Available online at www.sciencedirect.com
ScienceDirect
Biochemical and Biophysical Research Communications 366 (2008) 239–243
BBRC
www.elsevier.com/locate/ybbrc

Minimization of a eukaryotic mini-intein

Skander Elleuche, Kristin Döring, Stefanie Pöggeler *

Institute of Microbiology and Genetics, Department of Genetics of Eukaryotic Microorganisms, Georg-August-University of Göttingen, Grisebachstrasse 8, 37077 Gottingen, Germany

Received 5 November 2007
Available online 3 December 2007

Abstract

Inteins are internal protein splicing elements that can autocatalytically self-excise from their host protein and ligate the protein flanks (exteins) with a peptide bond. Large inteins comprise independent protein splicing and endonuclease domains whereas mini-inteins lack the central endonuclease domain. To identify mini-intein domains that are essential for protein splicing, deletions were introduced at different sites of the 157-aa PRP8 mini-intein of *Penicillium chrysogenum*. The removal of eight and six amino acids at two different sites resulted in a functional eukaryotic mini-intein of only 143 aa.

Keywords: Intein; Protein splicing; *Penicillium chrysogenum*; PRP8 protein; Minimization

An intein is a self-catalytic sequence that is embedded in-frame within a precursor protein. To produce a mature protein the intein precisely self-excises from the host protein. It thereby ligates the flanking sequences (N- and C-exteins) [1,2]. The post-translational removal of the intein is termed protein splicing and depends on three highly conserved residues. Most inteins begin with a serine, cysteine, or threonine residue at the N terminus, and end with an asparagine residue at the C terminus. The first position of the C-terminal extein is either a hydroxyl-containing (serine) or a thiol-containing (cysteine) residue. Protein splicing typically involves four steps: (i) an N-O or N-S acyl shift at the N-extein/intein junction; (ii) a transesterification which transfers the N-extein to the side-chain of the first residue (+1) of the C-extein; (iii) cyclization of a conserved asparagine residue at the C terminus of the intein and cleavage of the peptide bond, resulting in the release of the intein; and (iv) rearrangement to a peptide bond of the ester/thioester bond linking the N- and C-exteins [3–6].

Inteins defined as being large or minimal (mini-) depending on whether or not they contain an endonuclease domain [2]. The splicing domain of large inteins seems to be split by the endonuclease domain into an N-terminal and a C-terminal subdomain. Because several splicing-efficient mini-inteins have been engineered from large inteins by deleting the central endonuclease domain, it is obvious that the endonuclease domain is not involved in protein splicing [7–10]. Both intein subdomains contain conserved blocks of amino acids (blocks A, N2, B, and N4 for the N-terminal subdomain and blocks G and F for the C-terminal subdomain) [11–14]. These domains can also be identified in mini-inteins. Three-dimensional structures of inteins reveal that the N- and C-terminal splicing domains form a common horseshoe-like 12-β-strand scaffold termed the Hedgehog/intein (Hint) module [15–18].

Most inteins have been found in the genomes of prokaryotes. In eukaryotes, they have been identified in ~50 taxa, encoded within the nuclear genomes of fungi and the plastomes of some unicellular algae [12]. In fungi, most of the inteins (13 large and 19 mini-inteins) have been detected in the *prp8* gene [12,19]. The PRP8 protein is one of the largest and most highly conserved nuclear proteins occupying a central position in the catalytic core of the spliceosome [20].

In a previous study [21], we identified that the PRP8 protein of *Penicillium chrysogenum* (*Pch*), the major pro-

* Corresponding author. Fax: +49 551 3910123.
E-mail address: spoegge@gwdg.de (S. Pöggeler).

0006-291X/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2007.11.126

ducer of the β-lactam antibiotic penicillin, possesses a 157-aa mini-intein in the PRP8 protein. The *Pch* PRP8 mini-intein can undergo autocatalytic protein splicing when heterologously expressed in a model host protein in *Escherichia coli* [21]. Thus, the *Pch* PRP8 intein is among the smallest known nuclear-encoded active splicing protein elements. Moreover, we recently demonstrated that the *Pch* intein is capable of protein splicing in *trans* [22]. Several studies of the sequence and structural requirements of protein splicing have been conducted for large inteins [7–10]. However, little is known about the catalytic and structural elements that participate in protein splicing of naturally occurring mini-inteins. Here, we report the minimization of the eukaryotic *Pch* PRP8 mini-intein. Amino acid sequences of this intein can be deleted at two different sites without affecting splicing activity. One site corresponds to the insertion site of the endonuclease domain in large allelic PRP8 inteins. The other site was detected at a new position corresponding to the insertion site of a putative tongs domain of a large fungal PRP8 intein. The smallest functional intein found in our study comprises only 143 residues and is the smallest functional eukaryotic intein engineered so far.

Materials and methods

Sequence analysis. Protein sequence alignments were performed either using the ClustalX program [23] or the LALIGN program [24]. Protein secondary structure prediction was done with the PSIPRED server (http://bioinf.cs.ucl.ac.uk/psipred/) [25].

