APPLIED AND ENVIRONMENTAL MICROBIOLOGY, Dec. 2007, p. 7997–8000 Vol. 73, No. 24
0099-2240/07/$08.00+0 doi:10.1128/AEM.00956-07

Reconstitution of Glyphosate Resistance from a Split 5-Enolpyruvyl Shikimate-3-Phosphate Synthase Gene in *Escherichia coli* and Transgenic Tobacco▽†

Bao-Qing Dun,[1,2]‡ Xu-Jing Wang,[1]‡ Wei Lu,[1] Zhong-Lin Zhao,[1,3] Song-Na Hou,[1] Bao-Ming Zhang,[2] Gui-Ying Li,[2] Thomas C. Evans, Jr.,[4] Ming-Qun Xu,[4] and Min Lin[1]*

Biotechnology Research Institute, Chinese Academy of Agriculture Sciences, Beijing 100081, China[1]; National Key Facility for Crop Gene Resources and Genetic Improvement, Institute of Crop Sciences, Chinese Academy of Agriculture Sciences, Beijing 100081, China[2]; College of Biological Sciences, China Agricultural University, Beijing 100094, China[3]; and New England Biolabs, Inc., Ipswich, Massachusetts 01938[4]

Received 28 April 2007/Accepted 10 October 2007

A highly *N*-phosphonomethylglycine (glyphosate)-resistant *Pseudomonas fluorescens* G2 5-enolpyruvyl shikimate-3-phosphate synthase (EPSPS) was mapped to identify potential split sites using a transposon-based linker-scanning procedure. Intein-mediated protein complementation was used to reconstitute glyphosate resistance from the genetically divided G2 EPSPS gene in *Escherichia coli* strain ER2799 and transgenic tobacco.

N-Phosphonomethylglycine, commonly referred to as glyphosate, is a popular broad-spectrum, nonselective herbicide used for the control of weeds. Glyphosate kills weeds and crops by inhibiting 5-enolpyruvyl shikimate-3-phosphate synthase (EPSPS) (EC 2.5.1.19), a key enzyme in the shikimate pathway (1, 19, 23). Recently, a new gene encoding a highly glyphosate-resistant EPSPS, which was identified from *Pseudomonas fluorescens* G2 isolated from a storage area with a history of glyphosate pollution (29). The G2 EPSPS gene, when expressed in tobacco, corn, cole, and cotton, provided resistance to glyphosate (our unpublished results).

More than 200 protein splicing elements, termed inteins, have been identified from the genome of eubacteria, eukaryotes, and archaea (12, 17). A naturally occurring *trans*-splicing intein, which consists of a 123-amino-acid N-terminal splicing domain and a separate 36-amino-acid C-terminal splicing domain, was discovered to be in the DnaE gene of *Synechocystis* sp. strain PCC6803 (3, 10, 24, 27, 28). An intein-based split-gene technique, termed intein-mediated protein complementation (IPC), has been utilized to reconstitute the activity of a transgenic product with the goal of limiting the spread of transgenes from genetically modified plants to weedy relatives (6, 7, 13, 16, 18, 26). Theoretically, the split-transgene

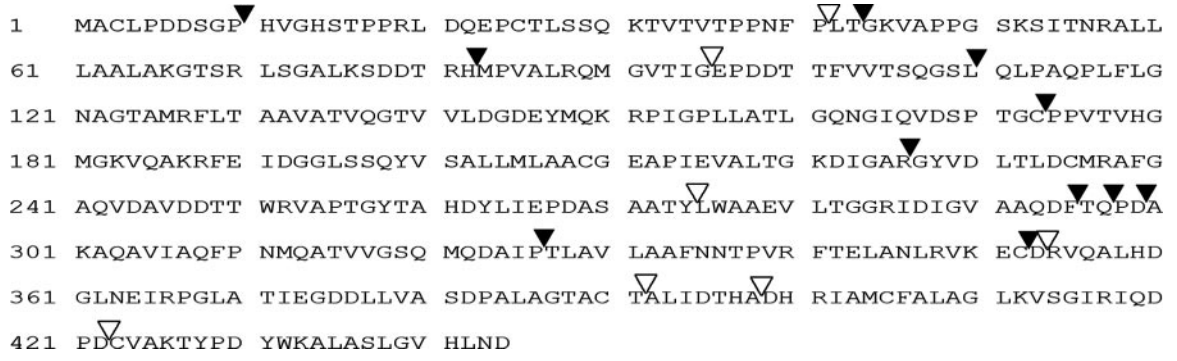

FIG. 1. Pentapeptide insertion into G2 EPSPS. Positions that tolerate 5-amino-acid insertions are indicated by thick arrowheads. Sites that did not tolerate an insertion are indicated by empty arrowheads.

