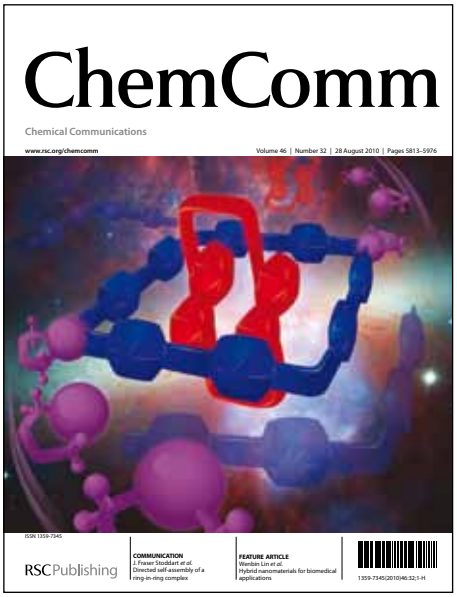
ChemComm

Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the RSC Publishing peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, which is prior to technical editing, formatting and proof reading. This free service from RSC Publishing allows authors to make their results available to the community, in citable form, before publication of the edited article. This *Accepted Manuscript* will be replaced by the edited and formatted *Advance Article* as soon as this is available.

To cite this manuscript please use its permanent Digital Object Identifier (DOI®), which is identical for all formats of publication.

More information about *Accepted Manuscripts* can be found in the **Information for Authors**.

Please note that technical editing may introduce minor changes to the text and/or graphics contained in the manuscript submitted by the author(s) which may alter content, and that the standard **Terms & Conditions** and the **ethical guidelines** that apply to the journal are still applicable. In no event shall the RSC be held responsible for any errors or omissions in these *Accepted Manuscript* manuscripts or any consequences arising from the use of any information contained in them.

Cite this: DOI: 10.1039/c0xx00000x

www.rsc.org/xxxxxx

ARTICLE TYPE

Circular permutation of *E. coli* EPSP synthase: increased inhibitor resistance, improved catalytic activity, and an indicator for protein fragment complementation

Xiongfeng Dai,‡ Manlu Zhu‡ and Yi-Ping Wang*

Received (in XXX, XXX) Xth XXXXXXXXX 20XX, Accepted Xth XXXXXXXXX 20XX
DOI: 10.1039/b000000x

We performed the first circular permutation analysis for *E. coli* 5-enolpyruvylshikimate-3-phosphate synthase, and identified one circular permutant with notably increased resistance to its specific inhibitor and several others with moderately improved catalytic activity. Valid circular permutation sites can be used as effective split sites of protein fragment complementation.

DNA shuffling and error-prone PCR have long been widely applied to enzyme engineering.[1-3] In contrast, circular permutation, which is the rearrangement of a protein primary polypeptide chain, has been largely neglected. Circular permutation involves covalently connecting the native N-terminus and C-terminus of a protein with a linker and cleaving the protein at a different site (called the circular permutation site) to introduce new termini.[4,5] While this termini relocation does not change the amino acid composition of the protein, the primary sequence arrangement can perturb protein local conformation and changed certain properties such as enhanced protein catalytic activity or ligand affinity,[6-10] modified substrate preference,[11,12] changed oligomerization states,[13,14] and improved stability upon proteolytic digestion. [15] Circular permutation has also been used to engineer switchable enzymes. [16, 17]

5-enolpyruvylshikimate-3-phosphate (EPSP) synthase (EPSPS or AroA) is a key metabolic enzyme involved in aromatic amino acid synthesis among bacteria, fungi and higher plants.[18] It converts shikimate-3-phosphate (S3P) and phosphoenolpyruvate (PEP) to EPSP and inorganic phosphate.[18] It is well known as the target of glyphosate, a broad-spectrum herbicide. Glyphosate takes effect by acting as a competitive inhibitor of PEP.[19] Due to the wide application of glyphosate, researchers pay attention to glyphosate-resistant EPSPS and their related transgenic crops. For *E. coli* EPSPS, point mutation has already yielded several glyphosate resistant mutants. However, their substrate affinity was also greatly impaired.[20,21] Thus they provide relatively poor transgenic application value.

