Supplementary Information

Glycine betaine monooxygenase, an unusual Rieske-type oxygenase system, catalyzes the oxidative *N*-demethylation of glycine betaine in *Chromohalobacter salexigens* DSM 3043

Ya-Hui Shao[a]†, Li-Zhong Guo[a]†, Yu-Qing Zhang[a], Hao Yu[a], Bai-Suo Zhao[b], Hai-Qiang Pang[c], Wei-Dong Lu[a]

Shandong Provincial Key Laboratory of Applied Mycology, College of Life Sciences, Qingdao Agricultural University, Qingdao, 266109, PR China [a]; Graduate School, Chinese Academy of Agricultural Sciences, Beijing, 100081, PR China [b]; Shandong Ruiying Pioneer Pharmaceutical Co., Ltd., HeZe, 274000, PR China [c].

†Contributed equally to this work
Author for corresponding: Wei-Dong Lu, luweidong401@hotmail.com
Tel.: +86-532-86080640

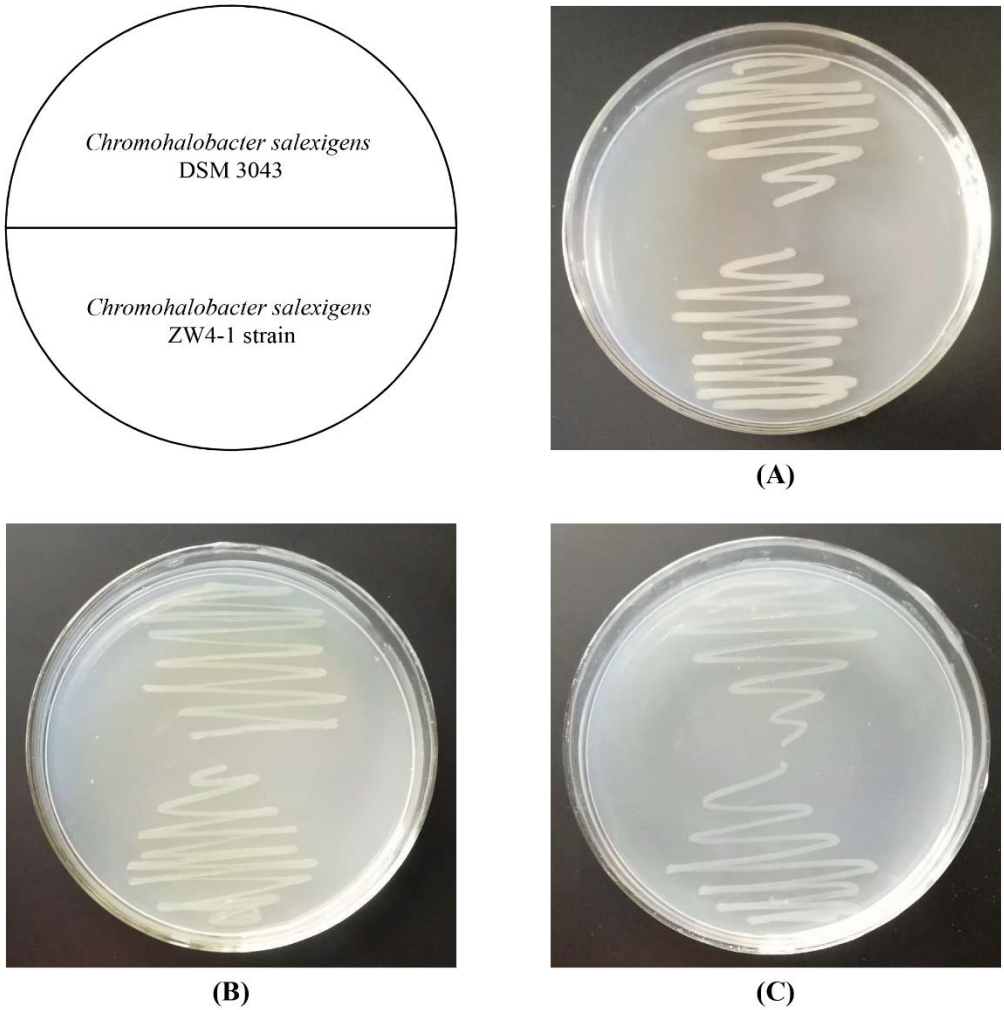

FIG S1 Growth of *C. salexigens* DSM 3043 and ZW4-1 strains on C-M63 agar plates (**A**) as described in MATERIALS AND METHODS, and on the nitrogen-free C-M63 agar plates supplemented with glycine betaine (20 mM) as the sole nitrogen source (**B**), and on the carbon- and nitrogen-free C-M63 agar plates supplemented with glycine betaine (20 mM) as the sole carbon and nitrogen source (**C**). The cultures were incubated at 37°C for 72 h.

