Supplementary information

Photosynthetic production of ethanol from carbon dioxide in genetically engineered cyanobacteria

Zhengxu Gao[a,b,#], Hui Zhao[a,#], Zhimin Li[a], Xiaoming Tan[a] and Xuefeng Lu*[a]

[1]Key Laboratory of Biofuels, Shandong Provincial Key Laboratory of Energy Genetics, Qingdao Institute of Bioenergy and Bioprocess Technology, Chinese Academy of Sciences, No. 189 Songling Road, Qingdao 266101, China

[2]Graduate School of Chinese Academy of Sciences, Beijing 100049, China

Contents

Table S1 Stains and plasmid used and constructed in this study

Strain	**Relevant genotype**	**Reference**
PCC6803	*Synechocystis* sp.PCC6803 wild type	[1]
PCC7120	*Anabaena* sp. PCC 7120 wild type	[1]
PCC7942	*Synechococcus* sp.PCC7942 wild type	[1]
Syn-LY2	*slr0168*::Omega	[1]
Syn-XT43	*slr0168*::Omega P_{rbc} *pdc* and *adh II*	This study
Syn-ZG25	slr0168:: Omega P_{rbc} *pdc* and *slr1192*	This study
Syn-HZ23	*slr9394*::Kan P_{rbc} *pdc* and *slr1192*	This study
Syn-HZ24	*slr9394*::Kan P_{rbc} *pdc* and *slr1192* slr0168:: Omega P_{rbc} *pdc* and *slr1192*	This study
Plasmid	**Relevant properties**	
pMD18-T	AP^r; cloning vector	Takara
pET28-b	Km^r; containing 6×His tag	Novagen
pFQ20	Ap^r Sp^r; slr0168 targeting vector; Omega	[1]
pMSD15	Cm^r; pACYA184 derivative	[2,3]
pXT5	Km^r; P_{T7}::*adh II*	This study
pXT113A	Km^r; P_{T7}::*slr1192*	This study
pZG35	Km^r; P_{T7}:: Synpcc7942_0459	This study
pZG36	Km^r; P_{T7}::*all0879*	This study
pZG37	Km^r; P_{T7}::*alr0895*	This study
pZG38	Km^r; P_{T7}::*alr0897*	This study
pZG39	Km^r; P_{T7}::*slr0942*	This study
pZG40	Km^r; P_{T7}::*sll0990*	This study
pZG41	Km^r; P_{T7}::*all2810*	This study
pZG42	Km^r; P_{T7}::*all5334*	This study
pZG62	Cm^r Sp^r; P_{T7}::*pdc*	This study
pZG63	Km^r; P_{T7}:: *slr1192*, *alr0895*	This study
pZG64	Km^r; P_{T7}:: *adh II*, *slr1192*	This study
pZG65	Km^r; P_{T7}:: *alr0895*, *adh II*	This study
pZG66	Km^r; P_{T7}:: *alr0895*, *adh II*, *slr1192*	This study
pZG25	Ap^r Sp^r; *slr0168* targeting; P_{rbc}::*pdc*, *slr1192*	This study
pHZ22	Ap^r Km^r; containing *slr9394* homologous arm	This study
pHZ23	Ap^r Km^r; *slr9394* targeting; P_{rbc}::*pdc*, *slr1192*	This study

Table S2 Oligonucleotide primers used in this study

Primers	**Sequence (5'⟶3')**
pdcF	GGCATATGAGTTATACTGTCGGTACCTATTTAGCGG
pdcR	GGTACTAGTCTAGAGGAGCTTGTTAACAGGCT
1192F	TACATATGATTAAAGCCTACGCTGCCCT
1192R	TTTCTCGAGCTAATTTTTACTATGGCTGAGCACTAC
93F	GGGGACCATCCTGACTACACGG
93R	TTCCAATGCCATGGGTTGGGAT
94F	GTAGCCATGGTAGCTATGTCACCG
94R	TGTTGATGGTGGGTATCGTGGTG
rbcL-T	TAATAACTGTCTCTGGGGCGACGG
0168-1	ACCTCTCCACGCTGAATTAG
0168-2	TTCCAGGCCACATTGTTGTC
dc92F	CTGGATGAACCCTTTACCCGCTTA
dc92R	TTTCTCGAGCTAATTTTTACTATGGCTGAGCACTAC
adhF	GGCATATGGCTTCTTCAACTTTTTATATTCCTTTCGTC
adhR	GCGCTCGAG TTAGAAAGCGCTCAGGA
0459F	TACATATGAAATCACGCGCTGCGAT
0459R	TTAAAAGGTGATGACGGTGCGAAT
0879F	GACATATGCGCGCCATGATTTTAGA
0879R	ATCTCGAGTCACTTCATCACTAAAACG
0895F	AGCATATGAAAGCAGTCTGCTGG
0895R	ATCTCGAGTTACGGTTTGAGTACAACT
0897F	TACATATGAAAGCAGTTTGCTGGC
0897R	ATCTCGAGTTAGGGTTTGAGTACAAC
0942F	GCCATATGCAGAGTTTCAATAGG
0942R	AGCTCGAGTTAAATTTCATCCCATAGG
0990F	TACATATGAAATCCCGTGCCGCC
0990R	ATCTCGAGTTAGTAGTGGATCACACT
2810F	CACATATGGAAGTGAAAGCAGCAA
2810R	CTCTCGAGTTAAAAAGTCACCACACT