In this study, we did the first circular permutation analysis of *E. coli* EPSPS. Seven selected EPSPS circularly permutated (cpEPSPS) variants were subjected to detailed analysis. One variant acquired notable enhancement of glyphosate resistance without sacrificing its substrate affinity and catalytic activity. To our knowledge, this is the first report that circular permutation

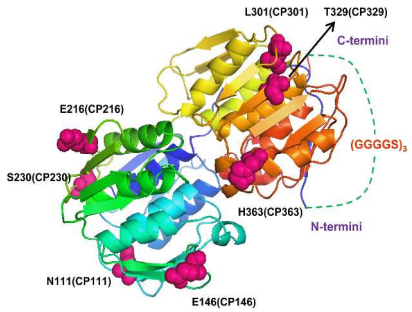


Fig. 1 Seven selected circular permutants of *E. coli* EPSPS. A 15aa $(GGGGS)_3$ flexible linker (green dashed line) was used to bridge the gap of the native termini of EPSPS (PDB ID: 2AA9). The positions of the new N terminus are labeled (magenta).

can enhance the resistance of a given enzyme to its specific inhibitor. Several other variants exhibited moderately improved enzyme catalytic efficiency or substrate affinity. Moreover, we surveyed the relation between circular permutation and protein fragment complementation (PFC), which means the reconstitution of the activity of a given protein from its two separate inactive fragments.[22-24] PFC has already been applied to detect protein-protein interactions[25-30] and modulate the function of biomolecules though small molecules.[31] It has also been proposed that application of PFC to transgenic crops can effectively limit gene flow to other plant species by separately expressing two inactive fragments at the nucleus genome and chloroplast genome.[32-34] Hence, PFC is of great theoretical significance as well as application value. Nevertheless, an efficient approach to identifying split sites of PFC is still lacking. Here we demonstrate for the first time, that all seven selected circular permutation sites can be directly used as valid split sites for leucine zipper-mediated PFC.

Bioinformatics prediction and an *E. coli* complementation assay were combined to identify the viable cpEPSPS variants. First, Cpred[35], a web server, was used to predict the viable circular permutation sites of EPSPS (Fig. S1). An *aroA* tandem fusion template was prepared for PCR amplification of all the selected cpEPSPS variants. In addition, we chose a 15aa flexible linker $(GGGGS)_3$ to bridge the gap of native EPSPS termini, which is about 34 Å from its crystal structure (PDB ID: 2AA9). All the cpEPSPS variants were driven by a constitutive *tet* promoter and transformed into *E. coli* BL21(DE3) *aroA* deleted strain

Cite this: DOI: 10.1039/c0xx00000x

www.rsc.org/xxxxxx

ARTICLE TYPE

Table 1 Kinetic parameters of wild type (WT) EPSPS and its seven selected circular permutants

Enzyme Variant[a]	Circular permutation site[b]	Specific Activity (nkat/mg)	k_{cat} (min^{-1})	K_m(PEP) (μM)	K_m(S3P) (μM)	k_{cat}/K_m (PEP) ($min^{-1}μM^{-1}$)	k_{cat}/K_m(S3P) ($min^{-1}μM^{-1}$)	K_i(glyphosate) (μM)
WT	–	688.5±54.2	(3.03±0.14)×10^3	91±8	137±14	33	21.9	3.75±0.23
CP111	S110-N111	425.9±24.6	(1.85±0.12)×10^3	33±5	36±3	56	51.4	2.3±0.21
CP146	Q145-E146	258.7±30.1	(0.99±0.1)×10^3	31±4	23±4	32	43	0.43±0.02
CP216	I215-E216	852.9±75.1	(3.57±0.22)×10^3	99±11	164±18	36	21.8	3.2±0.41
CP230	Q229-S230	611.5±50.7	(2.62±0.33)×10^3	95±9	107±9	27.6	24.5	8.02±0.57
CP301	E300-L301	419.8± 37.5	(1.6±0.18)×10^3	53±5	28±4	30.2	57.1	2.52±0.35
CP329	T328-T329	389.4± 35.5	(1.41±0.09)×10^3	52±6	32±3	27	44.1	2.89±0.23
CP363	G362-H363	802.1±46.7	(3.27±0.24)×10^3	85±8	91±6	38	35.9	59.1±6.7