* Corresponding author. Mailing address: Biotechnology Research Institute, Chinese Academy of Agriculture Sciences, Beijing 100081, China. Phone: 86-10-62139578. Fax: 86-10-62136981. E-mail: linmin57@vip.163.com.

† Supplemental material for this article may be found at http://aem.asm.org/.

‡ B.-Q.D. and X.-J.W. contributed to the study equally.

▽ Published ahead of print on 19 October 2007.

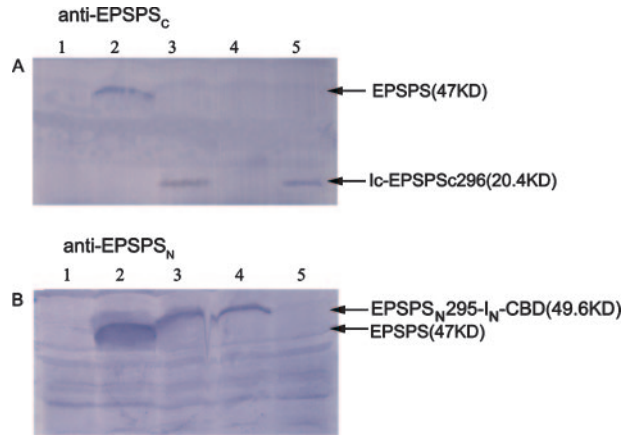

FIG. 2. Western blot analysis of the split EPSPS proteins extracted from *E. coli* ER2799 and its transformants. (A) Western blot analysis using antibodies against the C terminus of EPSPS. (B) Western blot analysis using antibodies against the N terminus of EPSPS. Lanes: 1, ER2799 control; 2, ER2799 containing pMEPS; 3, ER2799 containing pMEPS$_N$295I$_N$ plus pKEPS$_C$296I$_C$; 4, ER2799 containing pMEPS$_N$295I$_N$; 5, ER2799 containing pKEPS$_C$296I$_C$.

methodology is very straightforward, but it is not trivial to find a site to split a protein and subsequently have the truncated protein fragments form an active complex (9, 20, 22, 25).

In this study, a G2 EPSPS linker-scanning library was constructed using a transposon-based linker-scanning method (3, 4). Linker scanning was carried out according to instructions provided by the supplier (New England Biolabs). The plasmids used in this study are listed in Table S1 in the supplemental material. *Escherichia coli* EPSPS gene mutant strain ER2799 did not grow on M9 minimal medium unless complemented with an active EPSPS protein (3, 21) and was utilized in the assay for a functional EPSPS protein. To identify potential split sites within G2 EPSPS, the linker-scanning library was transformed into mutant strain ER2799, and active EPSPS proteins were identified by the ability to restore growth of the mutant cells on minimal medium. Twelve unique sites that allowed growth with 5 amino acid residues inserted were identified, suggesting that these positions were potential sites to divide G2 EPSPS. Seven sites that did not tolerate the 5-amino-acid insertion were found (Fig. 1) (see Table S2 in the supplemental material).

The suitabilities of these 12 positions within the G2 EPSPS protein for IPC were explored using two modified plasmids, pKEB12(E) and pMEB2(B), which contained genes encoding C-terminal and N-terminal splicing domains of strain PCC6803 DnaE intein, respectively (3, 8). The G2 EPSPS gene was divided into two fragments and placed into separate modified plasmids for expression. The plasmids encoding the C-terminal strain PCC6803 DnaE intein splicing domain fused to the C terminus of the gene encoding G2 EPSPS were created by inserting the appropriate G2 EPSPS gene fragments into the EcoRI-to-PstI sites in pKEB12(E). The complementary plasmids that encode the N-terminal intein splicing domain fused to an N-terminal fragment of G2 EPSPS were generated by inserting the appropriate portions of the G2 EPSPS gene into the BamHI-to-XhoI sites in pMEB2(B). The most successful complementary plasmids used G2 EPSPS that was split between amino acid residues F295 and T296, termed pMEPS$_N$295I$_N$ and pKEPS$_C$296I$_C$, respectively (see Fig. S2a and S2b in the supplemental material). Subsequently, the intein splicing domain was deleted from pKEPS$_C$296I$_C$ to generate a control plasmid, termed pKEPS$_C$296I$_C$(−), and a plasmid harboring the complete G2 EPSPS gene was also constructed with pMEB2(B) to yield pMEPS. Cotransformation of *E. coli* ER2799 cells with the two plasmids pMEPS$_N$295I$_N$ and pKEPS$_C$296I$_C$ rescued cell growth on M9 minimal plates supplemented with 50 mM glyphosate and 0.3 mM isopropyl-1-thio-D-galactopyranoside (IPTG). Transformation of the cells with plasmids pMEPS$_N$295I$_N$ and pKEPS$_C$296I$_C$(−), containing both EPSPS protein fragments but lacking the C-terminal 36 amino acids of strain PCC6803 DnaE intein did not permit cell growth in M9 minimal medium (see Fig.S3 in the supplemental material).