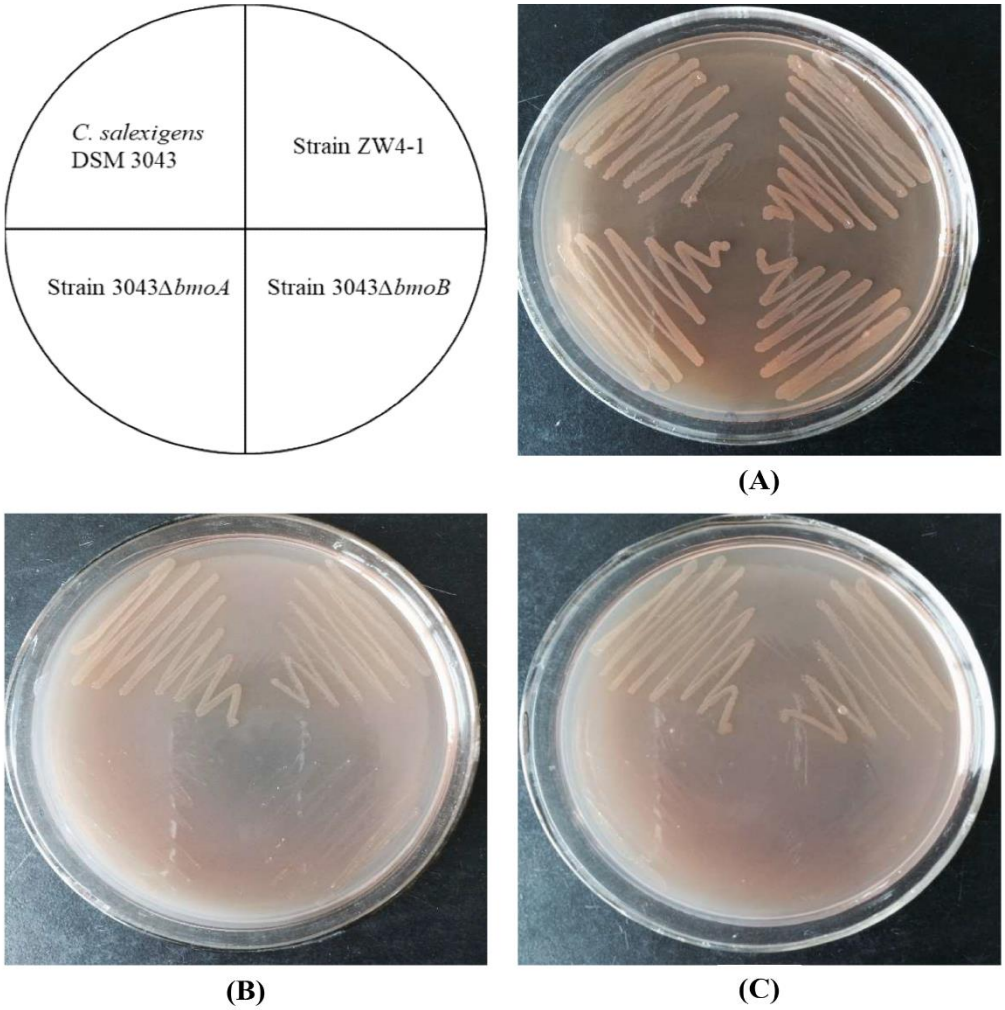

FIG S2 Carbon utilization tests. *C. salexigens* DSM 3043, strain ZW4-1, and single-gene deletion mutants (3043Δ*bmoA* and 3043Δ*bmoB* strains) were streaked on C-M63 agar plates with glucose (20 mM) (**A**), glycine betaine (20 mM) (**B**) or dimethylglycine (20 mM) (**C**) as the sole carbon source. The cultures were incubated at 37°C for 72 h.

Note: rifampicin (50 μg ml^{-1}) was added to C-M63 agar plates.