[a] EPSPS variants nomenclature: CP111 denotes cpEPSPS variant whose new N terminus starts with amino acid 111 of the wild type enzyme. [b] Circular permutation site refers to the new termini of the circular permutants (e.g. S110-N111 means the new N terminus and C terminus of the CP111 variant is N111 and S110 of the wild type enzyme, respectively).

BD2100[31] to test whether its growth defect in M63 minimal medium could be rescued. Using this strategy, we identified 15 viable cpEPSPS variants (Table S1, Fig. S2), among which seven (CP111, CP146, CP216, CP230, CP301, CP329, and CP363) (Fig. 1) could support BD2100 with a comparable or higher growth rate than the wild type (WT) enzyme (Table S1). Furthermore, BD2100 harboring three variants, CP363, CP230, and CP111, could survive in M63 minimal medium added with 50 mM, 20 mM, 10 mM glyphosate, respectively (Fig. S3, Fig. S4), suggesting that they might have gained elevated glyphosate resistance. The seven cpEPSPS variants mentioned above, along with the WT enzyme were purified for further *in vitro* analysis.

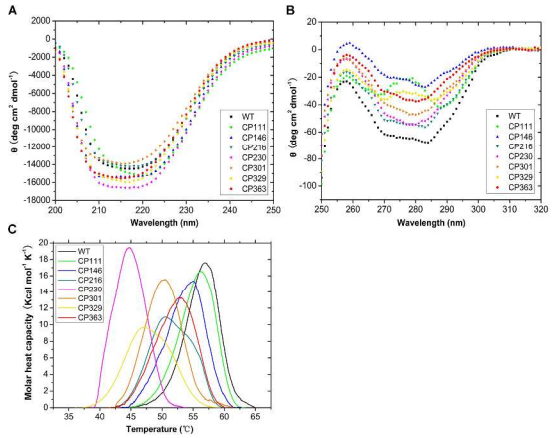


Fig. 2 CD spectra and DSC spectra of WT EPSPS and the selected cpEPSPS variants (A) Far-UV CD spectra (B) Near-UV CD spectra (C) DSC thermograms..

The purified proteins (Fig. S5) were first subjected to *in vitro* kinetic assay (Table 1). CP216 and CP363 exhibited slightly higher catalytic constant (k_{cat}) and specific activity than the WT enzyme. CP111, CP146, CP301, and CP329 showed higher substrate affinity than the WT enzyme, ranging from two- to three-fold for PEP and four- to six-fold for S3P. For catalytic efficiency (k_{cat}/K_m), all seven variants exhibited either equal or higher values than WT enzyme, with CP111 being of the highest efficiency (1.7-fold higher k_{cat}/K_m(PEP) and 2.5-fold higher k_{cat}/K_m(S3P) over WT enzyme). In agreement with *in vivo* result, CP363 had a 17-fold higher K_i(glyphosate) than did the WT enzyme, indicating its significantly enhanced glyphosate resistance. The fact that its specific activity and substrate affinity was similar to the WT enzyme is different from that of glyphosate insensitive mutants generated by point mutations whose activity and substrate affinity were obviously impaired.[20,21]