Western blot analysis was performed using crude cell extracts and antibodies against the N- or C-terminal peptides of G2 EPSPS. The reaction was visualized with a BCIP (5-bromo-4-chloro-3-indolylphosphate)/nitroblue tetrazolium color development substrate (Promega). Full-length EPSPS, about 47 kDa, indicating *trans*-splicing activity, was not detected when ER2799 was cotransformed with pMEPS$_N$295I$_N$ and pKEPS$_C$296I$_C$, which express residues 1 to 295 of EPSPS fused to the N-terminal splicing domain of strain PCC6803 DnaE intein (EPSPS$_N$295-I$_N$) and the C-terminal domain of strain PCC6803 DnaE intein fused to residues 296 to 445 of EPSPS (I$_C$-EPSPS$_C$296), respectively. But the two unreacted EPSPS$_N$295-I$_N$ and I$_C$-EPSPS$_C$296 fragments were present (Fig. 2). These results indicated that the G2 EPSPS protein split between F295 and T296 reconstitutes enzyme activity by IPC in *E. coli*. In order to further evaluate the effectiveness of IPC in this research system, kinetic characterization of crude proteins extracted from ER2799 containing different plasmids was undertaken using the malachite green dye assay method as previously described (2, 5, 11, 14). The standard reaction was carried out at 28°C in a final

TABLE 1. Kinetic constants of *E. coli*-expressed intact G2 EPSPS and reconstituted G2 EPSPS[a]

Enzyme	Mean sp act (U/mg) ± SD	K_m (PEP) (μM) ± SD	K_i (glyphosate) (μM) ± SD	K_i/K_m (PEP) ± SD	V_{max} (U/mg) ± SD
Intact G2 EPSPS	7.12 ± 0.13	95.20 ± 5.73	49.5 ± 3.17	0.52	7.24 ± 0.16
Reconstituted G2-EPSPS	4.48 ± 0.79	96.80 ± 11.01	35.0 ± 5.78	0.36	4.761 ± 0.92
EPSPS$_N$295I$_N$	—	ND	ND	ND	ND
I$_C$EPSPS$_C$296	—	ND	ND	ND	ND

[a] The results presented are averages of two sets of experiments performed in triplicate. —, the specific enzyme activity is less than 0.1% of that of intact enzyme; ND, not determined.

