*

Fachbereich Chemie/Biochemie, Philipps-Universität Marburg, Hans-Meerwein-Strasse, D-35032 Marburg, Germany

Received August 25, 2005; Revised Manuscript Received December 9, 2005

ABSTRACT: In protein trans-splicing, an intein domain split into two polypeptide chains mediates linkage of the flanking amino acid sequences, the N- and C-terminal exteins, with a native peptide bond. This process can be exploited to assemble proteins from two separately prepared fragments, e.g., for the segmental labeling with isotopes for NMR studies or the incorporation of chemical and biophysical probes. Split inteins can be artificially generated by genetic means; however, the purified intein^N and intein^C fragments usually require a denaturation and renaturation treatment to fold into the active intein, thus preventing their application to proteins that cannot be refolded. Here, we report that the purified fragments of the artificially split DnaB helicase of *Synechocystis* spp. PCC6803 (*Ssp* DnaB) intein are active under native conditions. The first-order rate constant of the protein trans-splicing reaction was 7.1×10^{-4} s^{-1}. The previously described split vacuolar ATPase of *Saccharomyces cerevisiae (Sce* VMA) intein is the only other artificially split intein that is active under native conditions; however, it requires induced complex formation of the intein fragments by auxiliary dimerization domains for efficient protein trans-splicing. In contrast, fusion of the dimerization domains to the split *Ssp* DnaB intein fragments had no effect on activity. This difference was also reflected by a higher thermostability of the split *Ssp* DnaB intein. Further investigations of the split *Sce* VMA intein under optimized conditions revealed a first-order rate constant of 9.4×10^{-4} s^{-1} for protein trans-splicing and 1.7×10^{-3} s^{-1} for C-terminal cleavage involving a Cys1Ala mutant. Finally, we show that the two split inteins are orthogonal, suggesting further applications for the assembly of proteins from more than two parts.

Protein splicing is the process of an autocatalytic excision of an internal protein sequence, the intein, out of a larger precursor protein with concomitant joining of the flanking sequences, the N- and C-terminal exteins, by a native peptide bond (see Figure 1A) (*1*−*4*). The principle steps in the reaction mechanism and the involvement of the key catalytic residues are well-understood and have been reviewed comprehensively (*1*, *2*). Meanwhile, more than 200 inteins have been identified in all three kingdoms of life (*5*). While the physiological role of protein splicing is still a matter of debate, inteins have been exploited for a variety of protein engineering approaches (*6*−*11*). Importantly, they are in general very promiscuous toward their flanking extein sequences and can be inserted into a heterologous protein context by genetic manipulation.

Split inteins represent a very valuable tool for the manipulation of the protein structure. As illustrated in Figure 1B, the intein domain reassembles from two separate polypeptides, the intein^N and intein^C fragments, and joins the N- and C-terminal exteins in a protein trans-splicing reaction. This allows for the separate production of the two polypeptides and thereby offers many ways to manipulate just one part of the final splice product. Applications include, for example, the segmental isotope labeling of proteins for NMR studies (*12*−*14*) and the incorporation of synthetic peptides that may contain unnatural amino acids into proteins (*15*−*18*) by using a chemically synthesized variant of the smaller intein^C fragment. The *Ssp* DnaE intein is the only characterized native split intein, and the intein fragments DnaE^N and DnaE^C bind to each other with high affinity (*19*).

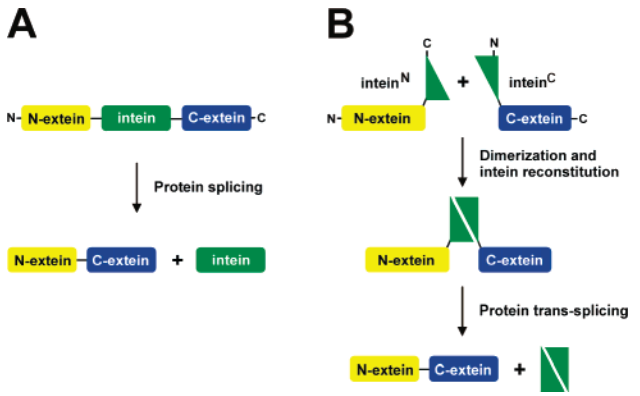

FIGURE 1: Schematic illustration of protein splicing (A) and protein trans-splicing (B).

Split inteins can be generated artificially from regular inteins; however, as generally found for split protein frag-

† Financial support was provided by an Emmy Noether Grant from the Deutsche Forschungsgemeinschaft and by the Fonds der Chemischen Industrie.

* To whom correspondence should be addressed. E-mail: mootz@chemie.uni-marburg.de. Telephone: +49-6421-2825404. Fax: +49-6421-2822191.

10.1021/bi051697+ CCC: $33.50
Published on Web 01/19/2006

ments, they usually behave very differently than their intact parent proteins. When split protein fragments are individually expressed, they tend to misfold or aggregate, which often leads to insolubility. In many cases, an active protein can be reassembled from two fragments only by coexpression in the same cell (which allows cofolding and probably is supported by endogenous chaperones) or by a renaturing procedure from a denaturing solution containing both purified fragments (*20*−*22*). This holds true in particular for those protein fragments that have been created from a compact folding domain, as is the case with inteins. In these cases, the individual N- and C-terminal fragments cannot form extensive well-packed globular structures but are likely to expose hydrophobic areas that will lead to misfolding and aggregation or even proteolytic degradation.

Several artificially split inteins have been studied *in vitro*, including the *Mtu* RecA (*23*), the *Psp* Pol-1 (*24*), PI-*Pfu*I (*12*−*14*), and PI-*Pfu*II (*13*) inteins, that required a denaturation-refolding treatment of the mixed intein halves to become active in protein trans-splicing. In a particularly well-studied example, the artificially split *Mtu* RecA intein was found to trans-splice with high efficiency when both halves were coexpressed in *Escherichia coli* (*25*), whereas the same intein fragments aggregated during purification when expressed separately and required a refolding step in the presence of the corresponding half to reconstitute the active intein (*23*).

The artificially split vacuolar ATPase of *Saccharomyces cerevisiae* (*Sce* VMA)[1] intein presents the only exception from this rule reported thus far. However, efficient protein trans-splicing between the purified VMA^N and VMA^C fragments was effected by fused dimerization domains, which probably induced the folding of the active intein by the induced proximity (*26*−*29*). The auxiliary domains FKBP12 and FKBP−rapamycin binding domain (FRB) form a dimer only in the presence of the small molecule rapamycin (*30*), whereas the FKBP12(F36M) mutant (F_M domain) spontaneously forms a homodimer that can be dissociated by the addition of rapamycin (*31*). Fusion of the VMA^N and VMA^C fragments with these domains could thus be exploited for the development of ligand-activated (*26*) and ligand-inhibited (*29*) trans-splicing inteins, respectively.