5334F	GCCATATGAAAGCAGTTGTTTTT
5334R	ACCTCGAGCTAAACATTAGCTAAAGC

Table S3 Nine alcohol dehydrogenases from three different cyanobacteria

Plasmid	**Gene name**	**Cyanobacteria type**
pXT113A	*slr1192*	*Synechocystis* sp. PCC6803
pZG35	Synpcc7942_0459	*Synechococcus* sp. PCC7942
pZG36	*all0879*	*Anabaena* sp. PCC7120
pZG37	*alr0895*	*Anabaena* sp. PCC7120
pZG38	*alr0897*	*Anabaena* sp. PCC7120
pZG39	*slr0942*	*Synechocystis* sp. PCC6803
pZG40	*sll0990*	*Synechocystis* sp. PCC6803
pZG41	*all2810*	*Anabaena* sp. PCC7120
pZG42	*all5334*	*Anabaena* sp. PCC7120

Table S4 The doubling time of *E. coli* strains with expression of pyruvate decarboxylase and different alcohol dehydrogenases

Strain	Genotype	Doubling time (g/h)	Final OD_{600}
E. coli (pZG62)	*pdc*	3.54±0.38	0.94
E. coli (pZG62+pXT5)	*pdc+adhII*	2.91±0.27	2.02
E. coli (pZG62+pXT113A)	*pdc+slr1192*	2.91±0.30	2.67
E. coli (pZG62+pZG35)	*pdc*+Synpcc7942_0459	3.24±0.08	2.22
E. coli (pZG62+ pZG36)	*pdc+all0879*	3.05±0.35	2.53
E. coli (pZG62+ pZG38)	*pdc+alr0897*	2.54±0.37	2.43
E .coli (pZG62+ pZG42)	*pdc+all5334*	2.68±0.14	2.49
E. coli (pZG62+ pZG37)	*pdc+alr0895*	3.16±0.22	2.65
E. coli (pZG62+ pZG39)	*pdc+slr0942*	2.82±0.29	3.18
E. coli (pZG62+ pZG40)	*pdc+sll0990*	3.41±0.16	2.27
E. coli (pZG62+ pZG41)	*pdc+all2810*	3.08±0.39	2.43
E. coli (pZG62+ pZG63)	*pdc+alr0895+slr1192*	2.65±0.30	2.53
E. coli (pZG62+ pZG64)	*pdc+adhII+slr1192*	3.24±0.27	2.66
E. coli (pZG62+ pZG65)	*pdc+adhII+alr0895*	2.64±0.21	2.25
E. coli (pZG62+ pZG66)	*pdc+adhII+alr0895+slr1192*	2.84±0.34	2.00