We further examined the impact of circular permutation on the structure and thermostability of EPSPS. Circular dichroism (CD) spectroscopy in far-UV range (Fig. 2A) and near-UV range (Fig. 2B) revealed certain changes in both secondary and tertiary structures for the seven cpEPSPS variants compared with the WT enzyme. However, strong near-UV CD signals were observed for all seven cpEPSPS variants, suggesting that well-folding tertiary structures were maintained for all of them. Protein thermostability was studied using differential scanning calorimetry (DSC) (Fig. 2C and Table S2). CP111 and CP146 showed T_m similar to the WT enzyme. CP363, the glyphosate-resistant variants had a T_m value of 52.9 °C, which was slightly lower than the WT enzyme. The other cpEPSPS variants appeared less stable, with T_m value dropping by 6 ~ 11 °C relative to the WT enzyme.

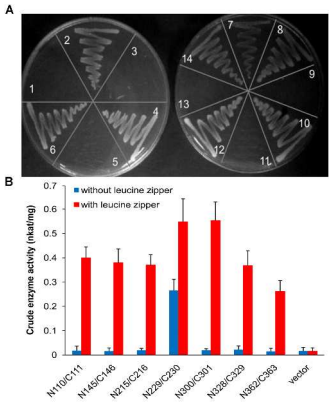


Fig. 3 (A) Survival assay of *E. coli* BD2100 harboring seven EPSPS fragment pairs in M63 minimal medium. For fragment pairs without leucine zippers: (1) N110/C111 (3) N145/C146; (5) N215/C216; (7) N229/C230; (9) N300/C301; (11) N328/C329; (13) N362/C363. For fragment pairs with leucine zippers: (2) N110/C111; (4) N145/C146; (6) N215/C216; (8) N229/C230; (10) N300/C301; (12) N328/C329; (14) N362/C363. EPSPS fragments were named according to their split sites (e.g. N110/C111 denotes the N-terminal fragment from residues 1–110 and the C-terminal fragment from residues 111–427). (B) EPSPS activity in the crude extracts of BD2100 harboring fragment pairs either with or without leucine zippers growing in LB broth.

To contribute to the transgenic risk control of EPSPS and many

other transgenic-related enzymes, an efficient approach to establishing PFC is needed.[36,37] We next tested the feasibility of directly taking circular permutation sites as split sites of PFC. We dissected the WT EPSPS into two fragments at seven selected circular permutation sites (S110-N111, Q145-E146, I215-E216, Q229-S230, E300-L301, T328-T329, and G362-H363) and designated the seven fragment pairs as N110/C111, N145/C146, N215/C216, N229/C230, N300/C301, N328/C329, and N362/C363, respectively. All seven fragment pairs were fused either with or without GCN4 leucine zippers and transformed into BD2100. Surprisingly, BD2100 harboring all seven fragment pairs fused with leucine zippers survived in M63 minimal medium (Fig. 3A, Table S3), suggesting the successful reconstitution of EPSPS activity *in vivo* for all fragment pairs. For fragment pairs without leucine zippers, only one pair, N229/C230, was able to self-complement to enable BD2100 to grow in M63 minimal medium (Fig. 3A, Table S3). For the other six pairs, even when either the N- or C-terminal fragment was fused with one leucine zipper, they still failed to enable BD2100 to survive in M63 minimal medium (Fig. S6), indicating that fragment reassembly of these six pairs required assistance from leucine zipper dimerization. We also measured EPSPS activity in the crude extracts of BD2100 growing in LB. Cells transformed with all seven pairs with leucine zippers showed obvious EPSPS activity, while cells containing all combinations without leucine zippers (except for N229/C230) exhibited activities that were not significantly different from vector-transformed cell (Fig. 3B). Enzyme activity in the crude extracts of BD2100 harboring the seven leucine zipper-fused fragment pairs was much lower compared with that of purified WT EPSPS (Fig. 3B, Table 1), which could be due to their relatively low expression levels driven by the *tet* promoter. When they were expressed by the T7 promoter, crude enzyme activities for all seven leucine zipper-fused fragment pairs were comparable to that of the WT enzyme (Fig. S7). Our discovery suggests that circular permutation might be closely related to leucine zipper-mediated PFC, at least for EPSPS. The underlying mechanism might be that: the former is essentially the reconstitution of two fragments *in cis*, while the latter promotes fragment association *in trans* by bringing fragments into proximity through non-covalent dimerization.[25-28] Thus they both facilitate enzyme structure refolding by increasing the local concentration of the two fragments. Overall, the rapid identification of active circular permutation sites provides an efficient approach to developing PFC for a given protein. This saves great deal of efforts as previous searches for valid split sites mainly relied on library screening or empirical design.[28,32,33,38,39] Leucine zipper-mediated PFC can even be applied to proteins that are not amenable to self-complementation[38,39], and it can be used to detect protein-protein interaction as well.[25,39]