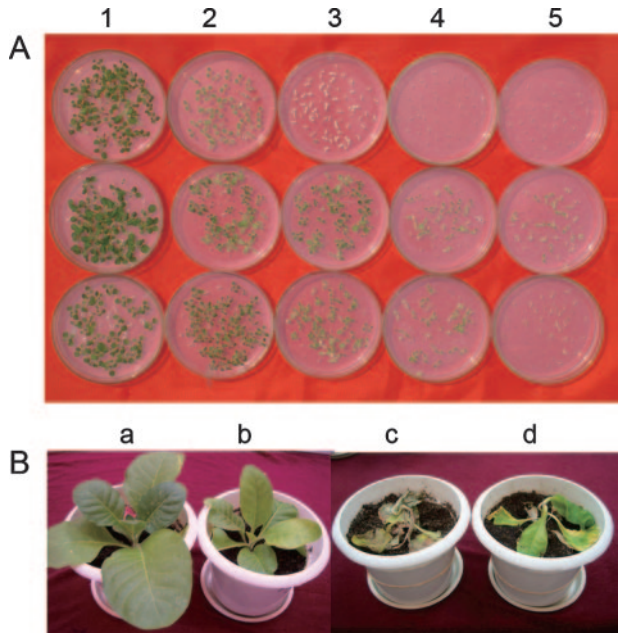

FIG. 3. Glyphosate resistance of transgenic lines versus wild-type tobacco plants. (A) Photograph taken after 28 days of culture on MS_o medium containing 0 to 1 mM glyphosate (columns 1 to 5 contained 0, 0.05 mM, 0.1 mM, 0.5 mM, and 1 mM glyphosate, respectively). The top row shows the seeds of the wild type, the middle row shows the seeds of T_1 generations transformed with the intact G2 EPSPS gene, and the bottom row shows the seeds of T_1 generations transformed with the split G2 EPSPS gene (E_NI_N and I_CE_C). (B) The six- to eight-leaf-stage transgenic plants were sprayed with the herbicide Roundup (isopropylamine salt of glyphosate as the active ingredient [41.0%]) at a dose equal to 1 liter/ha. Injury was observed visually 10 days after the application. (a) Transgenic tobacco containing G2 EPSPS coding sequence; (b) transgenic tobacco containing both I_CE_C and E_NI_N gene fragments; (c) transgenic tobacco containing only the I_CE_C gene fragment; (d) transgenic tobacco containing only the E_NI_N gene fragment.

volume of 100 μl containing 50 mM HEPES (pH 7.0), 1 mM shikimate-3-phosphate, 1 mM phosphoenolpyruvate (PEP), and 5 μl crude extracts. As can be seen in Table 1, there are no significant kinetic differences between the intein-reconstructed G2 EPSPS and the intact G2 EPSPS, suggesting that the EPSPS fragments were brought together by the intein splicing domains to generate a fully active G2 EPSPS (Table 1).

The G2 EPSPS gene was genetically divided into N-terminal and C-terminal inactive fragments from the F295/T296 sites by PCR and then fused to the N terminus and C terminus of strain PCC6803 DnaE intein by overlap extension PCR, namely, $EPSPS_N$-I_N (E_NI_N) and I_C-$EPSPS_C$ (I_CE_C). E_NI_N, I_CE_C, the full-length G2 EPSPS gene were inserted into the plant expression vector pBI121 (Clontech, Palo Alto, CA). The expression vectors pBE_NI_N, pBI_CE_C, and pBEPSPS were introduced into *Agrobacterium* sp. strain LBA4404 (Clontech, Palo Alto, CA) and transformed into *Nicotiana tabacum* var. NC89 (15). Transgenic plants were obtained, and glyphosate resistances of transgenic plants were compared at different growth stages. T_1 generation seeds were germinated on half-strength MS_o medium (21) containing 100 mg/liter of kanamycin and glyphosate with different concentrations from 0 to 1 mM for 4 weeks (Fig. 3A). The six- to eight-leaf-stage transgenic plants were sprayed with a 1% (vol/vol) solution of the herbicide Roundup containing 41.0% glyphosate isopropylamine salt (Monsanto Inc.) at a dose of 1 liter/ha (Fig. 3B). Transgenic plant seedlings from pBEPSPS or both pBE_NI_N and pBI_CE_C exhibited almost equal levels of resistance to glyphosate, suggesting that the strain PCC6803 DnaE intein N- and C-terminal splicing domains can act as in vivo affinity domains and reconstitute the EPSPS activities in plant cells.

Nucleotide sequence accession number. The GenBank accession number of the G2 EPSPS gene sequence is EF155478.

We thank Zhang-Lin Lin for making many useful suggestions, Ming Chen for technical help and useful comments, and Ning Su for reading the manuscript.

This work was supported by the Ministry of Science and Technology of China (National Basic Research Program 2007CB109203 and 2007CB707805 and National High-Tech Program 2007AA021304 and 2006AA020101).