In this paper, we show that the purified halves of the artificially split DnaB helicase of *Synechocystis* spp. PCC6803 (*Ssp* DnaB) intein are active in protein trans-splicing under native conditions. In contrast to other artificially split inteins, the $DnaB^N$ and $DnaB^C$ halves assemble spontaneously to the active intein and do not require dimerization domains. Both the DnaB and the conversely behaving VMA intein are further biochemically characterized.

EXPERIMENTAL PROCEDURES

General Procedures. Standard chemicals and rapamycin were purchased from Sigma−Aldrich (Munich, Germany) and Carl Roth (Karlsruhe, Germany). Restriction enzymes and amylose resin were obtained from New England Biolabs (Frankfurt am Main, Germany). Ni^{2+}-NTA resin was ordered from Qiagen (Hilden, Germany), *Pfu* DNA polymerase was ordered from Stratagene (Amsterdam, The Netherlands); and synthetic oligonucleotides were purchased from Operon (Cologne, Germany). Standard protocols were applied for handling of recombinant DNA and proteins. Mass analysis was performed on an ESI time-of-flight spectrometer from Applied Biosystems type QStar Pulsar i.

DNA Cloning of Expression Plasmids. The construction of the expression plasmids pHM41, pHM45, pHM81, pSB18, and pSB19 was reported previously (*27*−*29*). To create plasmid pTK27, a polymerase chain reaction (PCR) was carried out using pHM45 as a template DNA and oligonucleotides 5′-ATAACTAGTGCTTTTGCACGTGAGTGCCGC-3′ and 5′-ATAAAGCTTAGTGATGGTGATGG-3′ (restriction sites are underlined). The PCR product was digested with *Spe*I and *Hin*dIII and cloned into pHM45. Plasmid pSB13 was generated by cloning a PCR product encoding $DnaB^N$ into the *Eco*RI and *Xba*I-treated vector pHM41. The PCR product was obtained using oligonucleotides 5′-ATAGAATTCTCCGGCTGCATCAGTGGAGATAG-3′ and 5′-ATAAAGCTTTTATCTAGATAAAGAGGAGCTTTCTAGTTTACG-3′ and pMST (described in ref *32*; a kind gift from Xiang-Qin Liu) as the DNA template. The DNA fragment encoding $DnaB^C$ was PCR-amplified from pMST using oligonucleotides 5′-ATACCATGGGCACTAGTTCACCAGAAATAGAAAAGTTGTC-3′ and 5′-ATAGGTACCAGATCTAATACTGTTATGGACAATGATGTC-3′ and ligated into the *Spe*I and *Bgl*II sites of plasmid pHM24 (*26*). From this intermediate plasmid, the *Xba*I−*Hin*dIII fragment was excised and cloned into pHM45 to give pSB15. Two native, directly flanking extein residues SG and SI were included with the $DnaB^N$ and $DnaB^C$ fragments, respectively. All plasmids were confirmed by DNA sequencing. For the construction of pTK56, the human FKBP12 (FK506-binding protein) (FKBP) encoding fragment of pSB13 was excised using the restriction enzymes *Xba*I and *Spe*I and religated. The plasmid pTK55 was obtained through cloning of a *Spe*I−*Hin*dIII fragment from a pQE60 vector (Qiagen), coding for $DnaB^C$−hexahistidin tag (His_6), into the *Xba*I−*Hin*dIII-treated vector pHM45. The $DnaB^C$ fragment was previously PCR-amplified as described above and ligated into the pQE60 vector via its *Nco*I and *Bgl*II sites. Therefore, the C-terminal extein sequences of the DnaB constructs **2** and **7**, encoded by pSB15 and pTK55, differ slightly (SIRS−His_6 versus SIRSRS−His_6).

Expression and Purification of Proteins. For the expression of constructs **11**, **12**, **16**, **14**, **13**, **1**, **2**, **15**, **7**, and **6** (see Figure 2B), *E. coli* BL21 cells were transformed with the plasmids pHM41, pHM45, pHM81, pSB18, pSB19, pSB13, pSB15, pTK27, pTK55, and pTK56, respectively. Expression was carried out in 500 mL of LB medium containing ampicillin (100 μg/mL) and glucose (0.2%), which was inoculated from a 5 mL overnight culture in the same medium. Cells were grown at 37 °C to an OD (600 nm) of about 0.7. Then, the temperature was shifted to 30 °C, and isopropyl-β-D-thiogalactopyranoside was added to a final concentration of 0.2 mM. Cells were pelleted by centrifugation after 3−4 h and resuspended in wash buffer (50 mM Tris/HCl at pH 8.0 and 300 mM NaCl) with 5 mM imidazole and frozen at −80 °C.

[1] Abbreviations: FKBP, human FKBP12 (FK506-binding protein); F_M, F36M mutant of FKBP12; FRB, FKBP−rapamycin binding domain; MBP, *E. coli* maltose-binding protein; His_6, hexahistidin tag; *Sce* VMA, intein of the vacuolar ATPase of *Saccharomyces cerevisiae*; *Ssp* DnaB, intein of the DnaB helicase of *Synechocystis* spp. PCC6803; VMA^N and VMA^C, N- and C-terminal intein halves of the split *Sce* VMA intein; $DnaB^N$ and $DnaB^C$, N- and C-terminal intein halves of the split *Ssp* DnaB intein.