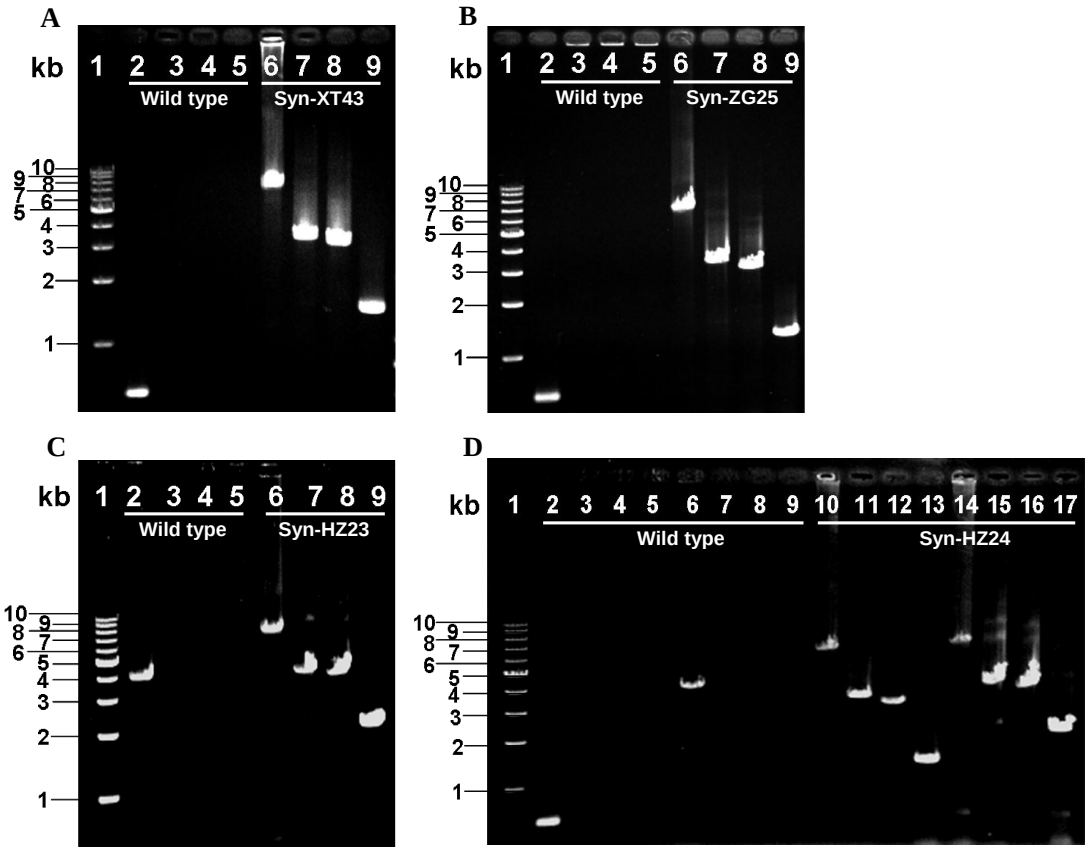

Figure S1 PCR Identification of the genotype of four ethanol producing strains. The primers used in this section were listed in Table S2.

A) Genotype analysis of Syn-XT43. Lane1: DNA Ladder. Lane 2-5: wild type DNA was used as the template in PCR. Lane 6-9: the genome DNA of Syn-XT43 was used as the template. Primers 0168-1 and 0168-2, which are specific to the N- and C-terminal of *slr0168* locus respectively, were used in lane 2 and 6. This result indicated the expression cassette was inserted into the *slr0168* site of the genome with an approximate length of 7,500 bp, and the mutant Syn-XT43 was completely segregated. Primers rbcL-T (specific to P_{rbc} promoter) and 0168-2 (specific to the *slr0168* C-terminal) were used in lane 3 and 7. Primers pdcF(specific to the N-terminal of *pdc* gene) and 0168-2 (specific to the *slr0168* C-terminal) were used in lane 4 and 8. Primers adhF (specific to the N-terminal of *adhII* gene) and 0168-2 (specific to the *slr0168* C-terminal) were used in lane 5 and 9. These results indicated the *pdc* and *adh II* expression cassette was successfully inserted into the *slr0168* site of the genome in the order of P_{rbc}, *pdc* and *adhII*.

B) Genotype analysis of Syn-ZG25. Lane1: DNA Ladder. Lane 2-5: wild type DNA was used as the template. Lane 6-9: the genome DNA of Syn-ZG25 was used as the template. Primers 0168-1 and 0168-2 were used in lane 2 and 6. This result indicated the expression cassette was inserted into the *slr0168* site of the genome with an approximate length of 7,300 bp, and the mutant Syn-ZG25 was completely segregated. Primers rbcL-T and 0168-2 were used in lane 3 and 7. Primers pdcF and 0168-2 were used in lane 4 and 8. Primers 1192F (specific to the N-terminal of *slr1192* gene) and 0168-2 were used in lane 5 and 9. These PCR results indicated the *pdc* and *slr1192* expression cassette was successfully inserted into the *slr0168* site of the genome in the order of P_{rbc}, *pdc* and *slr1192*.

