**Dockerin-containing protease inhibitor protects** **key cellulosomal cellulases from proteolysis in *Clostridium cellulolyticum***

**Supporting Information**

**Tao Xu1,2,Yongchao Li1,2, Zhili He1,5,andJizhong Zhou 1,3,4,5,\***

1*Institute for Environmental Genomics,* 2*Department of Microbiology and Plant Biology,* *University of Oklahoma, Norman, Oklahoma 73071, USA;*

3*Earth Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA;*

4*Department of Environmental Science and Engineering, Tsinghua University, Beijing 100084, China; and* 5*Virtual Institute for Microbial Stress and Survival*

\*For correspondence. E-mail [jzhou@ou.edu](mailto:jzhou@ou.edu); Tel. (+1) 405 325 6073; Fax (+1) 405 325-7552

**Running title:** Role of Dpi in protecting cellulases

**Key words:** cellulosome; cellulase; protease inhibitor; biofuels; *Clostridium cellulolyticum***Table S1 Mass spectrometry analysis of gel slices from SDS-PAGE.**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Hit** | **Accession** | **Mascot Score** | **Mass (Da)** | **Avg. Intensity** | **Gene number** | **Predicted functions** |
| B1 | gi|220928179 | 2608 | 80608 | 5.611e+4 | Ccel\_0729 | Cel48F |
| B2 | gi|220928182 | 3301 | 97127 | 9.375e+4 | Ccel\_0732 | Cel9E |
| B3 | gi|220928185 | 1271 | 85039 | 1.659e+4 | Ccel\_0735 | Cel9J |
| B4 | gi|220928187 | 1058 | 58027 | 2.347e+4 | Ccel\_0737 | Cel9M |

**Table S2 Oligonucleotide primers in this study.**

|  |  |  |
| --- | --- | --- |
| Primer name | Sequencea | Application |
| EBSu | CGAAATTAGAAACTTGCGTTCAGTAAAC | Intron modification |
| Dpi-171|172a-IBS | AAAACCCGGGATAATTATCCTTAGACCTCATTATCGTGCGCCCAGATAGGGTG | Intron modification |
| Dpi-171|172a-EBS1d | CAGATTGTACAAATGTGGTGATAACAGATAAGTCATTATCACTAACTTACCTTTCTTTGT | Intron modification |
| Dpi-171|172a-EBS2 | TGAACGCAAGTTTCTAATTTCGATTAGGTCTCGATAGAGGAAAGTGTCT | Intron modification |
| Cel48F-764a-IBS | AAAACCCGGGATAATTATCCTTACTGGACCATTTGGTGCGCCCAGATAGGGTG | Intron modification |
| Cel48F-764a-EBS1d | CAGATTGTACAAATGTGGTGATAACAGATAAGTCCATTTGTATAACTTACCTTTCTTTGT | Intron modification |
| Cel48F-764a-EBS2 | TGAACGCAAGTTTCTAATTTCGGTTTCCAGTCGATAGAGGAAAGTGTCT | Intron modification |
| Cel9E-653a-IBS | AAAACCCGGGATAATTATCCTTACTACCCGATTCAGTGCGCCCAGATAGGGTG | Intron modification |
| Cel9E-653a-EBS1d | CAGATTGTACAAATGTGGTGATAACAGATAAGTCGATTCAAATAACTTACCTTTCTTTGT | Intron modification |
| Cel9E-653a-EBS2 | TGAACGCAAGTTTCTAATTTCGGTTGGTAGTCGATAGAGGAAAGTGTCT | Intron modification |
| Dpi-171 F | TTGCTCCGGCAAAAGTAAAC | Mutant identification |
| Dpi-171 R | CACTGATAGCCCGTTGATCC | Mutant identification |
| Cel48F F | GATGAACATAAATTTGGTGGACAGT | Mutant identification |
| Cel48F R | TGCATAGTACCATGAAAGCAGATAA | Mutant identification |
| Cel9E F | CTGGAATTACAGGCTAATACTCCAA | Mutant identification |
| Cel9E R | TGCAATACCACCATTAACAACATAC | Mutant identification |
| Intron F1 | CCTATGGGAACGAAACGAAA | Mutant identification |
| Intron R1 | CGAGTACTCCGTACCCTTGC | Mutant identification |
| NtDpi F(NdeI) | GGAATTCCATATGGTGGTAGGAAGTTATACACTTTTCGG | Construct pET28a(+)-Dpi expression vector |
| NtDpi R(Not I) | ATAGTTTAGCGGCCGCTTAAATTACATTTATTTCACATTGG | Construct pET28a(+)-Dpi expression vector |
| Dpi-over F | CGCGGATCCCCCGGGATGGAAAAGAATTACACACCAA | Construct pClostron3-Dpiover complementation vector |
| Dpi-over R | TTATTTCGATCGTTAAATTACATTTATTTCACA | Construct pClostron3-Dpiover complementation vector |
| RTrecA F | GCAAAGAAACTTGGGGTTGA | *recA* qPCR |
| RTrecA R | TGAGACATCAGCCTTGCTTG | *recA* qPCR |
| RTcipC F | TACTGGCGTCGTATCAGTGC | *cipC* qPCR |
| RTcipC R | TGTCCGCATCCTGAGTGTAA | *cipC* qPCR |
| RTcel48F F | AACAAACCGGCTACATACGC | *cel48F* qPCR |
| RTcel48F R | GGTTCCATCAGCTCTTGCTC | *cel48F* qPCR |
| RTcel9E F | ACCTGGACCGTAATGAATGC | *cel9E* qPCR |
| RTcel9E R | TCATGAGCTTTGTGGTGAGC | *cel9E* qPCR |
| RTcel8C F | GGATACGGTTTGCTGCTTTC | cel8C qPCR |
| RTcel8C R | AGCAAACACAAGGGATACCG | cel8C qPCR |
| RTorfX F | AAGCAGCAACAGTGGTAAGG | *orfX* qPCR |
| RTorfX R | AATGCACCGGAAGTACCTTG | *orfX* qPCR |