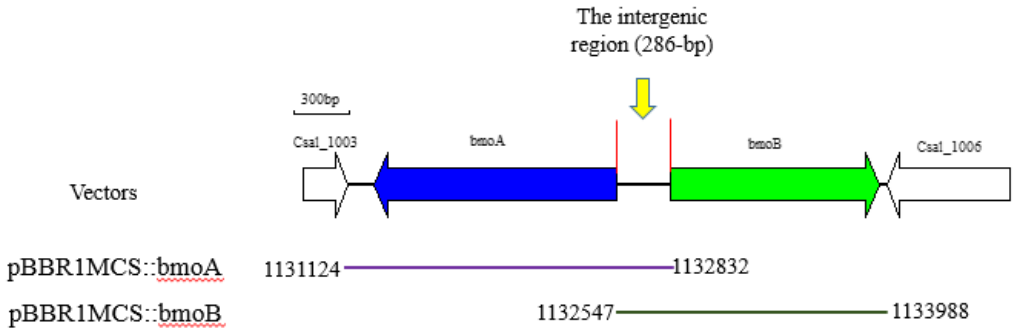

FIG S3 Schematic diagram for the complementation plasmids and the 286-bp intergenic region. The arrows indicate the length and transcription orientation of the ORFs. The DNA fragment inserted in the complementation plasmids are indicated by the lines. Numbers indicate the locations in the genome sequence of *C. salexigens* DSM 3043 (GenBank accession number: NC_007963.1).

```
286-OR.seq    ..GAAGGCATCCTCTTCGTGCGCCGGTTCGATAACGGATGGCGCCAATGACATTGAATGG     58
286-RC.seq    GGAAAGTCTACCTGGTCGCGCGACGGT.....ATCCGACGTTTCAGCGCATTGTGATCAT     55

286-OR.seq    GGTATGCGTGAGAGGCGCGAGTGACAAGGAGTGTGGCGTATCGCCACTCAAGCGAATTGT    118
286-RC.seq    GCGGCTCAGTGGCCACTTGCCTGCCAACGACTCACGCATGCCAGC.CCCGGCCACGCCTT    114

286-OR.seq    CTGCTCGC.GACACCGGCCGATGCAGTGACGACGCCCATCGAAGACAGCGACTGCAAGGC    177
286-RC.seq    GCAGTCGCTGTCTTCGATGGGCGTCGTCACTGCATCGGCCGGTGTC.GCGAGCAGACAAT    173

286-OR.seq    GTGGCCGGG.GCTGGCATGCGTGAGTCGTTGGCAGGCAAGTGGCCACTGAGCCGCATGAT    236
286-RC.seq    TCGCTTGAGTGGCGATACGCCACACTCCTTGTCACTCGGCCCTCTCACGCATACCCCATT    233

                  +1                                  RBS
286-OR.seq    CACAATGCGCTGAAACGTCGGATA.....CCGTCGGCGCACCAGGTAGACTTTC         285
286-RC.seq    CAATGTCATTGGCGCCATCCGTTATCGAACCGGCGCACGAAGAGGATGCCTTC.         286
```

FIG S4 Sequence alignment of the 286-bp intergenic region (286-OR) between the *bmoA* and *bmoB* coding regions with its reverse complement sequence (286-RC). The identical amino acids residues in the two sequences are shown in black. With the promoter prediction software NNPP version 2.2 (http://www.fruitfly.org/seq_tools/promoter.html), a possible promoter region of *bmoB* (Csal_1005) gene was detected, which was indicated by a red horizontal line, and the predicted transcription start was shown in +1. The potential ribosome-binding sites (RBS) for the *bmoA* and *bmoB* genes were marked with a purple box.