C) Genotype analysis of Syn-HZ23. Lane1: DNA Ladder. Lane 2-5: wild type DNA was used as the template. Lane 6-9: the genome DNA of Syn-HZ23 was used as the template. Primers 93F and 94R, which are specific to the N- and C-terminal of *slr9394* (or *phaAB*) locus respectively, were used in lane 2 and 6. This result indicated the expression cassette was completely inserted into the *phaAB* site of the genome with an approximate length of 8,000 bp, and the mutant Syn-HZ23 was completely segregated. Primers rbcL-T and 94R were used in lane 3 and 7. Primers pdcF and 94R were used in lane 4 and 8. Primers 1192F and 94R were used in lane 5 and 9. These PCR results indicated the *pdc* and *slr1192* expression cassette was successfully inserted into the *phaAB* site of the genome in the order of P_{rbc}, *pdc* and *slr1192*.

D) Genotype analysis of Syn-HZ24. Lane1: DNA Ladder. Lane 2-9: wild type DNA was used as the template. Lane 10-17: the genome DNA of Syn-HZ24 was used as the template. Primers 0168-1 and 0168-2 were used in lane 2 and 10. Primers rbcL-T and 0168-2 were used in lane 3 and 11. Primers pdcF and 0168-2 were used in lane 4 and 12. Primers 1192F and 0168-2 were used in lane 5 and 13. These PCR result indicated the *pdc* and *slr1192* expression cassette was completely inserted into the *slr0168* site of the genome in the order of P_{rbc}, *pdc* and *slr1192*. Primers 93F and 94R were used in lane 6 and 14. Primers rbcL-T and 94R were used in lane 7 and 15. Primers pdcF and 94R were used in lane 8 and 16.

Primers 1192F and 94R were used in lane 9 and 17. And these PCR result indicated the *pdc* and *slr1192* expression cassette was completely inserted into the *phaAB* site of the genome in the order of P_{rbc}, *pdc* and *slr1192*.

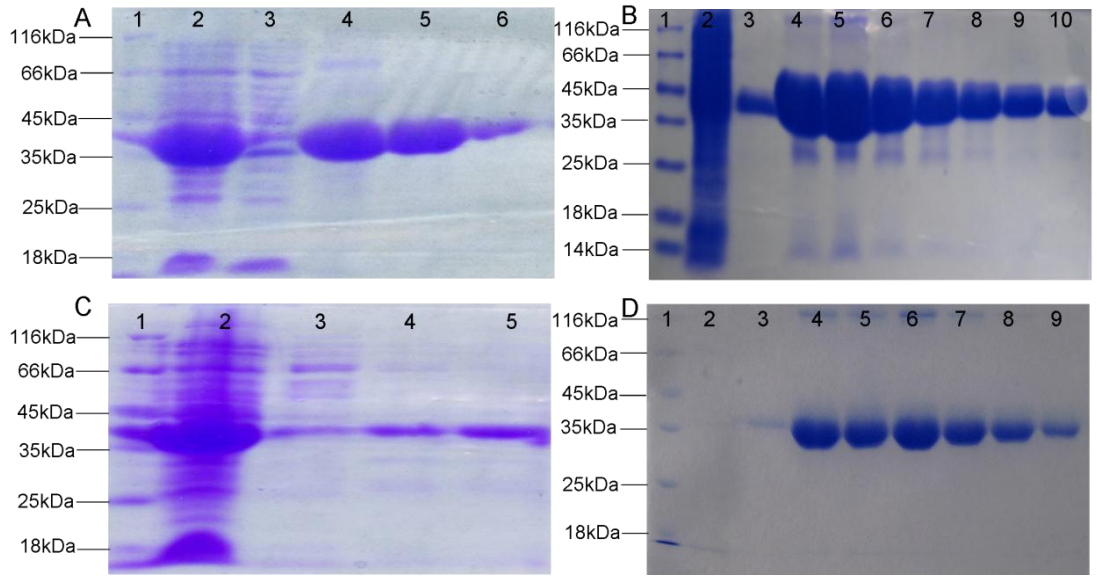

Figure S2 The SDS-PAGE gel of *adh Ⅱ*, *slr1192*, *all0879* and Synpcc7942_0459 purification. (A) The gel of *adh Ⅱ*, lane 1: protein marker, lane 2: crude extract, lane 3: the protein through the Ni-NTA resin, lane 4-6: the target protein eluted by Tris buffer containing 100mM iminazole (B) The gel of *slr1192*, lane 1: protein marker, lane 2: crude extract, lane 3-10: the protein eluted by Tris buffer containing 250 mM iminazole (C) The gel of *all0879*, lane1: protein marker, lane 2: crude extract, lane 4 and 5: the target protein eluted by Tris buffer containing 100 mM iminazole (D) The gel of Synpcc7942_0459, lane 1: protein marker, lane 2-9: the target protein eluted by Tris buffer containing 250mM iminazole.

