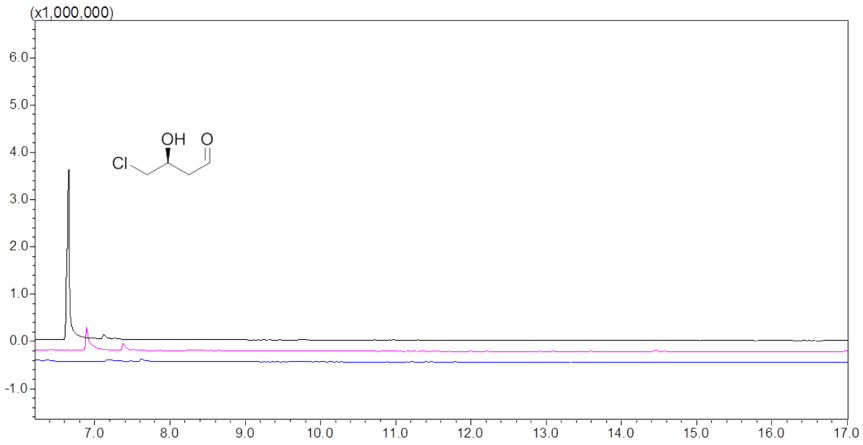
Supporting Information

Synthesis of statin precursor in high space-time yield by a new aldehyde-tolerant aldolase identified from *Lactobacillus brevis*

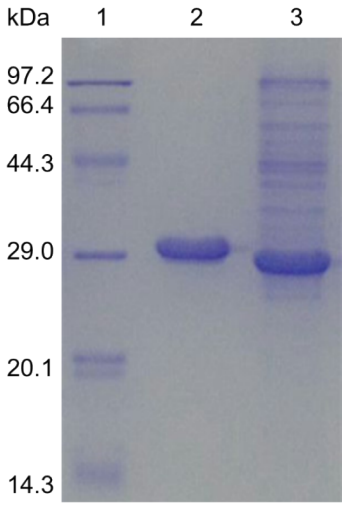
Laboratory of Biocatalysis and Synthetic Biotechnology, State Key Laboratory of Bioreactor Engineering, East China University of Science and Technology, 130 Meilong Road, Shanghai 200237, China.

* To whom correspondence should be addressed. Fax: +86-21-6425-0840; E-mail: panjiang@ecust.edu.cn (J.P.); jianhexu@ecust.edu.cn (J.H.X.).

