



STRUCTURAL BIOLOGY
COMMUNICATIONS

Volume 74 (2018)

Supporting information for article:

Crystal structure of an inferred ancestral, bacterial pyruvate decarboxylase

Lisa Buddrus, Emma S. V. Andrews, David J. Leak, Michael J. Danson, Vickery L. Arcus and Susan J. Crennell

S1. ANC27 DNA sequence

ATGACGTATACGGTCGGCCATTATTTGGCGACGCGCTTGGCGCAAATTGGCTTGAAACAT
CATTTTGCGGTCGCGGGCGATTATAACTTGGTCTTGTTGGATCAATTGTTGAAAAACAAA
GATTTGGAACAAGTCTATTGCTGCAACGAATTGAACTGCGGCTTTAGCGCGGAAGGCTAT
GCGCGCGCGAACGGCGTCGGCGCGGGCGGTCGTCACGTTTAGCGTCGGCGCGTTGAGCGC
GTTTAACGCGATTGGCGGCGCGTATGCGGAAAACCTTGCCGGTCATTTTGATTAGCGGCGC
GCCGAACACGAACGATCATGGCAGCGGCCATATTTTGCATCATACGATTGGCACGACGG
ATTATGGCTATCAATTGGAAATGGCGAAACAAATTACGTGCGCGGGCGGTCAGCATTACG
CATGCGGAAGATGCGCCGGCGTTGATTGATCATGCGATTTCGCACGGCGTTGCGCGAAAA
AAAACCGGCGTATATTGAAATTGCGTGCAACGTCGCGGGCGCAACCGTGCGCGCGCCCCGG
GCCCCGTCAGCGCGTTGTTGAACGAACCGACGAGCGATGAAGAAACGTTGAAAGCGGCG
GTCGAAGCGGCGTTGGATTTTATTGAAAAACGCGAAAAACCGGTCTTGTTGGTCGGCGG
CAAATTGCGCGCGGCGGGCGCGGAAGAAGCGGTCGTCGAATTGGCGGATGCGTTGGGCT
GCGCGGTGCGGACGATGGCGGCGGCGAAGAGCTTTTTTCCGGAAGATCATCCGGGCTAT
GTCGGCACGTATTGGGGCGAAGTCAGCAGCCCCGGGCGTCGAAGAAATTGTCGATTGGGC
GGATGGCATTATTTGCTTGGGCCCGGTCTTTAACGATTATAGCACGGTCGGCTGGACGGC
GTGGCCGAAAGGCGAAAACGTCGTCTTGGTCGATCCGCATCATATTACGGTCGGCGGCG
AAGAATTTACGGGCATTCATTTGAAAGATTTTTTTGACGGCGTTGACGGAACGCGTCCCGA
AAAAAGATGCGACGTTGGATCAATTTAAAGCGCGCGTCGGCAAACCGGCGGCGGAAAA
AGTCCCGGCGGCGGACCCGAACGCGCCGTTGACGCGCGCGGAATTGTGCCGCCAAATTC
AAGGCTTGTTGAACCCGAACACGACGTTGATTGCGGAAACGGGCGATAGCTGGTTTAAC
GCGATGCGCATGAAATTGCCGCATGGCGCGCGCGTCGAATTGGAAATGCAATGGGGCCA
TATTGGCTGGAGCGTCCCGGCGACGTTTGGCTATGCGGTGCGGGAACCGGAACGCCGCA
ACGTCTTGATGGTCGGCGATGGCAGCTTTCAATTGACGGCGCAAGAAGTCGCGCAAATG
GTCCGCCGCAAATTGCCGATTATTATTTTTTTGATTAACAACCGCGGCTATACGATTGAA
GTCAAAATTCATGATGGCCCGTATAACAACATTA AAAACTGGGATTATGCGGGCTTGATG
GAAGTCTTTAACGCGGAAGATGGCAAAGGCTTGGGCTTGAAAGCGACGACGGGCGGCGA
ATTGGCGGAAGCGATTAAAAAAGCGTTGGCGCATCGCGAAGGCCCGACGTTGATTGAAT
GCGTCATTGATCGCGATGATTGCACGCCGGAATTGGTCACGTGGGGCAAAAAAGTCGCG
ACGGCGAACGCGCGCCCCGCCGAAGCGATTCTGGTGCCGCGCGGCAGCGGCGGCGGCCT
CGAGCACCACCACCACCACCACTGA

The sequence encoding the additional purification tag is underlined.