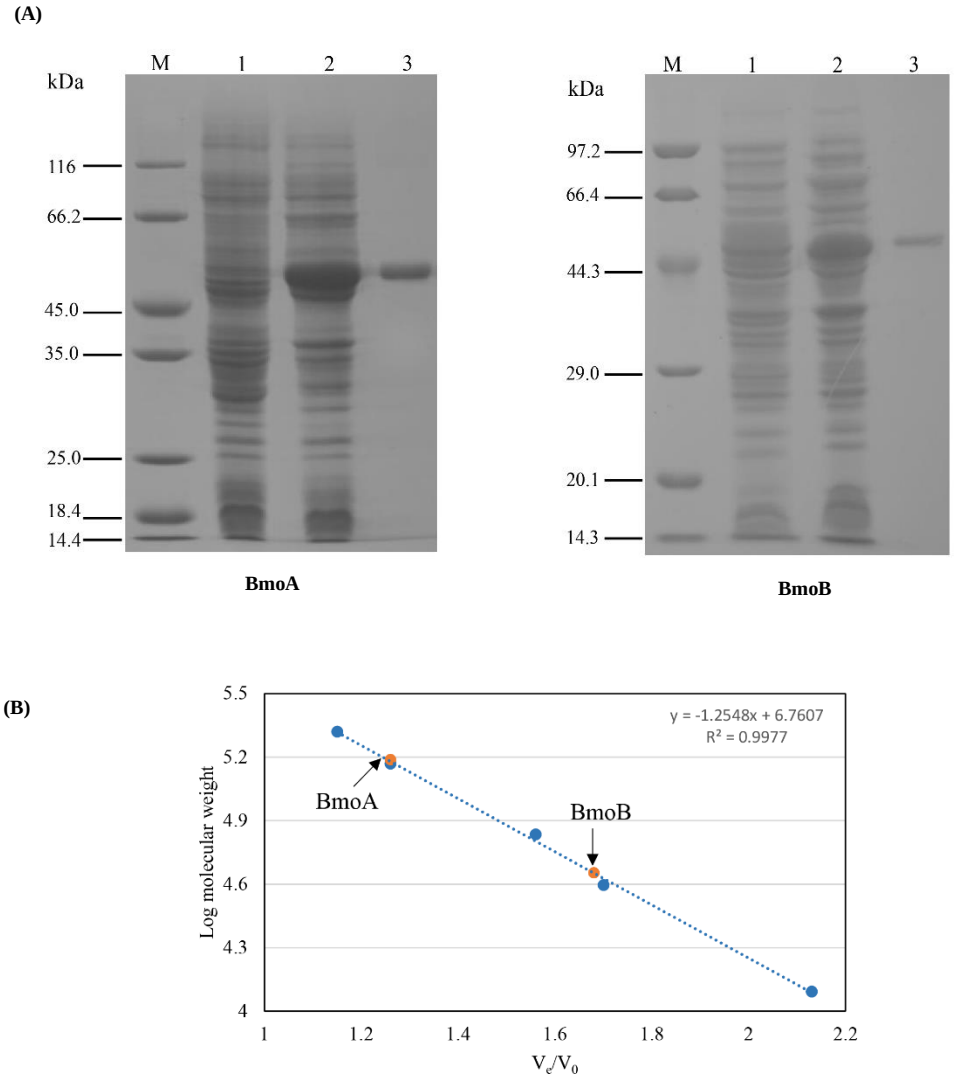

FIG S5 Estimation of molecular masses of the purified enzyme subunits (A) and molecular weight determination of the native recombinant enzymes (B). (**A**) SDS-polyacrylamide gel electrophoresis of BmoA-His_6 (right) and His_6-BmoB (left) expressed in *E. coli* BL21(DE3). lane M, molecular mass standards; lane 1, soluble fractions of *E. coli* BL21(DE3) harboring pET28a; lane 2, soluble fraction of *E. coli* BL21(DE3) harboring pET28a-bmoA (right) or *E. coli* BL21(DE3) harboring pET28a-bmoB (left); lane 3, purified BmoA-His_6 (109 μg, right) or His_6-BmoB (67 μg, left). (**B**) Molecular weight determination of BmoA-His_6 and His_6-BmoB by Sephacryl S-200 gel chromatography. Protein molecular weight markers (shown in parentheses) were as follows: sweet potato β-amylase (200,000), yeast alcohol dehydrogenase (150,000), bovine serum albumin (66,000), ovalbumin (43,000), and cytochrome c from horse heart (12,400). Arrows indicated the purified BmoA-His_6 and His_6-BmoB.

.

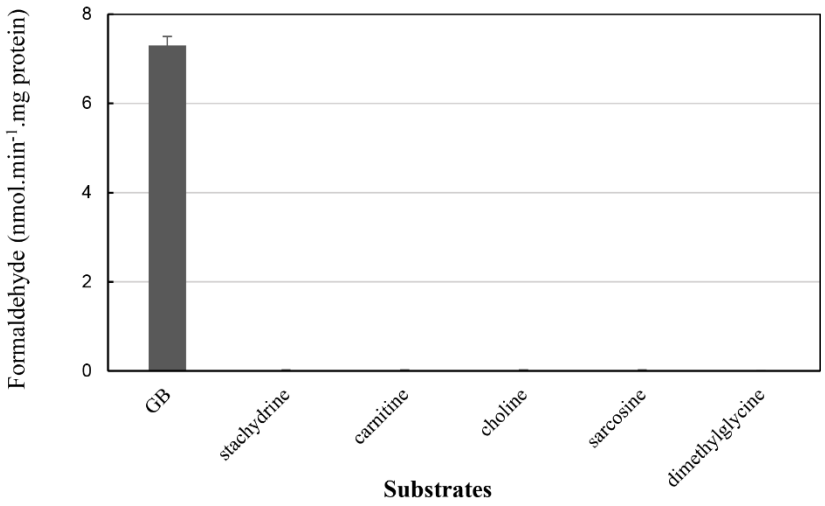

FIG S6 Substrate specificity assay of glycine betaine monooxygenase (BmoA/BmoB). Reactions were carried out under the assay conditions as described in MATERIALS AND METHODS except that glycine betaine was replaced by other tested substrates as indicated, and production of formaldehyde was measured spectrophotometrically.