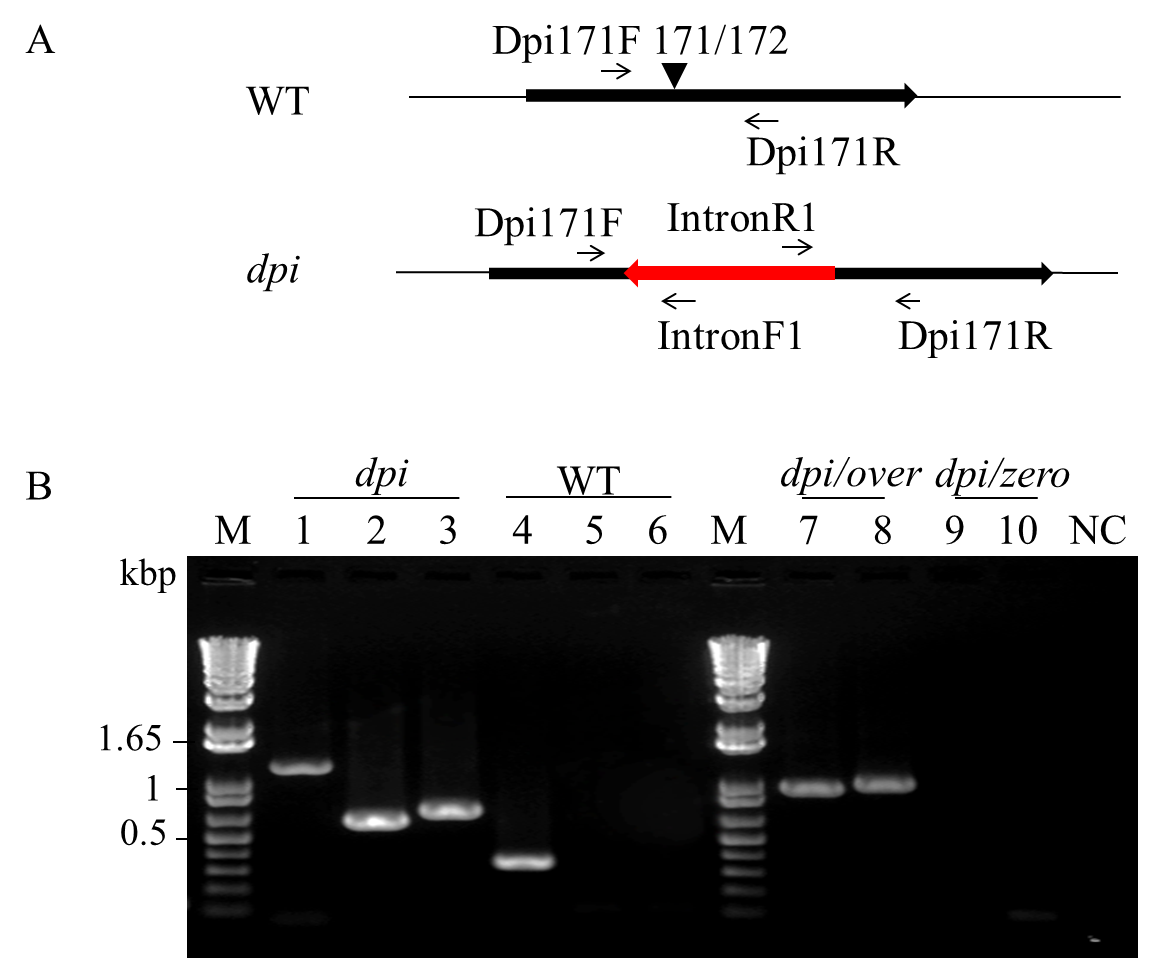


Fig. S1 Diagram of intron insertion in *dpi* gene and strains identification. A. The group II intron (red arrow) potentially inserted into the *dpi* ORF (bold black arrow) at 171/172nt in the anti-sense direction. Small arrows indicated locations of four primers (Dpi171F, Dpi171R, IntronF1 and Intron R1) that were applied to identify the anticipated intron insertion. B. PCR identification. Primers used in each PCR reaction are as follows: Dpi171F-Dpi171R (lane 1 and 4); Dpi171F-IntroF1 (lane 2 and 3), IntronR1-Dpi171R (lane 5 