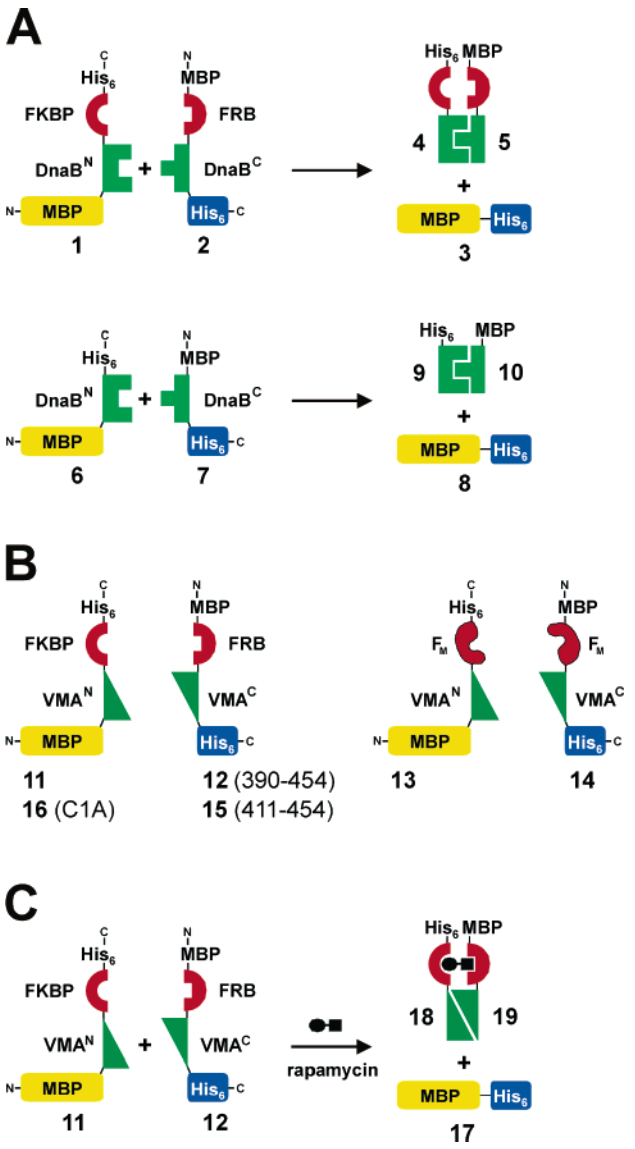

FIGURE 2: Overview of all protein constructs used in this study. (A) Scheme of the protein trans-splicing reactions of the constructs involving the split DnaB intein. (B) Nomenclature of the constructs containing the split VMA intein. The numbering refers to the entire native VMA intein (454 amino acids). The VMA^N fragment consists of amino acids 1−184, whereas two different VMA^C fragments were investigated as shown. (C) Reaction scheme of the protein trans-splicing reaction of proteins **11** and **12** with rapamycin.

For protein purification, cells were thawed on ice and disrupted in a French pressure cell. Insoluble cell material was pelleted by centrifugation at 30000*g* for 15 min. The soluble cell extract was loaded on a Ni^{2+}-NTA agarose column previously equilibrated with 12 column volumes of wash buffer (plus 5 mM imidazole). The subsequent washing steps were performed with 15 column volumes of wash buffer with 5 mM imidazole, 5 column volumes of wash buffer with 20 mM imidazole, and 2 column volumes of wash buffer with 40 mM imidazole. Proteins were eluted with elution buffer (50 mM Tris/HCl at pH 8.0, 300 mM NaCl, and 250 mM imidazole). The fractions containing the protein in an adequate concentration were pooled and dialyzed against amylose column buffer (20 mM Tris/HCl at pH 7.4, 200 mM NaCl, and 1 mM EDTA). Subsequently, the protein solutions were loaded on an equilibrated amylose column (NEB). The column was washed with 15 column volumes of the amylose column buffer, and proteins were eluted with the same buffer containing maltose (10 mM). The fractions with pure protein were pooled and dialyzed against assay buffer (50 mM Tris/HCl at pH 7.0, 300 mM NaCl, 1 mM EDTA, 10% glycerine, and 2 mM DTT).

Protein concentrations were determined using the calculated absorbances at 280 nm. Protein solutions were stored in aliquots at −80 °C. Once an aliquot was thawed, it could be stored on ice for several days without losing splicing activity.

Protein-Splicing Assays. All protein-splicing reactions were performed in an assay buffer at 25 °C unless stated otherwise. Standard protein concentrations were 2 μM of each protein. DTT was freshly added to a final concentration of 2 mM just before the reaction mixtures were prepared. All experiments were carried out at least in duplicate. Kinetic measurements were performed by adding rapamycin from a 1 mM stock solution in DMSO to a final concentration of 10 μM directly to the reaction mixture of constructs **11** and **12**, **11** and **15**, as well as **16** and **12**. The uninduced reactions were performed with 1% (v/v) DMSO. Constructs **13** and **14** with F_M domains were separately incubated for 5 min in an assay buffer containing 12.5 μM rapamycin or the equivalent amount of DMSO [1.25% (v/v)]. Mixing the proteins represented the respective starting point of the protein-splicing reactions. Samples were taken at different time points and stopped with 4× SDS−PAGE loading buffer, boiled for 5 min, and stored at −20 °C.

For determination of the temperature dependence, each protein was preincubated at the respective temperature for 5 min before the reaction was initiated by mixing the complementary pairs of constructs. The reaction mixtures of constructs **11** and **12** also contained 10 μM rapamycin. The yield of the splice product was monitored after 2 h.

The different sets of cross-splicing experiments were conducted either by mixing the constructs containing the VMA^N and $DnaB^C$ fragments (**11** and **2**) or vice versa (**1** and **12**) and inducing the dimerization by the addition of rapamycin. Alternatively, the four constructs were mixed, and rapamycin was added. Reactions were stopped after 2 h.

SDS−PAGE and Densitometric Analysis. All protein-splicing reactions were analyzed by densitometric analysis of SDS−PAGE gels (10% polyacrylamide) stained with Coomassie Brilliant Blue. The stained gels were scanned and analyzed using "Scion image" (www.scioncorp.com). The relative percentages of spliced proteins for all constructs of the split VMA intein were calculated as previously described (*29*). For the split DnaB intein, the relative increase in intensity of the product band of *E. coli* maltose-binding protein (MBP)−His_6 (**3** or **8**) was quantified and related to the intensity of the protein bands in unreacted samples at 0 min. For the calculation of the reaction progress, a correction factor reflecting the relative differences in protein size was taken into account. First-order rate constants were determined using the equation $c = c_0 + Ae^{-kt}$, where c is the reactant concentration at the time point t, c_0 is the asymptote for large t values, A is the amplitude, and k represents the rate constant.

RESULTS AND DISCUSSION

Purified Fragments of the Artificially Split Ssp DnaB Intein Are Active under Native Conditions. To generate a new artificially split intein, we investigated the DnaB intein from *Synechocystis* spp. PCC6803. The native DnaB intein consists of 429 amino acids and includes a centrally located homing endonuclease domain. Liu and co-workers described the removal of 275 amino acids covering the endonuclease and showed that the resulting DnaB mini-intein of 154 amino acids was still functional (*32*). They also showed that the DnaB mini-intein can be split into two halves that are active for protein trans-splicing when coexpressed in *E. coli*. To this end, a short sequence corresponding to a termination codon, a Shine−Dalgarno sequence, and a new initiation codon for translation were inserted into a plasmid encoding the mini-intein. Expression from this two-gene operon and *in vivo* protein trans-splicing was monitored by Western blotting using antibodies directed against the model N- and C-terminal extein sequences (*32*). While this study demonstrated the activity of the intein halves when coexpressed in *E. coli*, we reasoned that protein trans-splicing of the purified constructs for *in vitro* applications still remained questionable.