REFERENCES

1. **Boocock, M. R., and J. R. Coggins.** 1983. Kinetics of 5-enopyruvylshikimate-3-phosphate synthase inhibition by glyphosate. FEBS Lett. **154:**127–133.
2. **Bradford, M. M.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72:**248–254.
3. **Chen, L., S. Pradhan, and T. C. Evans.** 2001. Herbicide resistance from a divided EPSPS protein: the split *Synechocystis* DnaE intein as an in vivo affinity domain. Gene **263:**39–48.
4. **Chin, H. G., G. D. Kim, I. Marin, F. Mersha, T. C. Evans, L. X. Chen, M.-Q. Xu, and S. Pradhan.** 2003. Protein *trans*-splicing in transgenic plant chloroplast: reconstruction of herbicide resistance from split genes. Proc. Natl. Acad. Sci. USA **100:**4510–4515.
5. **Copeland, R. A.** 1996. Enzymes: a practical introduction to structure, mechanism, and data analysis. VCH Publishers Inc., New York, NY.
6. **Daniell, H.** 2002. Molecular strategies for gene containment in transgenic crops. Nat. Biotechnol. **20:**581–586.
7. **Evans, T. C., M.-Q. Xu, and S. Pradhan.** 2005. Protein splicing elements and plants: from transgene containment to protein purification. Annu. Rev. Plant Biol. **56:**375–392.
8. **Evans, T. C., D. Martin, R. Kolly, D. Panne, L. Sun, I. Ghosh, L. Chen, J. Benner, X.-Q. Liu, and M.-Q. Xu.** 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.
9. **Hakansson, M., and S. Linse.** 2002. Protein reconstitution and 3D domain swapping. Curr. Protein Pept. Sci. **3:**629–642.
10. **Hayes, F., B. Hallet, and Y. Cao.** 1997. Insertion mutagenesis as a tool in the modification of protein function. J. Biol. Chem. **272:**28833–28836.
11. **He, M., Z. Y. Yang, Y. F. Nie, J. Wang, and P. Xu.** 2001. A new type of class I bacterial 5-enopyruvylshikimate-3-phosphate synthase mutants with enhanced tolerance to glyphosate. Biochim. Biophys. Acta **1568:**1–6.
12. **Henry, P.** 2000. Protein splicing and related forms of protein autoprocessing. Annu. Rev. Biochem. **69:**447–496.
13. **Khan, M. S., A. M. Khalid, and K. A. Malik.** 2005. Intein-mediated protein *trans*-splicing and transgene containment in plastids. Trends Biotechnol. **23:**217–220.
14. **Lanzetta, P. A., L. J. Alvarez, P. S. Reinach, and O. A. Candia.** 1979. An improved assay for nanomole amounts of inorganic phosphate. Anal. Biochem. **100:**95–97.
15. **Limanton-Grevet, A., and M. Jullien.** 2001. *Agrobacterium*-mediated transformation of *Asparagus officinalis* L: molecular and genetic analysis of transgenic plants. Mol. Breeding **7:**141–150.
16. **Muth, N.** 2003. Gene flow might turn wimps into superweeds. Nature **421:** 785–786.
17. **Perler, F. B.** 2002. InBase: the intein database. Nucleic Acids Res. **30:**383–384.
18. **Puricelli, E., and D. Tuesca.** 2005. Weed density and diversity under glyphosate-resistant crop sequences. Crop Prot. **24:**533–542.
19. **Rogers, S. G., L. A. Brand, S. B. Holder, E. S. Sharps, and M. J. Brackin.** 1983. Amplification of the *aroA* gene from *Escherichia coli* results in tolerance to the herbicide glyphosate. Appl. Environ. Microbiol. **46:**37–43.
20. **Rost, B., and C. Sander.** 1993. Prediction of secondary structure at better than 70% accuracy. J. Mol. Biol. **232:**584–599.
21. **Sambrook, J., and D. W. Russell.** 2001. Molecular cloning: a laboratory manual, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

22. **Shiba, K., and P. Schimmel.** 1992. Functional assembly of a randomly cleaved protein. Proc. Natl. Acad. Sci. USA **89:**1880–1884.
23. **Steinrucken, H. C., and N. Amrhein.** 1980. The herbicide glyphosate is a potent inhibitor of 5-enolpyruvylshikimic acid-3-phosphate synthase. Biochem. Biophys. Res. Commun. **94:**1207–1212.
24. **Sun, L., I. Ghosh, H. Paulus, and M.-Q. Xu.** 2001. Protein *trans*-splicing to produce herbicide-resistant acetolactate synthase. Appl. Environ. Microbiol. **67:**1025–1029.
25. **Sun, Y. C., Y. Li, H. Zhang, H. Q. Yan, D. N. Dowling, and Y. P. Wang.** 2006. Reconstitution of the enzyme AroA and its glyphosate tolerance by fragment complementation. FEBS Lett. **580:**1521–1527.
26. **Watrud, L. S., E. H. Lee, A. Fairbrother, C. Burdick, J. R. Reichman, M. Bollman, M. Storm, G. King, and K. V. Peter.** 2004. Evidence for landscape-level, pollen-mediated gene flow from genetically modified creeping bentgrass with CP4 EPSPS as a marker. Proc. Natl. Acad. Sci. USA **101:**14533–14538.
27. **Wu, H., Z. Hu, and X. Q. Liu.** 1998. Protein trans-splicing by a split intein encoded in a split DnaE gene of *Synechocystis* sp. PCC6803. Proc. Natl. Acad. Sci. USA **95:**9226–9231.
28. **Yang, J., G. C. Fox, Jr., and T. V. Henry-Smith.** 2003. Intein-mediated assembly of a functional β-glucuronidase in transgenic plants. Proc. Natl. Acad. Sci. USA **100:**3513–3518.
29. **Zhu, Y., Z. L. Yu, and M. Lin.** 2003. Bioresistance or biodegradation of glyphosate and construction of transgenic plants. Mol. Plant Breeding **1:**435–441.