and 6); pClostron3RBSF-Dpi overexpR (lane 7 and 9); pClostron3RBSF-pClostron3seqR (lane 8 and 10). NC indicates negative control without any templates in the PCR system.

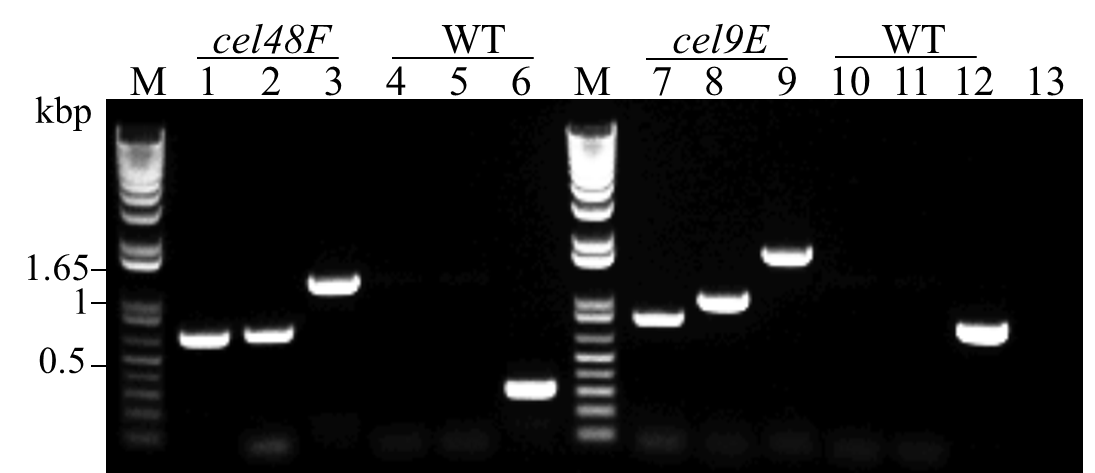


Fig. S2 Identification of *cel48F* and *cel9E* anti-sense mutants by PCR. Primers for each PCR reaction are as follows: 1, Cel48FF- intronF1; 2, intronR1-Cel48FR; 3, Cel48FF-Cel48FR; 4, Cel48FF-intronF1; 5, intronR1-Cel48FR; 6, Cel48FF-Cel48FR; 7, Cel9EF-intronF1; 8, intronR1-Cel9ER; 9, Cel9EF-Cel9ER; 10, Cel9EF-intronF1; 11, intronR1-Cel9ER; 12, Cel9EF-Cel9ER.