To test *in vitro* protein trans-splicing of the split DnaB intein, we therefore carried out separate expression and purification of two different protein fragments. For this purpose, similar intein halves $DnaB^N$ (amino acids 1−104) and $DnaB^C$ (amino acids 382−429) as described by Liu and co-workers were chosen and incorporated into the fusion constructs **1** (MBP−$DnaB^N$−FKBP−His_6) and **2** (MBP−FRB−$DnaB^C$−His_6) as shown in Figure 2A. In these constructs, the underlined MBP and a hexahistidine tag (His_6) served as N- and C-terminal exteins, respectively. The FKBP and FRB domains at the other end of the intein halves were included to heterodimerize the two polypeptides with rapamycin, to test if induced proximity is required to form the active intein.

Both proteins were expressed in *E. coli* and purified from the soluble fraction. The purified constructs were mixed at stoichiometric ratios (2 μM in each) in the absence or presence of rapamycin (10 μM). Surprisingly, protein trans-splicing was observed in both cases, independent of the presence of rapamycin. As shown in Figure 3A, the expected products were identified in Coomassie-stained SDS−PAGE gels. The formation of the correct products was also confirmed by ESI−MS analysis of the proteins (see Figure 3D). Parts B and C of Figure 3 show the time courses of the same reactions in the absence or presence of rapamycin. Both reactions proceeded very similarly, which implied that the rapamycin-induced dimerization had no or only a negligible effect. After 60 min, the product formation reached a plateau at about 40% completion, which is typical for split inteins. It is generally assumed for trans-splicing inteins that the process of dimerization of the intein halves is fast compared to a slow subsequent protein-splicing step (*4*). The latter rate-determining step would then follow a pseudo-first-order reaction. Consistent with these considerations, the data of the reaction without rapamycin fit very well to first-order kinetics, giving rise to a rate constant of $7.1 \pm 1.1 \times 10^{-4}$ s^{-1} (see also Table 1 for all rate constants discussed in this paper).

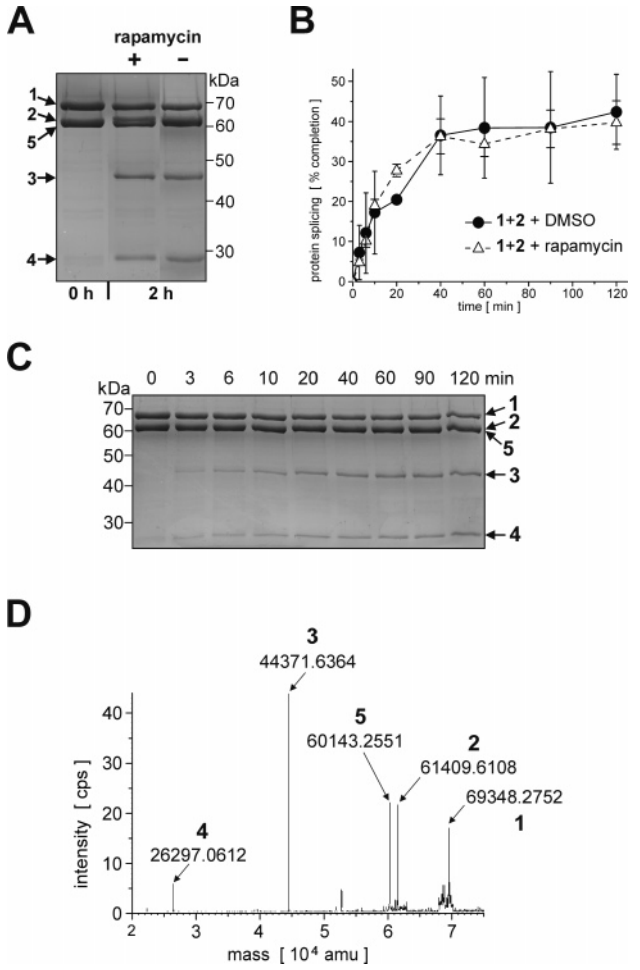

FIGURE 3: Analysis of the protein trans-splicing reactions involving the split DnaB intein with the FKBP and FRB domains. (A) Coomassie-stained SDS−PAGE gel showing the reaction starting materials and products in the presence or absence of rapamycin. Compare with Figure 2A for product designations. (B) Time courses with rapamycin and the DMSO control. (C) Coomassie-stained SDS−PAGE gel corresponding to a time-course reaction quantified in B. (D) Spectrum of an ESI−MS analysis of a reaction mixture after 2 h, verifying the reactants **1** and **2** as well as the products **3**, **4**, and **5**. The calculated masses are 69 351.6 Da (**1**), 61 405.3 Da (**2**), 60 138.9 Da (**5**), 44 368.9 Da (**3**), and 26 266.0 Da (**4**). Protein **4** was further verified by an independent measurement optimized for proteins of lower molecular weight that revealed a mass of 26 266.0 Da. All protein trans-splicing reactions were carried out at 25 °C, and equimolar amounts of each construct were used.

Table 1: First-Order Rate Constants for the Investigated Protein trans-Splicing Reactions

reaction	rate constant k (s^{-1})	rapamycin
1 + **2**	$7.1 \pm 1.1 \times 10^{-4}$	−
6 + **7**	$9.9 \pm 0.8 \times 10^{-4}$	−
11 + **12**	$9.4 \pm 2.0 \times 10^{-4}$	+
11 + **15**	$1.2 \pm 0.1 \times 10^{-3}$	+
16 + **12**[a]	$1.7 \pm 0.1 \times 10^{-3}$	+
13 + **14**	$2.0 \pm 0.3 \times 10^{-3}$	−
Ssp DnaE[b]	$6.6 \pm 1.3 \times 10^{-5}$	

[a] C-Terminal cleavage reaction. [b] Taken from ref *35*.

To rule out any secondary effect of the fused dimerization domains, we prepared a second pair of proteins that lacked the FKBP and FRB segments. Constructs **6** (MBP−$DnaB^N$−His_6) and **7** (MBP−$DnaB^C$−His_6) (see Figure 2A) were also purified from the soluble fraction after expression in *E. coli*.