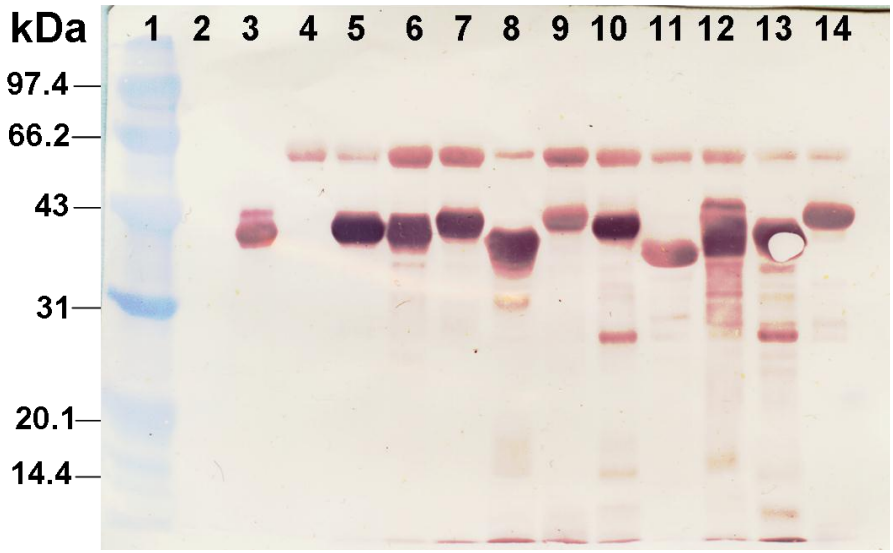

Figure S3 Western blot analysis of protein expression in *E. coli* mutants with heterologous overexpression of pyruvate decarboxylase from *Zymomonas mobilis* and different alcohol dehydrogenases.

Lane 1: protein marker. Lane 2: the total proteins in wild type *E.coli* as a negative control. Lane 3: the purified alcohol dehydrogenase (*slr1192*) from *Synechocystis* sp. PCC6803 as a positive control. Lane 4: the purified pyruvate decarboxylase (*pdc*) from *Zymomonas mobilis* as another positive control. Lane 5 through lane 14 showed the results of the total protein extracts from *E. coli* mutant harboring *pdc* and *adh II* genes (lane 5); *pdc* and *slr1192* genes (lane 6); *pdc* and Synpcc7942_0459 genes (lane 7); *pdc* and *all0879* genes (lane 8); *pdc* and *alr0895* genes (lane 9); *pdc* and *alr0897* genes (lane 10); *pdc* and *slr0942* genes (lane 11); *pdc* and *sll0990* genes (lane 12); *pdc* and *all2810* genes (lane 13); *pdc* and *all5334* genes (lane 14), respectively.

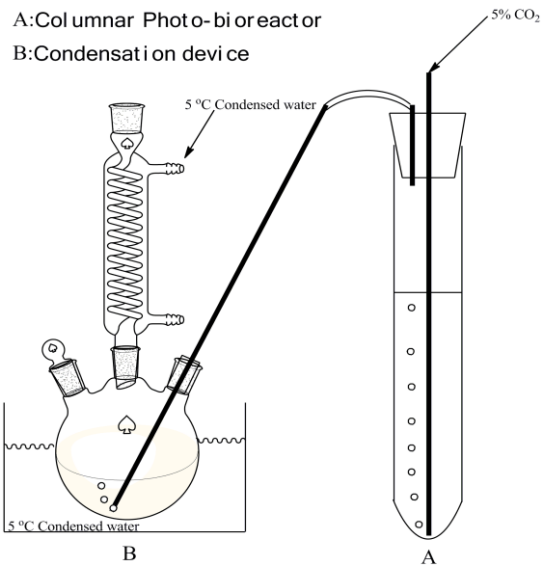

Figure S4 The column photo-bioreactor (A) and condensation device (B) used for exact determination of the ethanol production in cyanobacteria.

References

1. X. Tan, L. Yao, Q. Gao, W. Wang, F. Qi and X. Lu, *Metab Eng*, 2011, **13**, 169-176.
2. Y. Duan, Z. Zhu, K. Cai, X. Tan and X. Lu, *PLoS One*, 2011, **6**, e20265.
3. X. Lu, H. Vora and C. Khosla, *Metab Eng*, 2008, **10**, 333-339.