Construction of truncated P. chrysogenum inteins. Escherichia coli strain SURE (Stratagene, La Jolla, CA, USA) was used for general plasmid construction. Cloning and propagation was done under standard conditions [26]. Deletion of selected regions of the *Pch* PRP8 intein (AM042015) was done by inverse PCR using plasmid pGPch-1 as a circular template [21,27]. Primer pair Endo_for (CAAACGCACAGTTTCAAGA TTGAGCAAGTTAGCCTCG) and Endo_rev (CTTCTCAGCCCGGG GAGATGTAAAGGC, SmaI restriction site underlined) were used to construct plasmid pΔE. The SmaI restriction site in primer Endo_rev was used for the selection of positive clones. Amplification with primer pair TSD_for (CAAACGGTGGAGATCACTGCTGCCGAGTTTGCCGC GCTTTCTACC) and TSD_rev (CTCTCTATAAAGCACCAGAATA TGGTTCGGCGTCACCACAAGGTCC) resulted in plasmid pΔT and with TSD_for and TSD2_rev (CACCAGAATATGGTTCGGCGTCAC CACAAGGTCC) in plasmid pΔTΔ4. To construct plasmid pΔT-E, TSD_rev and Endo_for were used as primer pair. Plasmid pΔTΔE was generated by means of inverse PCR with plasmid pΔE as template and primers TSD_for and TSD-rev. Inverse PCRs were performed using Phusion™ High-Fidelity DNA Polymerase (New England Biolabs GmbH, Frankfurt am Main, Germany). Subsequently, amplicons were treated with Klenow Polymerase (Fermentas, St. Leon-Rot, Germany) to generate blunt-ended products and with polynucleotide kinase (Fermentas, St. Leon-Rot, Germany) to generate phosphorylated 5′-ends. After ligation of the linear PCR-fragments, plasmids were transformed into the bacterial host strain. Positive clones were screened by colony PCR [28] using MolTaq (Molzym GmbH & Co., KG, Bremen, Germany). The DNA sequences of all plasmids were verified by DNA sequencing at the G2L-sequencing service of the Göttinger Genom Labor (Georg-August University, Göttingen, Germany).

Protein expression in E. coli. Plasmid encoded intein constructs were overexpressed in *E. coli* strain BL21 (DE3) Gold (Stratagene, La Jolla, CA, USA) as described earlier [21]. In brief, protein production was carried out in 50 ml of Luria Broth (LB) medium at 37 °C containing ampicillin (60 μg/ml), which was inoculated from a 5 ml overnight culture. Cells were grown under constant shaking at 37 °C to a culture density of $A_{600} = 0.6$, subsequently, isopropyl-1-thio-β-D-galctopyranoside (IPTG) was added to a final concentration of 0.5 mM to induce production of recombinant proteins. After incubation for 4 h, cells were harvested by centrifugation and lysed in SDS-containing gel-loading buffer at 98 °C in a heating block.

SDS–polyacrylamide gel electrophoresis (PAGE) and Western blotting analysis. Proteins were separated in a 15% SDS–polyacrylamide gel and transferred to a polyvinylidene difluoride membrane (Biometra, Göttingen, Germany) by using a semidry blotting system (Biometra, Göttingen, Germany). Western blots were either done with a polyclonal anti-His-RGS antibody (Qiagen, Hilden, Germany) or with an anti-glutathione S-transferase (GST) HRP-conjugate (Amersham Bioscience, Europe GmbH, Freiburg, Germany). Signal detection was carried out using the chemiluminescence Western blotting kit (Roche, Mannheim, Germany) according to the supplier's recommendations.

Results and discussion

Selection of deletion sites

Little is known about the catalytic and structural elements that participate in the protein splicing of eukaryotic mini-inteins. As a guide to constructing deletion mutations, the *Pch* PRP8 mini-intein was aligned with the related large PRP8 intein of *Aspergillus fumigatus* (*Afu*) which is known to splice in a recombinant protein in *E. coli* [29]. *Afu* PRP8 intein is unusual because, in addition to a 454-aa endonuclease domain, it has a 222-aa extra sequence which is predicted to be a putative tongs subdomain [29]. This prediction was based solely on the location of this domain being the same as the tongs subdomain in the crystal structure of vacuolar ATPase subunit (VMA) intein of the yeast *Saccharomyces cerevisiae* (*Sce*). In the *Sce* VMA, the tongs domain is supposed to participate in DNA substrate binding [30]. Alignment of the *Pch* PRP8 mini-intein and the *Afu* PRP8 large intein without the tongs and endonuclease domains revealed that the proteins share a high degree of amino acid identity within the first 72 N-terminal residues and 40 C-terminal residues. By contrast, the middle parts displayed a low level of sequence conservation. Conserved motifs A, N2, and B were identified within the conserved N-terminal subdomains and motifs F and G within the conserved C termini (Fig. 1). The N4 motif, comprising 16 residues and including a highly conserved Asp or Glu residue [13] could not be allocated in either the *Pch* or the *Afu* PRP8 intein. The dependence of protein splicing on conserved residues within motifs A, B, F, and G has recently been assessed for the PRP8 mini-intein of the basidiomycete *Cryptococcus neoformans* [31]. Thus, a deletion mutation (pΔT-E, a deletion of 45 aa) was constructed so that its deleted area matched the poorly conserved region between the *Afu* and the *Pch* inteins. The protein expressed from pΔT-E was 112 aa and comprised only the conserved N- and C-terminal subdomains (Fig. 2).

To investigate amino acid residues surrounding putative tongs and/or endonuclease domain insertion sites, we constructed the plasmids pΔT (deletion of 8 aa), pΔE (deletion of 6 aa), and pΔTΔE (deletion of 8 aa and 6 aa). We took