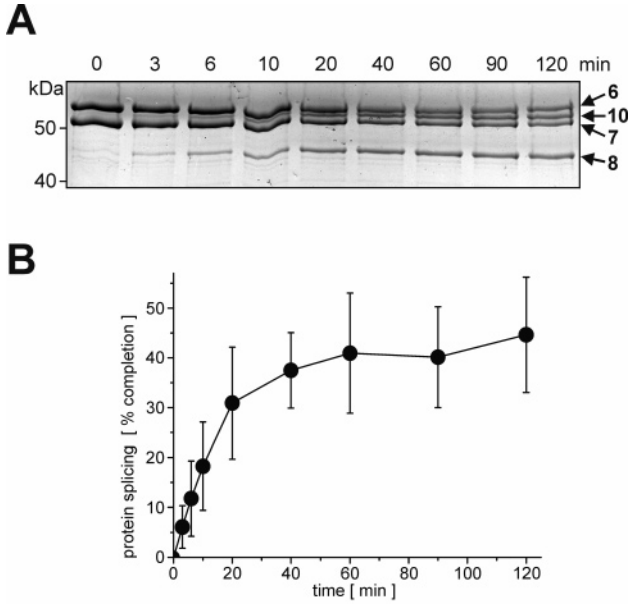

FIGURE 4: Analysis of the split DnaB intein constructs without the FKBP and FRB domains. (A) Time course of the reaction of **6** and **7** visualized on a Coomassie-stained SDS−PAGE gel. Product **9** is too small to be seen on the gel (see Figure 2A for numbering of proteins). (B) Kinetics of the protein-splicing reaction. The calculated masses are 56.4 kDa (**6**), 50.1 kDa (**7**), 48.6 kDa (**10**), 44.6 kDa (**8**), and 13.3 kDa (**9**).

The active intein formed spontaneously after mixing both constructs (each at 2 μM), and the expected products could be observed (see Figure 4A). ESI−MS analysis of a reaction mixture confirmed the identity of all reactants and products (data not shown; **6**, M (calculated) = 56 407.9 Da, M (observed) = 56 403 Da; **7**, M (calculated) = 50 122.3 Da, M (observed) = 50 127 Da; **8**, M (calculated) = 44 612.2 Da, M (observed) = 44 616 Da; **9**, M (calculated) = 13 322.3 Da, M (observed) = 13 324 Da; **10**, M (calculated) = 48 613.7 Da, M (observed) = 48 620 Da). A kinetic analysis shown in Figure 4B revealed no significant difference in comparison to the constructs carrying the FKBP and FRB domains. The protein trans-splicing reaction went to about 45% completion after 2 h and proceeded with a first-order rate constant of $9.9 \pm 0.8 \times 10^{-4}$ s^{-1}. We conclude that the DnaB^N and DnaB^C split intein fragments spontaneously form the active intein under native conditions. Obviously, they retained a higher affinity than the VMA^N and VMA^C pieces, possibly because of better folding properties or more suitable presentation of important recognition elements.

Split Ssp DnaB Intein Is Less Temperature-Sensitive Than the Split Sce VMA Intein. To further characterize the different properties of the split DnaB and VMA inteins, we investigated protein trans-splicing at different temperatures. For the VMA intein, two different and previously described sets of constructs were investigated (see parts B and C of Figure 2). On one hand, in constructs **11** and **12**, the intein halves VMA^N and VMA^C are fused to FKBP and FRB domains (*27*). They are thus presented in the same context as the DnaB^N and DnaB^C pieces in constructs **1** and **2**. On the other hand, in constructs **13** and **14**, the VMA^N and VMA^C fragments are each fused to one F_M domain, whose spontaneous dimerization can be inhibited by rapamycin (*29*).

Figure 5 shows the product yields of the three sets of constructs monitored after 2 h of protein trans-splicing at

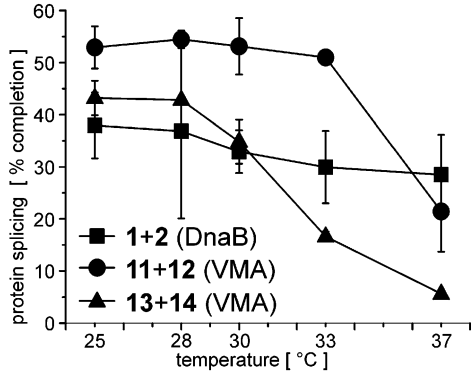

FIGURE 5: Comparison of the temperature dependence of the split DnaB and the split VMA inteins. The reaction progress was determined after a 2 h reaction time. The reaction mixture of the constructs **11** and **12** contained rapamycin, and the mixture of the constructs **13** and **14** contained the corresponding amount of DMSO. See Figure 2 for construct designations.

the indicated temperatures. The activity of the split DnaB intein decreased only slightly when the temperature was increased from 25 to 37 °C. In contrast, product yields for both split VMA inteins dropped dramatically at the higher temperature. The VMA constructs **13** and **14** with the F_M domains exhibited the highest sensitivity, with a pronounced loss of activity already above 30 °C. Apart from the intrinsic properties of the VMA intein, these observations might also be explained by the fused dimerization domains. Possibly, the spatial orientation of the termini in the FKBP−rapamycin−FRB complex is more favorable for folding of the fused VMA^N and VMA^C pieces compared to the complex of the F_M−F_M homodimer. Alternatively, the interaction between two F_M domains [K_D = 30 μM (*31*)] itself might be more temperature-sensitive than in the FKBP−rapamycin−FRB complex [K_D = 2 nM (*30*)]. In conclusion, these data suggest a more pronounced temperature-dependent unfolding of the split VMA intein compared to the more robust split DnaB intein.

Further Truncation of the VMA^C Intein Half and Kinetic Studies on the Split Sce VMA Intein. A closer examination of the VMA^C intein half in the above-described constructs **12** and **14** (amino acids 390−454, numbering according to the native VMA intein including the homing endonuclease domain) guided by the crystal structure of the VMA intein [PDB accession code 1VDE (*33*)] revealed that amino acids 390−410 are accounted for the inserted homing endonuclease domain or the linker between this domain and the C-terminal half of the intein. We speculated that these 21 amino acids could interfere with folding of the VMA^C piece or occlude the interaction interface in the initial contact with the VMA^N half and thus prohibit efficient recognition of the split intein halves. Construct **15** was generated containing a truncated VMA^C fragment (amino acids 411−454, see Figure 2B). Kinetic measurements of the protein trans-splicing with construct **11** showed that the splicing activities of the two different forms of VMA^C in the absence or presence of rapamycin were identical within the error margins of the assay (see Figure 6B). In both cases, about 70% completion of the reaction was reached after 2 h and up to 85% completion after 44 h (data not shown), when the small molecule inducer was present. From these data, first-order rate constants for protein trans-splicing of $9.4 \pm 2.0 \times 10^{-4}$