Table S1 Sequences used in ASR. *Zymobacter palmae* PDC (PDB entry 5euj) is the reference protein.

Strain	GenBank accession number (amino acid/nucleotide)	% amino acid identity to ZpPDC
<i>Acetobacter aceti</i>	KDE19620.1/JEOA01000009.1	76
<i>Acetobacter malorum</i>	KFL91998.1/JOJU01000008.1	63
<i>Acetobacter nitrogenifigens</i>	WP_026398870.1/NZ_AUBI01000020.1	68
<i>Acetobacter papaya</i>	WP_025859940.1/NZ_BAIN01000005.1	62
<i>Acetobacter pasteurianus</i>	AF368435.1	76
<i>Acetobacter persici</i>	WP_025828656.1/NZ_BAJW01000064.1	69
<i>Acetobacter pomorum</i>	WP_006115789.1/NZ_AEUP01000018.1	76
<i>Acetobacter tropicalis</i>	WP_006559524.1/NZ_BABS01000096.1	62
<i>Acidomonas methanolica</i>	GAJ29946.1/BAND01000085.1	71
<i>Beijerinckia indica</i>	WP_012386294.1/NC_010581.1	65
<i>Commensalibacter sp MX01</i>	EUK18520.1/ATSX01000001.1	56
<i>Gluconacetobacter diazotrophicus</i>	AIG13066.1/KJ746104.1	71
<i>Gluconobacter frateurii</i>	WP_010505240.1/NZ_BADZ01000024.1	70
<i>Gluconobacter morbifer</i>	WP_008852112.1/NZ_AGQV01000006.1	69
<i>Gluconobacter oxydans</i>	KF650839.1	67
<i>Gluconobacter thailandicus</i>	WP_007283613.1/NZ_BAON01000015.1	68
<i>Komagataeibacter europaeus</i>	WP_019085054.1/NZ_CADP01000006.1	65
<i>Komagataeibacter hansenii</i>	WP_003622049.1/NZ_CM000920.1	64
<i>Komagataeibacter medellinensis</i>	WP_014105323.1/NC_016027.1	64
<i>Komagataeibacter oboediens</i>	WP_010515737.1/NZ_CADT01000033.1	66
<i>Kozakia baliensis</i>	WP_029604671.1/NZ_JNAB01000024.1	78
<i>Ktedonobacter racemifer</i>	WP_007922190.1/NZ_ADVG01000005.1	52
<i>Sarcina ventriculi</i>	AF354297.1	31
<i>Zymomonas mobilis</i>	AAA27696.2/M15393.2	63

Table S2 Sequence comparison of ANC27 to all PDCs used as ASR inputs. This information was calculated in Geneious using ClustalW alignments.

