SUMMARY

The term intein is used for an internal protein domain, that catalyses its own excision out of a precursor protein and the concomitant linkage of the two flanking sequences, the N- and C-extein, with a native peptide bond. In split inteins, the intein domain is divided into two parts that undergo fragment association followed by protein splicing *in trans*. The specificity and sequence promiscuity of split inteins make them a generally useful biochemical tool, for example for the preparation of proteins out of two fragments. If the intein is used to ligate a synthetic and a recombinant fragment the splice product represents a semi-synthetic protein.

The new approach outlined in this work is based on the artificially split *Ssp* DnaB mini-intein. The splitting was performed at a variable loop position close to the N-terminus leading to a synthetically accessible N-terminal intein fragment (Int^N) and a recombinantly expressible C-terminal part (Int^C), which are 11 and 143 amino acids in length, respectively. This intein enables the chemo-enzymatic synthesis of N-terminally modified semi-synthetic proteins.

The feasibility of the new approach was shown experimentally using two model systems, wherein a synthetic N-extein peptide pentamer, including a 5,6-Carboxyfluorescein moiety, was ligated to the N-terminus of the 12 kD protein thioredoxin and the 31 kD protein β-lactamase. After 24 h of incubation at 25°C, under native conditions and using a peptide and protein concentration in the micromolar range, the desired splice product could be formed with a maximal yield of approximately 70%. Furthermore, the reaction also proceeded selectively in an *E. coli* cell lysate. The applicability in such complex mixtures together with the fact, that this intein enables the generation of N-terminally modified semi-synthetic proteins without a necessary incorporation of thioester groups and amino terminal cysteins, are the main advantages compared to *Expressed Protein Ligation*.

For a detailed characterisation of the semi-synthetic intein the minimal sequence requirements of Int^N were evaluated. Measurements of changes in fluorescence anisotropy were carried out to investigate the inherent affinity of both intein fragments. For that purpose an Int^C mutant blocked in protein splicing was used, leading to a dissociation constant of 1.1 μM and a rate constant of complex formation of 16.8 $M^{-1}s^{-1}$. Furthermore, based on an alanine scanning analysis of the Int^N(1-11) sequence, three for the complex association essential hydrophobic residues could be identified (Ile2, Ile8 and Leu10), whereas substitutions at other positions were tolerated.

Additionally, the semi-synthetic intein was well suited for further mechanistic studies of the protein splicing mechanism. Intriguingly, a Gly[−1]Ala-substitution of the first amino acid within the N-extein, was revealed to result in thiazoline ring formation involving the catalytic Cys1, likely by aberrant dehydration of an oxythiazolidine-anion intermediate. This finding provides the first experimental evidence for this postulated intermediate during the initial N,S → acyl shift. A second interesting side reaction within the splicing mechanism, the formation of a caprolactone, could be induced using an N-methylated amino acid at position [−1] and a serin residue at position [−2]. Further unnatural amino acids were successfully introduced at delicate positions within the synthetic Int^N-fragment as well as the N-extein sequence, and their impact on the spatially and temporally coordinated steps of the splicing mechanism was evaluated.