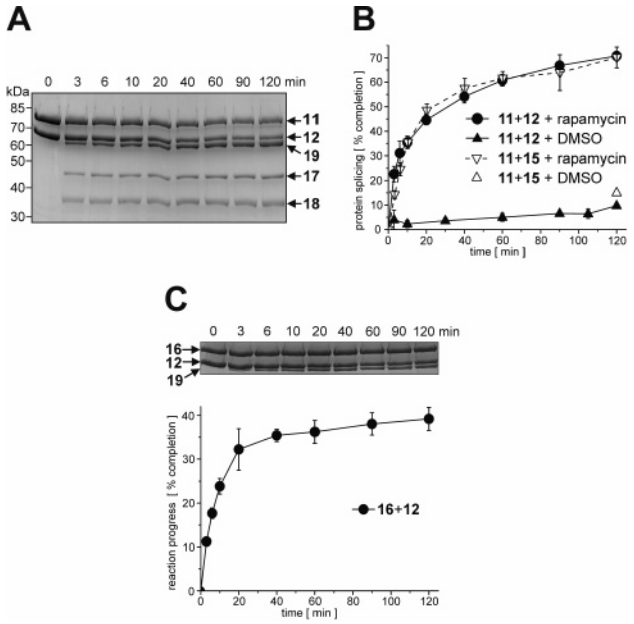

FIGURE 6: Kinetic characterization of truncated and mutant versions of the split VMA intein. (A) Coomassie-stained SDS−PAGE gel of a time-course experiment using proteins **11** and **12**. Calculated molecular masses are 78.7 kDa (**11**), 64.1 kDa (**12**), 62.1 kDa (**19**), 45.2 kDa (**17**), and 35.4 kDa (**18**). (B) Comparison of the longer and truncated form of the VMA^C fragment in time-course experiments in the presence and absence of rapamycin. (C) Time course of the C-terminal cleavage reaction involving the C1A mutant of VMA^N (protein **16**) and **12**. All reactions were performed at 25 °C. See parts B and C of Figure 2 for protein designations.

and $1.2 \pm 0.1 \times 10^{-3}$ s^{-1}, for the longer and truncated constructs, respectively, were determined. Also, the residual splicing activity in the absence of the ligand was comparable at 10−15% within 2 h. Thus, the removal of the 21 amino acids did not increase the spontaneous activity of the split VMA intein, providing further evidence for the inherently different behavior compared to the split DnaB intein. It is also notable that the VMA^C fragment could be reduced to 43 amino acids.

Interestingly, the rate constants of the rapamycin-dependent split VMA intein determined in this study were about 5-fold higher than previously reported ($k = 1.9 \times 10^{-4}$ s^{-1}, see ref *26*). These differences might be explained by an additional His_6 or MBP affinity tag included in constructs **11** and **12**, respectively, of the present study. These tags allowed a more gentle purification of the split intein constructs, which avoided an ion-exchange chromatography step as applied in the previous paper (*26*), which might account for the higher protein activity. It is conceivable that the probably partially unfolded intein fragments were further unfolded on the ion-exchange column. Importantly, the values reported in this paper are therefore likely to represent the full kinetic potential of the split VMA intein. They are also in good agreement with our previous kinetic investigation of the split VMA intein fused with the F_M-homodimerization domains (constructs **13** and **14**), which resulted in a first-order rate constant of $2.0 \pm 0.3 \times 10^{-3}$ s^{-1} (*29*) (see also Table 1).

Because we had identified optimized conditions for split VMA intein activity, we extended our kinetic investigation on the C-terminal cleavage reaction, which is particularly useful in the context of conditional inteins to liberate the C-terminal extein (*28*). The protein-splicing pathway of inteins can be shuttled to C-terminal cleavage when the catalytic residue C1 of the intein is blocked by mutation. Figure 6C shows the time course of C-terminal cleavage of the C1A mutant **16** with construct **12** in the presence of rapamycin. This curve fit to a first-order reaction with a rate constant of $1.7 \pm 0.1 \times 10^{-3}$ s^{-1} and is thus in the same range as protein trans-splicing. A previous study reported the kinetics of N-terminal thioester formation at a first-order rate of 1.9×10^{-3} s^{-1} (*4*, *34*) by measuring thiol-induced cleavage of the activated intermediate. It is a striking result that the rates of N-terminal thioester formation, C-terminal asparagine cyclization, and protein splicing are almost identical for the VMA intein. In contrast, differences of almost 2 orders of magnitude have been reported for the natively split *Ssp* DnaE intein (*35*). However, one should be aware that inteins probably have evolved to perfectly coordinate these individual reactions. Therefore, the dissection of these steps by introducing mutations at the key residues is likely to also affect the other catalytic center in the intein and will only allow us to study the reaction rates in first approximation.

Split Ssp DnaB and Split Sce VMA Inteins Are Orthogonal. For potential combined applications, we were interested whether the split DnaB and the split VMA inteins were orthogonal. Even though inteins exhibit a relatively low overall sequence similarity, the question of orthogonality is not trivial considering the high degree of structural conservation of the intein fold and the common general mechanism of protein splicing. Very recently, Shi and Muir showed that the split VMA intein is orthogonal to the natively split *Ssp* DnaE intein. They suggested that initial recognition between intein halves might be largely determined by electrostatic attraction. In fact, the calculated pI values for the split VMA and DnaE intein halves were reciprocal and would therefore favor orthogonality. However, in the case of the split DnaB ($DnaB^N$, pI = 9.5; $DnaB^C$, pI = 3.8) and split VMA intein halves (VMA^N, pI = 9.3; VMA^C, pI = 5.6) investigated in this study, the predicted pI values would be expected to also favor the formation of the mixed intein complexes. Nevertheless, we found that also these latter two split inteins are orthogonal. The inability to form active mixed complexes was first shown by the incubation of constructs **11** and **2** containing VMA^N and $DnaB^C$ as well as constructs **1** and **12** containing $DnaB^N$ and VMA^C. No formation of any splice products could be detected in the presence of rapamycin, which forces the heterologous intein fragments into close proximity by virtue of the fused FKBP and FRB domains (see lanes I and II in Figure 7). Next, we confirmed that the reactions of both split inteins can proceed parallelly without mutual interference. All four constructs, **1** and **2** as well as **11** and **12**, were mixed in the presence of rapamycin. Lane III in Figure 7 shows that splice products of the split DnaB and the split VMA intein can be identified (see parts A and C of Figure 2 for product designations). Only the bands corresponding to the constructs **2** and **12** and their corresponding intein fragments of the splicing products, **5** and **19**, respectively, lacking the His_6-tag C-terminal extein, could not be distinguished on the Coomassie-stained SDS−PAGE gel because of the very small differences in size. Thus, the split DnaB and the split VMA inteins were orthogonal; i.e., they were incapable of forming active inteins from mixed