	Pairwise amino acid sequence identity in %
<i>Acetobacter aceti</i>	68.4
<i>Acetobacter malorum</i>	74.0
<i>Acetobacter nitrogenifigens</i>	64.7
<i>Acetobacter papaya</i>	73.1
<i>Acetobacter pasteurianus</i>	67.0
<i>Acetobacter persici</i>	66.3
<i>Acetobacter pomorum</i>	68.2
<i>Acetobacter tropicalis</i>	73.1
<i>Acidomonas methanolica</i>	67.5
<i>Beijerinckia indica</i>	77.0
<i>Commensalibacter sp MX01</i>	67.1
<i>Gluconacetobacter diazotrophicus</i>	74.4
<i>Gluconobacter frateurii</i>	67.0
<i>Gluconobacter morbifer</i>	67.8
<i>Gluconobacter oxydans</i>	66.1
<i>Gluconobacter thailandicus</i>	66.7
<i>Komagataeibacter europaeus</i>	76.3
<i>Komagataeibacter hansenii</i>	77.7
<i>Komagataeibacter medellinensis</i>	78.1
<i>Komagataeibacter oboediens</i>	76.7
<i>Kozakia baliensis</i>	70.0
<i>Ktedonobacter racemifer</i>	60.9
<i>Sarcina ventriculi</i>	33.0
<i>Zymobacter palmae</i>	68.8
<i>Zymomonas mobilis</i>	71.1

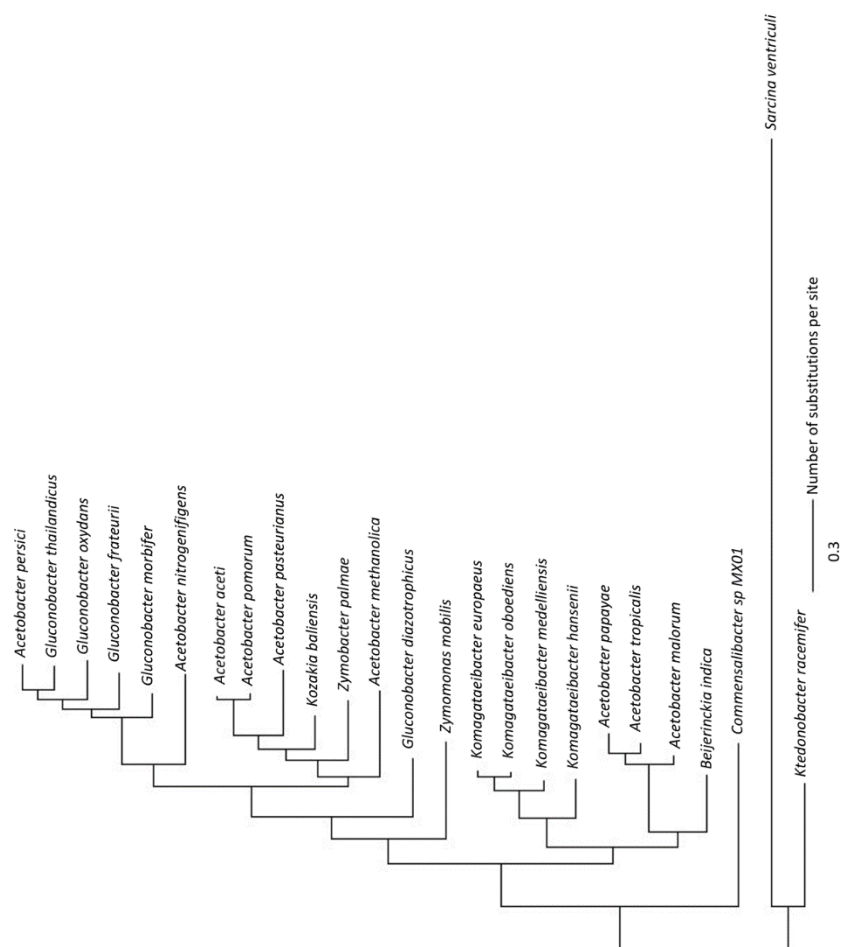


Figure S1 Maximum likelihood phylogram based on PDC amino acid sequences – the best ML GARLI tree. The WAG model of evolution and the PDC amino acid alignment based on phylogenetic information were implemented in GARLI to find the best ML tree. *S. ventriculi* and *K. racemifer* represent the outgroup.