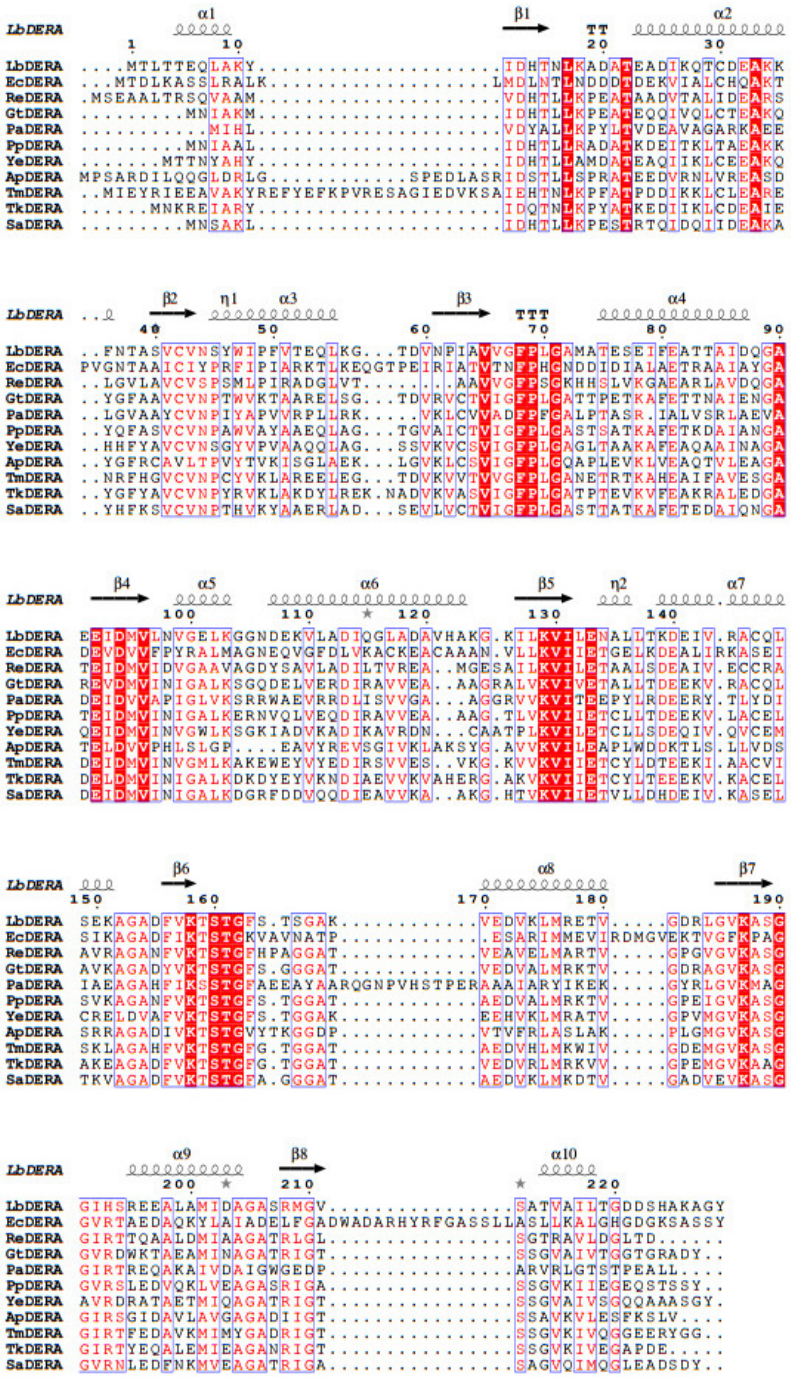
Table of Contents



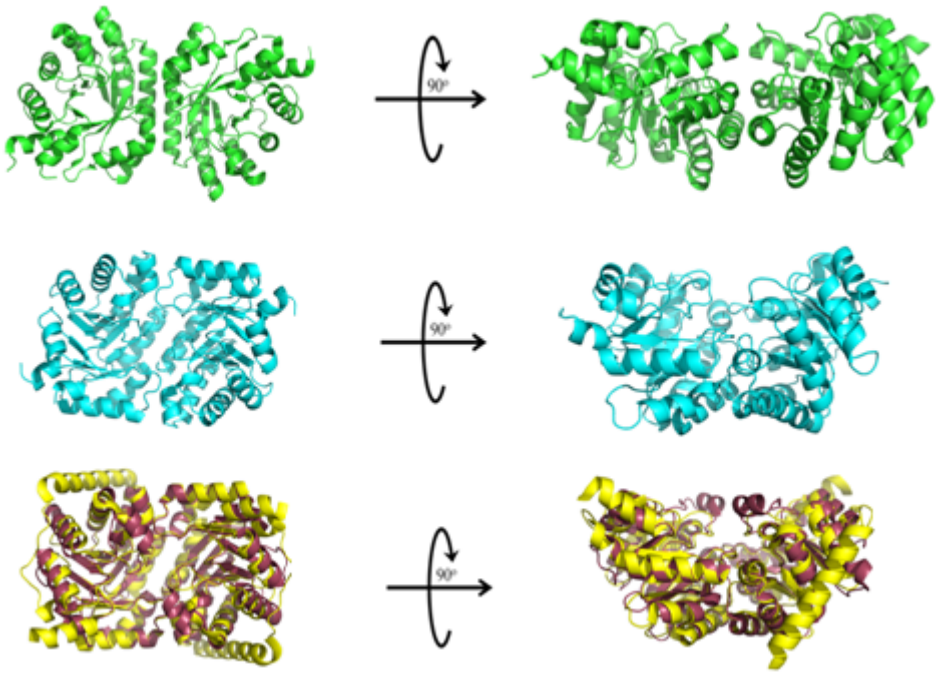
1. Figure S1 Screening of microorganisms for aldol condensation of acetaldehyde and chloroacetaldehyde using GC-MS. Reactions of DERAs were carried out as follows: Cells of 30 g DCW/L were resuspended in 1 mL KPB (100 mM, pH 7.0) with 100 mM acetaldehyde and 50 mM chloroacetaldehyde in 2-mL eppendorf tubes. The reaction mixture was shaken at 1000 rpm and 30°C for 12 h. Black line: *Lb*DERA; red line: *Bt*DERA; blue line: *Es*DERA.