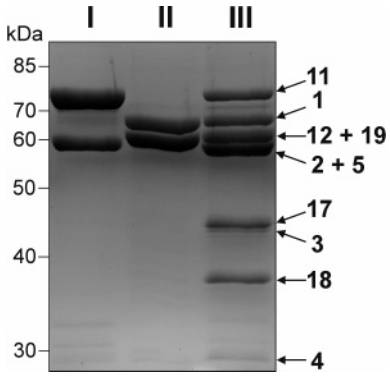

FIGURE 7: Orthogonality of the split DnaB and VMA inteins. Lane I of the Coomassie-stained SDS−PAGE gel shows construct **11** (VMA^N) incubated with construct **2** ($DnaB^C$) in the presence of the dimerizer rapamycin. The converse combination of **1** ($DnaB^N$) and **12** (VMA^C) is depicted in lane II. Lane III shows the reaction products of a one-pot incubation of both split DnaB constructs **1** and **2** and both split VMA constructs **11** and **12** in the presence of rapamycin. Reactions were carried out for 2 h at 25 °C. See parts A and C of Figure 2 for illustrations of both reactions. Molecular masses of all starting materials and products are given in the captions to Figures 3 and 5.

halves, and they still formed the active trans-splicing intein in the presence of the heterologous halves.

CONCLUSIONS

Split inteins are of particular interest for protein engineering approaches, because an active trans-splicing intein allows the assembly of a protein from two separately prepared polypeptide chains. For many of these applications, split intein fragments that can be purified and reacted under native conditions would be especially desirable, because they allowed us to circumvent denaturation of the protein of interest. Here, we show that the fragments of a split DnaB intein fulfill these criteria. To our knowledge, this is the only such example of an artificially split intein next to the VMA intein. The latter split intein, however, requires auxiliary domains to induce close proximity of the intein fragments for efficient protein trans-splicing, whereas the split DnaB intein fragments spontaneously form the active intein when mixed. The split DnaB intein appears particularly well-suited for protein engineering approaches. Importantly, for this split intein, a high extein tolerance has been demonstrated in a number of studies involving *in vivo* production of cyclic peptides and proteins (*36*, *37*). Moreover, this intein contains a serine instead of a cysteine as a C-terminal nucleophile, which generally appears to be a less problematic residue to be left behind at the splice junction in the generated protein of interest.

We can only speculate about the reason for the different behavior of the split DnaB and the split VMA inteins. Obviously, folding and packing properties of the individual intein fragments as well as presentation of recognition motives at their surface must come into play. The most striking difference between the two inteins on the structural level is an insertion of about 65 amino acids (residues 87−151) in the VMA^N fragment (*33*), which is not found in the minimal splicing domain of the DnaB intein (*38*). This region, parts of which have been implicated in DNA binding in cooperation with the homing endonuclease domain (*39*), is also not present in most other inteins, such as the *Mtu* RecA intein, as judged by sequence alignments (data not shown). Possibly, this part of the VMA^N fragments prevents effective spontaneous dimerization with VMA^C.

It is striking that both the artificially split DnaB and the VMA inteins exhibit rate constants for protein trans-splicing that significantly exceed the rate constant reported for the only naturally occurring split intein, the *Ssp* DnaE intein (6.6×10^{-5} s^{-1}) (*35*). The rate constants determined in this paper are about 15-fold (split DnaB, 9.9×10^{-4} s^{-1}) to 30-fold (split VMA, from 9.4×10^{-4} to 2.0×10^{-3} s^{-1}) higher. Probably, the $DnaE^N$ and $DnaE^C$ fragments have been mostly optimized for stability and affinity during evolution.

In summary, the split DnaB intein represents a new intein to assemble proteins of interest from two separate polypeptides under native conditions. We have also shown that it is orthogonal to the split VMA intein, which suggests that combinations of these inteins can be used for three-fragment protein trans-splicing reactions. We believe that such inteins are valuable additions to the protein chemist's toolbox.

ACKNOWLEDGMENT

We are grateful to Uwe Linne for support with mass spectrometry and Tim Sonntag for help with cloning procedures. We thank Mohamed Marahiel for general support. We thank Xiang-Qin Liu for providing plasmid pMST. ARIAD Pharmaceuticals provided DNA templates for FKBP and FRB domains.

REFERENCES

1. Noren, C. J., Wang, J., and Perler, F. B. (2000) Dissecting the chemistry of protein splicing and its applications, *Angew. Chem. Int. Ed. 39*, 450−466.
2. Paulus, H. (2000) Protein splicing and related forms of protein autoprocessing, *Annu. Rev. Biochem. 69*, 447−496.
3. Giriat, I., Muir, T. W., and Perler, F. B. (2001) Protein splicing and its applications, *Genet. Eng. 23*, 171−199.
4. Evans, T. C., Jr., and Xu, M.-Q. (2002) Mechanistic and kinetic considerations of protein splicing, *Chem. Rev. 102*, 4869−4883.
5. Perler, F. B. (2002) InBase: The intein database, *Nucleic Acids Res. 30*, 383−384.
6. Gerth, M. L., Patrick, W. M., and Lutz, S. (2004) A second-generation system for unbiased reading frame selection, *Protein. Eng. Des. Sel. 17*, 595−602.
7. Zeidler, M. P., Tan, C., Bellaiche, Y., Cherry, S., Hader, S., Gayko, U., and Perrimon, N. (2004) Temperature-sensitive control of protein activity by conditionally splicing inteins, *Nat. Biotechnol. 22*, 871−876.
8. Muir, T. W., Sondhi, D., and Cole, P. A. (1998) Expressed protein ligation: A general method for protein engineering, *Proc. Natl. Acad. Sci. U.S.A. 95*, 6705−6710.
9. Evans, T. C., Jr., Benner, J., and Xu, M. Q. (1998) Semisynthesis of cytotoxic proteins using a modified protein splicing element, *Protein. Sci. 7*, 2256−2264.
10. Chong, S., Mersha, F. B., Comb, D. G., Scott, M. E., Landry, D., Vence, L. M., Perler, F. B., Benner, J., Kucera, R. B., Hirvonen, C. A., Pelletier, J. J., Paulus, H., and Xu, M. Q. (1997) Single-column purification of free recombinant proteins using a self-cleavable affinity tag derived from a protein splicing element, *Gene 192*, 271−281.
11. Mathys, S., Evans, T. C., Chute, I. C., Wu, H., Chong, S., Benner, J., Liu, X. Q., and Xu, M. Q. (1999) 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 231*, 1−13.
12. Yamazaki, T., Otomo, T., Oda, N., Kyogoku, Y., Uegaki, K., Ito, N., Ishino, Y., and Nakamura, H. (1998) Segmental isotope labeling for protein NMR using peptide splicing, *J. Am. Chem. Soc. 120*, 5591−5592.