In summary, circular permutation has the potential to increase the resistance of a given enzyme to its specific inhibitor. Moreover, viable circular permutation sites could also be used as valid split sites of protein fragment complementation.

We thank Prof. Terrence Hwa (Department of Physics, UCSD, US) for critically reading the manuscript. This work was supported by the National Transgenic Research and Industrialization Special Foundation of China (2009ZX08009057B and 2008ZX08009-003-002).

Notes and references

State Key Laboratory of Protein and Plant Gene Research, School of Life Sciences, Peking University, Beijing, 100871, China.

E-mail: wangyp@pku.edu.cn; Fax: +86 10 6275 6325; Phone: +86 10 6275 8490

† Electronic Supplementary Information (ESI) available: [Exprimental details, supplementary figures and tables.]. See DOI: 10.1039/b000000x/

‡ These two authors contributed equally to this work.

1 O. Kuchner, F. H. Arnold, *Trends Biotechnol*, 1997, **15**, 523–530.
2 A. Crameri, S. A. Raillard, E. Bermudez, W. P. C. Stemmer, *Nature*, 1998, **391**, 288–291.
3 F. H. Arnold, A. A. Volkov, *Curr. Opin. Chem. Biol.*, 1999, **3**, 54–59.
4 Y. Yu, S. Lutz, *Trends Biotechnol.*, 2011, **29**, 18–25.
5 S. Bliven, A. Prlić, *PLoS Comput Biol,* 2012, **8**, e1002445.
6 A. L. Fishburn, J. R. Keeffe, A. V. Lissounov, D. H. Peyton, S. J. Anthony-Cahill, *Biochemistry*, 2002, **41**, 13318–13327.
7 Z. Qian and S. Lutz, *J. Am. Chem. Soc*., 2005, **127**, 13466–13467.
8 S. Reitinger, Y. Yu, J. Wicki, M. Ludwiczek, I. D'Angelo, S. Baturin, M. Okon, N. C. Strynadka, S. Lutz, S. G. Withers, L. P. McIntosh, *Biochemistry*, 2010, **49**, 2464–2474.
9 A. V. Cheltsov, M. J. Barber, G. C. Ferreira, *J. Biol. Chem*., 2001, **276**, 19141–19149.
10 P. Stephen, K. L. Tseng, Y. N. Liu, P. C. Lyu, *Chem. Commun.*, 2012, **48**, 2612–2614.
11 Y. Yu, S. Lutz, *Biotechnol. Bioeng*., 2010, **105**, 44–50.
12 G. Guntas, M. Kanwar, M. Ostermeier, *PloS ONE*, 2012, **7**, e35998.
13 T. Mizobata, T. Uemura, K. Isaji, T. Hirayama, K. Hongo, Y. Kawata, *PloS ONE*, 2011, **6**, e26462.
14 Z. Qian, J. R. Horton, X. Cheng, S. Lutz, *J. Mol. Biol*. 2009, **393**, 191–201.
15 T. A. Whitehead, L. M. Bergeron, D. S. Clark, *Protein Eng. Des. Sel.* 2009, **22**, 607– 613.
16 G. Guntas, S. F. Mitchell, M. Ostermeier, *Chem. Biol.,* 2004, **11**, 1483–1487.