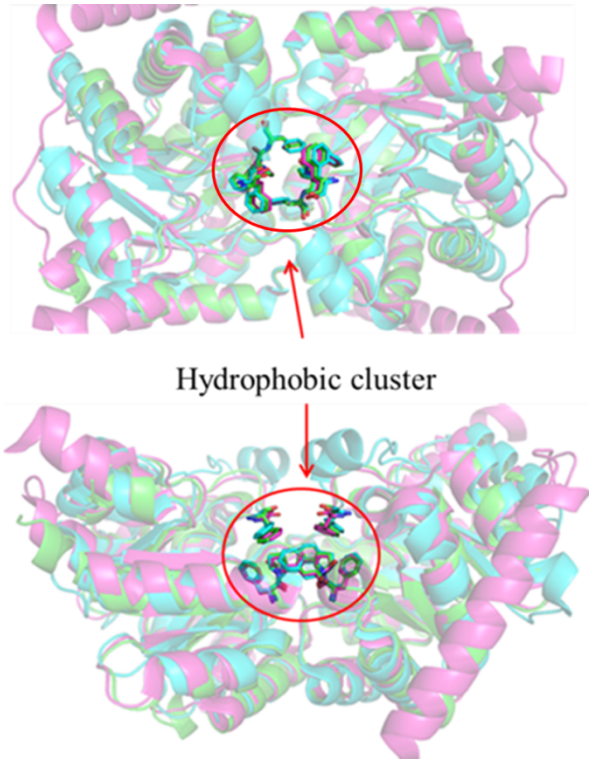
2. Figure S2 SDS-PAGE analysis of the purified *Ec*DERA. Lane 1, protein markers; lane 2, purified enzyme; lane 3, crude extract. Protein bands were visualized by Coomassie brilliant blue.



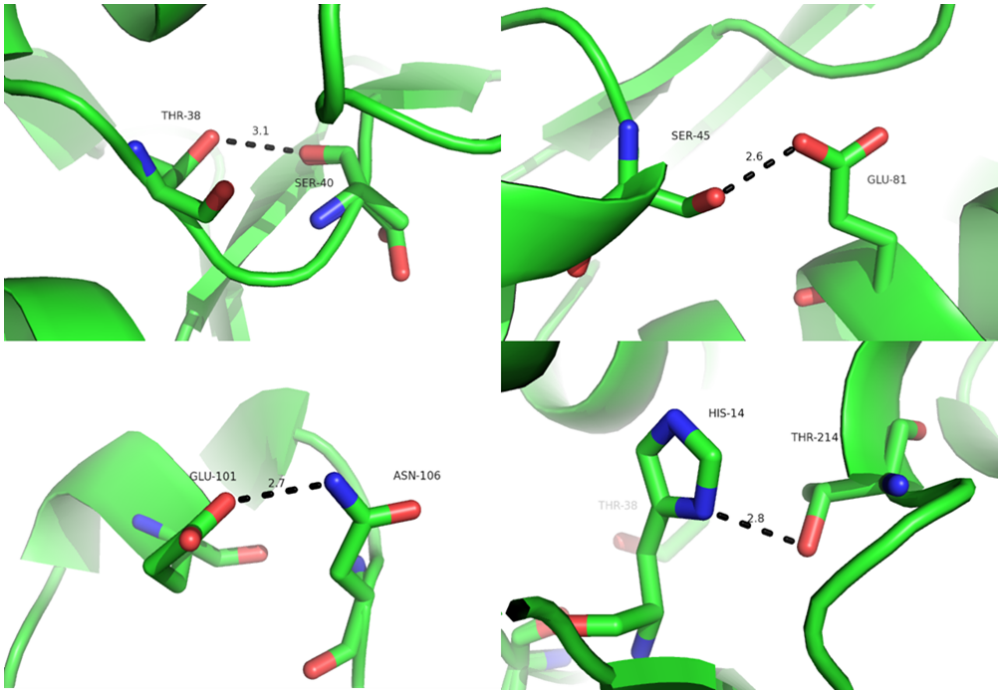
3. Figure S3 Multiple sequence alignment of *Lb*DERA with proteins from different sources. Conserved residues are shaded in red background. The catalytic residues are shown in blue stars. The schematic diagram of secondary structure of DERA is shown on the top of sequence. The α-helices, 3_{10}-helices and π-helices are displayed as medium, small and large squiggles respectively. β-strands are rendered as arrows, strict β-turns as TT letters and strict α-turns as TTT.