13. Otomo, T., Ito, N., Kyogoku, Y., and Yamazaki, T. (1999) NMR observation of selected segments in a larger protein: Central-segment isotope labeling through intein-mediated ligation, *Biochemistry 38*, 16040−16044.
14. Yagi, H., Tsujimoto, T., Yamazaki, T., Yoshida, M., and Akutsu, H. (2004) Conformational change of H^+-ATPase β monomer revealed on segmental isotope labeling NMR spectroscopy, *J. Am. Chem. Soc*. *126*, 16632−16638.
15. Lew, B. M., Mills, K. V., and Paulus, H. (1998) Protein splicing *in vitro* with a semisynthetic two-component minimal intein, *J. Biol. Chem. 273*, 15887−15890.
16. Lew, B. M., Mills, K. V., and Paulus, H. (1999) Characteristics of protein splicing *in trans* mediated by a semisynthetic split intein, *Biopolymers 51*, 355−362.
17. Evans, T. C., Jr., Martin, D., Kolly, R., Panne, D., Sun, L., Ghosh, I., Chen, L., Benner, J., Liu, X. Q., and Xu, M. Q. (2000) Protein trans-splicing and cyclization by a naturally split intein from the *dnaE* gene of *Synechocystis species* PCC6803, *J. Biol. Chem. 275*, 9091−9094.
18. Giriat, I., and Muir, T. W. (2003) Protein semi-synthesis in living cells, *J. Am. Chem. Soc. 125*, 7180−7181.
19. Shi, J., and Muir, T. W. (2005) Development of a tandem protein trans-splicing system based on native and engineered split inteins, *J. Am. Chem. Soc. 127*, 6198−6206.
20. Burbaum, J. J., and Schimmel, P. (1991) Assembly of a class I tRNA synthetase from products of an artificially split gene, *Biochemistry 30*, 319−324.
21. Kippen, A. D., Sancho, J., and Fersht, A. R. (1994) Folding of barnase in parts, *Biochemistry 33*, 3778−3786.
22. Gross, M., Wyss, M., Furter-Graves, E. M., Wallimann, T., and Furter, R. (1996) Reconstitution of active octameric mitochondrial creatine kinase from two genetically engineered fragments, *Protein. Sci. 5*, 320−330.
23. Mills, K. V., Lew, B. M., Jiang, S., and Paulus, H. (1998) Protein splicing in trans by purified N- and C-terminal fragments of the *Mycobacterium tuberculosis* RecA intein, *Proc. Natl. Acad. Sci. U.S.A. 95*, 3543−3548.
24. Southworth, M. W., Adam, E., Panne, D., Byer, R., Kautz, R., and Perler, F. B. (1998) Control of protein splicing by intein fragment reassembly, *EMBO J. 17*, 918−926.
25. Shingledecker, K., Jiang, S. Q., and Paulus, H. (1998) Molecular dissection of the *Mycobacterium tuberculosis* RecA intein: Design of a minimal intein and of a trans-splicing system involving two intein fragments, *Gene 207*, 187−195.
26. Mootz, H. D., and Muir, T. W. (2002) Protein splicing triggered by a small molecule, *J. Am. Chem. Soc. 124*, 9044−9045.
27. Mootz, H. D., Blum, E. S., Tyszkiewicz, A. B., and Muir, T. W. (2003) Conditional protein splicing: A new tool to control protein structure and function *in vitro* and *in vivo*, *J. Am. Chem. Soc. 125*, 10561−10569.
28. Mootz, H. D., Blum, E. S., and Muir, T. W. (2004) Activation of an autoregulated protein kinase by conditional protein splicing, *Angew. Chem. Int. Ed. 43*, 5189−5192.
29. Brenzel, S., and Mootz, H. D. (2005) Design of an intein that can be inhibited with a small molecule ligand, *J. Am. Chem. Soc. 127*, 4176−4177.
30. Choi, J., Chen, J., Schreiber, S. L., and Clardy, J. (1996) Structure of the FKBP12−rapamycin complex interacting with the binding domain of human FRAP, *Science 273*, 239−242.
31. Rollins, C. T., Rivera, V. M., Woolfson, D. N., Keenan, T., Hatada, M., Adams, S. E., Andrade, L. J., Yaeger, D., van Schravendijk, M. R., Holt, D. A., Gilman, M., and Clackson, T. (2000) A ligand-reversible dimerization system for controlling protein−protein interactions, *Proc. Natl. Acad. Sci. U.S.A. 97*, 7096−7101.
32. Wu, H., Xu, M. Q., and Liu, X. Q. (1998) Protein trans-splicing and functional mini-inteins of a cyanobacterial dnaB intein, *Biochim. Biophys. Acta 1387*, 422−432.
33. Duan, X., Gimble, F. S., and Quiocho, F. A. (1997) Crystal structure of PI-*Sce*I, a homing endonuclease with protein splicing activity, *Cell 89*, 555−564.
34. Chong, S., Williams, K. S., Wotkowicz, C., and Xu, M. Q. (1998) Modulation of protein splicing of the *Saccharomyces cerevisiae* vacuolar membrane ATPase intein, *J. Biol. Chem. 273*, 10567−10577.
35. Martin, D. D., Xu, M. Q., and Evans, T. C., Jr. (2001) Characterization of a naturally occurring trans-splicing intein from *Synechocystis* sp. PCC6803, *Biochemistry 40*, 1393−1402.
36. Kinsella, T. M., Ohashi, C. T., Harder, A. G., Yam, G. C., Li, W., Peelle, B., Pali, E. S., Bennett, M. K., Molineaux, S. M., Anderson, D. A., Masuda, E. S., and Payan, D. G. (2002) Retrovirally delivered random cyclic peptide libraries yield inhibitors of interleukin-4 signaling in human B cells, *J. Biol. Chem. 277*, 37512−37518.
37. Williams, N. K., Prosselkov, P., Liepinsh, E., Line, I., Sharipo, A., Littler, D. R., Curmi, P. M., Otting, G., and Dixon, N. E. (2002) *In vivo* protein cyclization promoted by a circularly permuted *Synechocystis* sp. PCC6803 DnaB mini-intein, *J. Biol. Chem. 277*, 7790−7798.
38. Ding, Y., Xu, M. Q., Ghosh, I., Chen, X., Ferrandon, S., Lesage, G., and Rao, Z. (2003) Crystal structure of a mini-intein reveals a conserved catalytic module involved in side chain cyclization of asparagine during protein splicing, *J. Biol. Chem. 278*, 39133−39142.
39. Christ, F., Steuer, S., Thole, H., Wende, W., Pingoud, A., and Pingoud, V. (2000) A model for the PI-*Sce*I×DNA complex based on multiple base and phosphate backbone-specific photocross-links, *J. Mol. Biol. 300*, 867−875.

BI051697+