Backbone circularization of *Bacillus subtilis* family 11 xylanase increases its thermostability and its resistance against aggregation

Max C. Waldhauer, Silvan N. Schmitz, Constantin Ahlmann-Eltze, Jan G. Gleixner, Carolin C. Schmelas, Anna G. Huhn, Charlotte Bunne, Magdalena Büscher, Max Horn, Nils Klughammer, Jakob Kreft, Elabeth Schäfer, Philipp A. Bayer, Stephen J. Krämer, Julia Neugebauer, Pierre Wehler, Matthias P. Mayer, Roland Eils and Barbara Di Ventura

Electronic Supplementary Information

Contents

Fig. S1: Schematic overview of intein-mediated protein splicing.................................................................2
Fig. S2: Analysis of expression of constructs in *E. coli*...................................................................................3
Fig. S3: Temperature dependence of catalytic activities for the circular and linear RGKCWE variants and native xylanase..................................................................................................................4
Fig. S4: Impurities present in the circular S variant after purification with CM Sephadex C-50 before incubation at 60 °C.........................................................................................................................5
Table S1: Oligonucleotides used for the construction of the plasmids..............................................................6
Protocol S1: Subcloning into pCIRC_intein....................................................................................................7
References.........................................................................................................................................................9
**Fig. S1** Schematic overview of intein-mediated protein a) cis- and b) trans-splicing. We depict here inteins artificially included in proteins to control their function. In this case, we indicate with N- and C-extein the residues that need to be introduced for efficient splicing and not the entire fragments that are joined after splicing.
Fig. S2 Analysis of expression of constructs in *E. coli*. SDS-PAGE of crude cell lysates of bacterial cultures expressing the indicated constructs. The different exteins used for circularization are indicated in bold. *Npu, Npu* DnaE split intein. Arrows point to the position in the gel where the linear (lin) or circular (circ) xylanases migrated.
Fig. S3  Temperature dependence of catalytic activities for the circular and linear RGKCWE variants and native xylanase. Enzymes were used at a final concentration of 2 µg/mL in the presence of 1.26 % glycerol and 0.25 mM dithiothreitol. The data points are normalized to the local maximum of a fitted 3rd degree polynomial for each variant. Error bars represent the standard error of the mean for three independent replicates.
Fig. S4  Impurities present in the circular S variant after purification with CM Sephadex C-50 before incubation at 60 °C. SDS-PAGE of the indicated construct. Other lanes in the gel are not shown and are indicated by a white space. Arrows point to the position in the gel where the linear (lin) or circular (circ) xylanases migrated.
<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence (5’− 3’)</th>
<th>Target</th>
<th>Amplicon usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>pSBX_F_taataa</td>
<td>TAATAACTGGTGAATGCGACAGCTGTAG</td>
<td>pSB1K30-T7RBS-IGEMHD</td>
<td>Vector fragment for pCIRC&lt;sub&gt;3pa&lt;/sub&gt; and pCIRC&lt;sub&gt;p21-l&lt;/sub&gt;</td>
</tr>
<tr>
<td>pSBX_R_stg</td>
<td>CATCTGATCTCTCCTCTCTTTAAATGTAAACAAATATTTTC</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| NpuDnaEC_F_pSB-T7RBS        | TCTAGAGAATATTGTTTATTAAGAAGGAGATCTAGAGTCATACATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCAAAGATGACGCA
Protocol S1  Subcloning into pCIRC\textsubscript{Intein}.

I. PCR-amplify the coding sequence of your protein of interest

1. Design primers for the coding sequence of your protein of interest using standard primer design guidelines, so that the amplicon starts at the codon of the first amino acid and ends at the codon of the last amino acid to be included in the circular protein. (This assumes that the coding sequence is read in the forward direction; otherwise, the roles of forward and reverse primer are interchanged.) The reading frame of the amplicon must not contain a stop codon.

2. Depending on the application, add extein residues to the 5' end of your primers. For the splicing reaction to succeed, the first amino acid must be a Cys, Ser, or Thr (but, depending on the intein, not all of them may work). When adding to the reverse primer, make sure to add the reverse complement of the sequence coding for the extein residues.

3. Further add the following sequence to the 5' end of your forward primer:
   \[
   \text{CTGGTCTCACAAAC}
   \]
   and the following sequence to the 5' end of your reverse primer:
   \[
   \text{CTGGTCTCTAGCA}
   \]

4. Order the primers and perform the PCR reaction with a polymerase of your choice. Refer to the manufacturer's instructions to determine the conditions of the amplification.

5. Purify the amplified DNA, e. g. by gel extraction, and determine the concentration, e. g. by spectrophotometric measurement.

II. Insert the coding sequence into pCIRC\textsubscript{Intein} (where ‘intein’ stands for the selected split intein to be used) via golden gate assembly

1. In a PCR tube, mix:
   - 100 ng pCIRC\textsubscript{Intein} DNA
   - An equimolar amount of the insert DNA. It can be calculated from the length of the insert amplicon and the length of the vector (pCIRC\textsubscript{Npu}: 3817 bp, pCIRC\textsubscript{gp41-1}: 3778 bp) as follows:
   \[
   100 \text{ ng} \cdot \frac{\text{length of insert}}{\text{length of vector}}
   \]
   - 1.5 µL 10X CutSmart Buffer (New England Biolabs, Ipswich, MA, USA)
   - ATP to a final concentration of 1 mM
   - 1 µL BsaI (New England Biolabs, Ipswich, MA, USA)
   - 1 µL T4 DNA Ligase (New England Biolabs, Ipswich, MA, USA)
   - H\textsubscript{2}O ad 15 µL
2. Place the tube into a thermocycler and run the following program:

- 37 °C 3 min
- 16 °C 4 min
- 50 °C 5 min
- 80 °C 5 min

25 cycles

3. Transform suitable competent *E. coli* with 5 µL of the reaction mixture

The procedure will work in a variety of buffers, including e.g. T4 DNA Ligase buffers (New England Biolabs, Ipswich, MA, USA or Thermo Fisher Scientific, Waltham, MA, USA), as long as sufficient amounts of both ATP (required by T4 DNA Ligase) and BSA (required by Bsal for optimal activity at 37 °C) are present.
References