4. Figure S4. Dimeric structure of four DERAs. *Ec*DERA (green), *Lb*DERA (blue), *Pa*DERA (yellow) and *Tm*DERA (aubergine).



5. Figure S5. Hydrophobic cluster of *Lb*DERA, *Pa*DERA and *Tm*DERA.



6. Figure S6. The hydrogen bonds or salt bridges between Thr38-Ser40, Ser45-Glu81, Glu101-Asn106 and Thr214-His14 were destroyed by the newly introduced mutations.

7. Table S1. Primers used in this work.

*Lb*DERA_for	ACGGAATTCATGACATTAACCACA (*EcoR* I)
*Lb*DERA_rev	CCCAAGCTTTTAGTAACCAGCTTT (*Hind* III)
*Bt*DERA_for	ACGGAATTC ATGAACATTGCAAAG (*EcoR* I)
*Bt*DERA_rev	CCCAAGCTTTTAGTAGTTATCTGT (*Hind* III)
*Es*DERA_for	ACGGAATTCATGAATTTAGCAGGA (*EcoR* I)
*Es*DERA_rev	CCCAAGCTTTTAGTAGTCGCTGTC (*Hind* III)
*Ec*DERA_for	ACGGAATTCATGACTGATCTGAAA (*EcoR* I)
*Ec*DERA_rev	CCCAAGCTTTTAGTAGCTGCTGGC (*Hind* III)
T29L_for	TTCTTAGCTTCGTCACATAATTGCTTGATATCTGCTTCAGTTGCATCTGCC
T29L_rev	GGCAGATGCAACTGAAGCAGATATCAAGCAATTATGTGACGAAGCTAAGAA
T38F_for	GTTTACGCAGACCGAAGCGAAGTTAAATTTCTTAGCTTCGTCACAAG
T38F_rev	CTTGTGACGAAGCTAAGAAATTTAACTTCGCTTCGGTCTGCGTAAAC
S45P_for	CTTCGGTCTGCGTAAACCCCTATTGGATTCCGTTT
S45P_rev	AAACGGAATCCAATAGGGGTTTACGCAGACCGAAG
I48V_for	TGCGTAAACTCCTATTGGGTTCCGTTTGTAACTGAGC
I48V_rev	GCTCAGTTACAAACGGAACCCAATAGGAGTTTACGCA
E78K_for	CGTGGCTTCAAAGATTTTACTTTCCGTTGCCATGG
E78K_rev	CCATGGCAACGGAAAGTAAAATCTTTGAAGCCACG
E101A_for	TATGGTCTTGAACGTAGGTGCATTAAAAGGCGGTAATGATG
E101A_rev	CATCATTACCGCCTTTTAATGCACCTACGTTCAAGACCATA