17 G. Guntas, T. J. Mansell, J. R. Kim, M. Ostermeier, *Proc. Natl. Acad. Sci. U.S.A*, 2005, **102**, 11224–11229.
18 R. Bentley, *Crit. Rev. Biochem. Mol. Biol.*, 1990, **25**, 307–384.
19 E. Schönbrunn, S. Eschenburg, W. A. Shuttleworth, J. V. Schloss, N. Amrhein, J. N. Evans, W. Kabsch, *Proc. Natl. Acad. Sci. U. S. A*., 2001, **98**, 13761380.
20 G. M. Kishore, D. M. Shah, *Annu. Rev. Biochem.*, 1988, 57, 627–663.
21 S. Eschenburg, M. L. Healy, M. A. Priestman, G. H. Lushington, E. Schönbrunn, *Planta*, 2002, **216**, 129–135.
22 H Taniuchi, CB Anfinsen, A Sodja, *Proc. Natl. Acad. Sci. U. S. A*., 1967, **58**, 1235–1242.
23 E. Bibi, H. R. Kaback, *Proc. Natl. Acad. Sci. U. S. A*., 1990, **87**, 4325–4329.
24 K Shiba, P Schimmel, *Proc. Natl. Acad. Sci. U. S. A*., 1992, **89**, 1880-1884.
25 J. N. Pelletier, F. X. Campbell-Valois and S. W. Michnick, *Proc. Natl. Acad. Sci. U. S. A*., 1998, **95**, 12141–12146.
26 I. Remy, S. W. Michnick , *Nat. Methods.*, 2006, **3**, 977–979.
27 S. W. Michnick, P. H. Ear, E. N. Manderson, I. Remy, E.Stefan, *Nat. Rev. Drug. Discov*., 2007, **6**, 569–582.
28 T. K. Kerppola, *Chem. Soc. Rev.*, 2009, **38**, 2876–2886.
29 C. D. Hu, T. K. Kerppola, *Nat. Biotechnol.*, 2003, **21**, 539–545.
30 P. Q. Nguyen, J. J. Silberg, *Protein Eng Des Sel*, 2010, **23**, 529–536.
31 H. D. Mootz, T. W. Muir, *J. Am. Chem. Soc*., 2002, **124**, 9044–9045.
32 L. X. Chen, S. Pradhan, TC. Evans. Jr, *Gene*, 2001, 263, 39–48.
33 Y. C. Sun, Y. Li, H. Zhang, H. Q. Yan, D. N. Dowling and Y. P. Wang, *FEBS Lett*., 2006, **580**, 1521–1527.
34 H. G. Chin, G. D. Kim, I. Marin, F. Mersha, T. C. Evans, Jr., L. X. Chen, M. Q. Xu, S. Pradhan, *Proc. Natl. Acad. Sci. U. S. A*., 2003, **100**, 4510–4515.
35 W. C. Lo, L. F. Wang, Y. Y. Liu, T. Dai, J. K. Hwang, P. C. Lyu, *Nucleic Acids Res*, 2012, **40**, W232–7.
36 D. G. Burgess, E. J. Ralston, W. G. Hanson, M. Heckert, M. Ho, T. Jenq, J. M. Palys, K. Tang, N. Gutterson, *Plant J*, 2002, **31**, 113–125.
37 Y. Gleba, S. Marillonnet, V. Klimyuk, *Biotechnol. Genet. Eng. Rev.*, 2004, **21**, 325–367
38 M. Ostermeier, A. E. Nixon, J. H. Shim, S. J. Benkovic, *Proc. Natl. Acad. Sci. U. S. A*., 1999, **96**, 3562–3567.
39 D. E. Paschon, Z. S. Patel, M. Ostermeier, *J. Mol. Biol.*, 2005, **353**, 26–37.