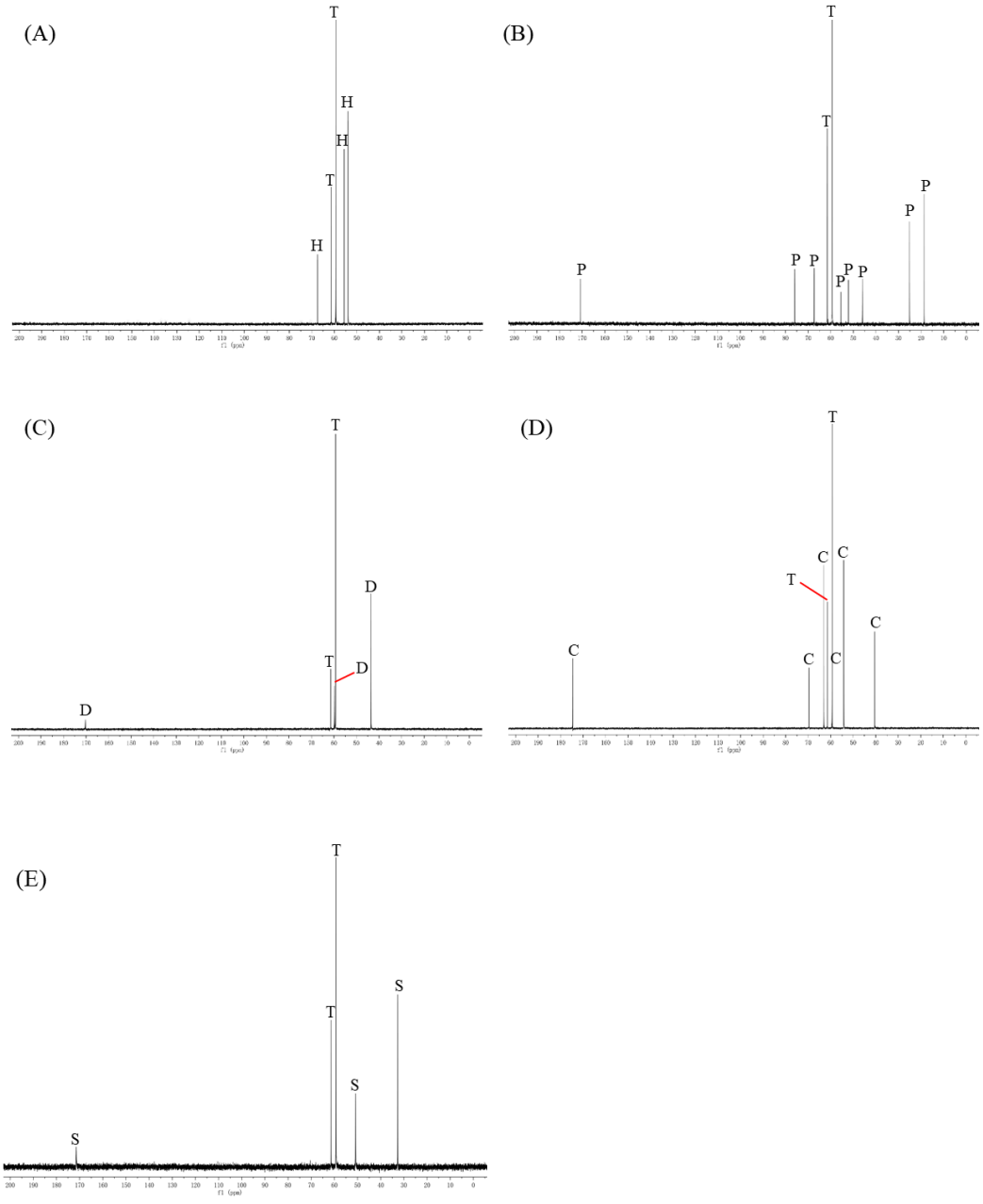

FIG S7 Natural-abundance ^{13}C-NMR spectra of the reaction products by glycine betaine monooxygenase (BmoA/BmoB) on various substrates. Reactions were carried out under the assay conditions as described in MATERIALS AND METHODS except that glycine betaine was replaced by other compounds [(A), choline; (B), stachydrine; (C), dimethylglycine; (D), L-carnitine; (E), sarcosine.] The signals were as follows: Tris-HCl (T), choline (H), stachydrine (P), dimethylglycine (D), L-carnitine (C), and sarcosine (S).