L128V_for	CAGTCCACGCTAAGGGTAAAATTGTAAAAGTTATCTTGGAAAATG
L128V_rev	CATTTTCCAAGATAACTTTTACAATTTTACCCTTAGCGTGGACTG
K139D_for	TGGAAAATGCGTTGTTAACGGATGATGAAATTGTCCGGGCTTG
K139D_rev	CAAGCCCGGACAATTTCATCATCCGTTAACAACGCATTTTCCA
S166G_for	ACCTTAGCACCTCCTGTTGAGAACCCAGTGGACGTCT
S166G_rev	AGACGTCCACTGGGTTCTCAACAGGAGGTGCTAAGGT
K169T_for	GGTTCTCAACATCAGGTGCTACGGTTGAAGATGTTAAGTT
K169T_rev	AACTTAACATCTTCAACCGTAGCACCTGATGTTGAGAACC
H193R_for	CGTCAAGGCTTCCGGTGGTATCCGCAGTCGTGAAGAAGCC
H193R_rev	GGCTTCTTCACGACTGCGGATACCACCGGAAGCCTTGACG
M209I_for	GTTGCACTGACACCTATCCGACTGGCACCAG
M209I_rev	CTGGTGCCAGTCGGATAGGTGTCAGTGCAAC
T214G_for	CGGATGGGTGTCAGTGCAGGCGTAGCCATCTTAAC
T214G_rev	GTTAAGATGGCTACGCCTGCACTGACACCCATCCG

8. Table S2 Aldehyde tolerability of the mutants based on consensus sequence approach.

	Residual activity (%) [a]
WT	33 ± 3
T29L	35
T38F	16
S45P	10
I48V	33
E78K	**52**
E101A	8
L128V	24
K139D	29
S166G	34
K169T	31
H193R	31
M209I	30
T214G	6

[a] The residual activity after kept in 300 mM acetaldehyde for 1 h, which was test with supernatant of cell lysate.

9. Table S3 X-ray data collection and refinement statistics.

	*Lb*DERA	*Lb*DERA$_{E78K}$
Datacollection statistics		
Wavelength (Å)	1.5418	1.5418
Space group	P41212	P41212
Cell axial lengths (Å)	a = b = 96.882 c = 119.593	a = b= 96.882 c = 119.593
Resolution rang (Å)[a]	44.898-1.951 (2.02-1.95)	36.572-2.170 (2.24-2.17)
Total/Unique reflections	42,101/1531	30,632/1963
Redundancy[a]	27.5 (27.2)	15.6 (15.3)
Average(I/δ)[a]	18.5 (3.0)	19.5 (4.3)
Completeness (%)[a]	99.9 (100)	100 (100)
R_{merge} (%)[a,b]	0.103 (0.684)	0.160 (0.745)
Refinement statistics		
PDB ID	4XBK	4XBS
No. of reflections	41993	30568
R-factor/*R*-free (%)[c]	16.4/18.6	15.6/18.3
R.m.s deviations from ideal geometry		
Bond lengths (Å)	0.008	0.007
Bond angles (°)	1.061	1.036
Ramachandran plot		
Favored (%)	98.39	98.40
Allowed (%)	1.61	1.60
Disallowed (%)	0	0

[a] Numbers in parentheses are values for the highest-resolution shell.

[b] $R_{merge}= \sum_{hkl}|I_i - I_m|/I_{hkl}I_m$, where I_i and I_m are the observed intensity and the mean intensity of related reflections, respectively.

[c] *R*-factor = $\sum||F_o– F_c||/ac F_o|$. $R_{free}= \sum_T||F_o– F_c||/e_T|F_o|$, where *T*is the summation over 5% of the reflections chosenrandomly.

10. Table S4 Aldehyde tolerability of the mutants related to hydrophobic cluster.

	Residual activity (%) [a]
WT	33 ± 3
F68A	1.1
P69A	1.4
L70A	1.6
F163A	16

[a] The residual activity after kept in 300 mM acetaldehyde for 1 h, which was test with supernatant of cell lysate.