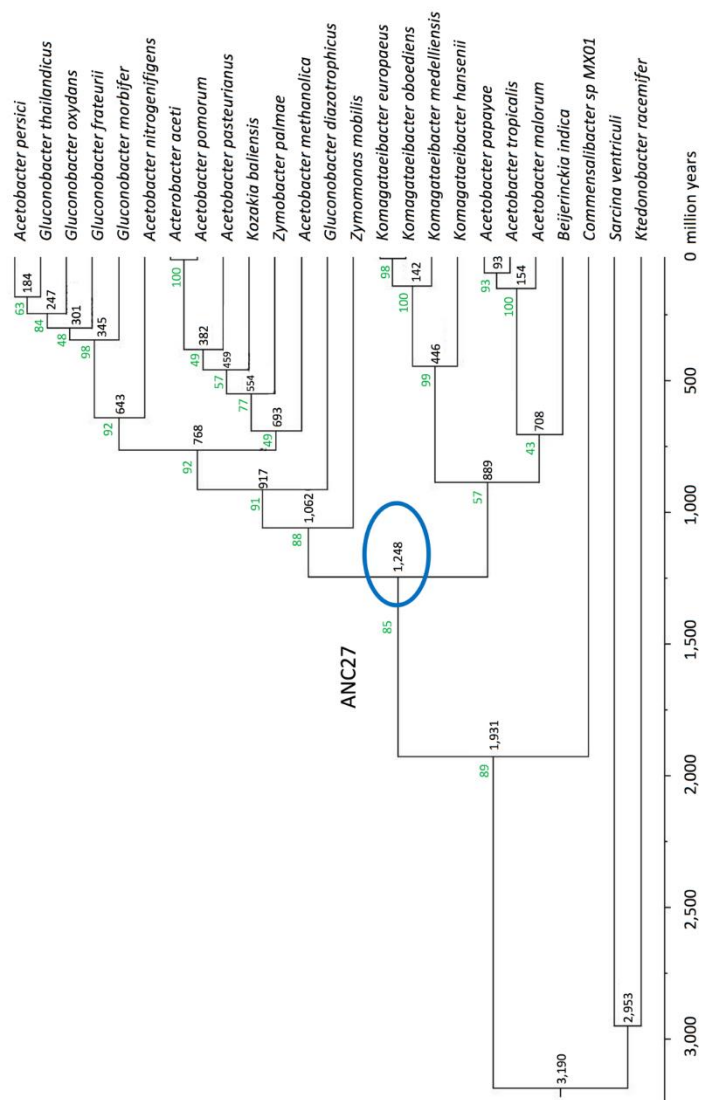


Figure S2 Maximum likelihood chronogram based on PDC amino acid sequences. ANC27 (1,248 million years old) is highlighted in blue. Inferred node ages are indicated in black. Bootstrap percentages were assessed by 1,024 bootstrap replicates and are given in green. *S. ventriculi* and *K. racemifer* represent the outgroup. The point of divergence of Proteobacteria and Firmicutes (3.19 billion years ago) was used as the calibration point for node age estimates.

Figure S3 PROMALS3D structural sequence alignment. Each representative sequence has a magenta name and is coloured according to PSIPRED secondary structure predictions (red: alpha-helix, blue: beta-strand). The last two lines show consensus amino acid sequence (Consensus_aa) and consensus predicted secondary structures (Consensus_ss). Consensus predicted secondary structure symbols: alpha-helix: h; beta-strand: e. Consensus amino acid symbols are: conserved amino acids are in **bold** and uppercase letters; aliphatic (I, V, L): *l*; aromatic (Y, H, W, F): @; hydrophobic (W, F, Y, M, L, I, V, A, C, T, H): *h*; alcohol (S, T): o; polar residues (D, E, H, K, N, Q, R, S, T): p; tiny (A, G, C, S): t; small (A, G, C, S, V, N, D, T, P): s; bulky residues (E, F, I, K, L, M, Q, R, W, Y): b; positively charged (K, R, H): +; negatively charged (D, E): -; charged (D, E, K, R, H): c. The ANC27 R-PP linker region is indicated by the black box.

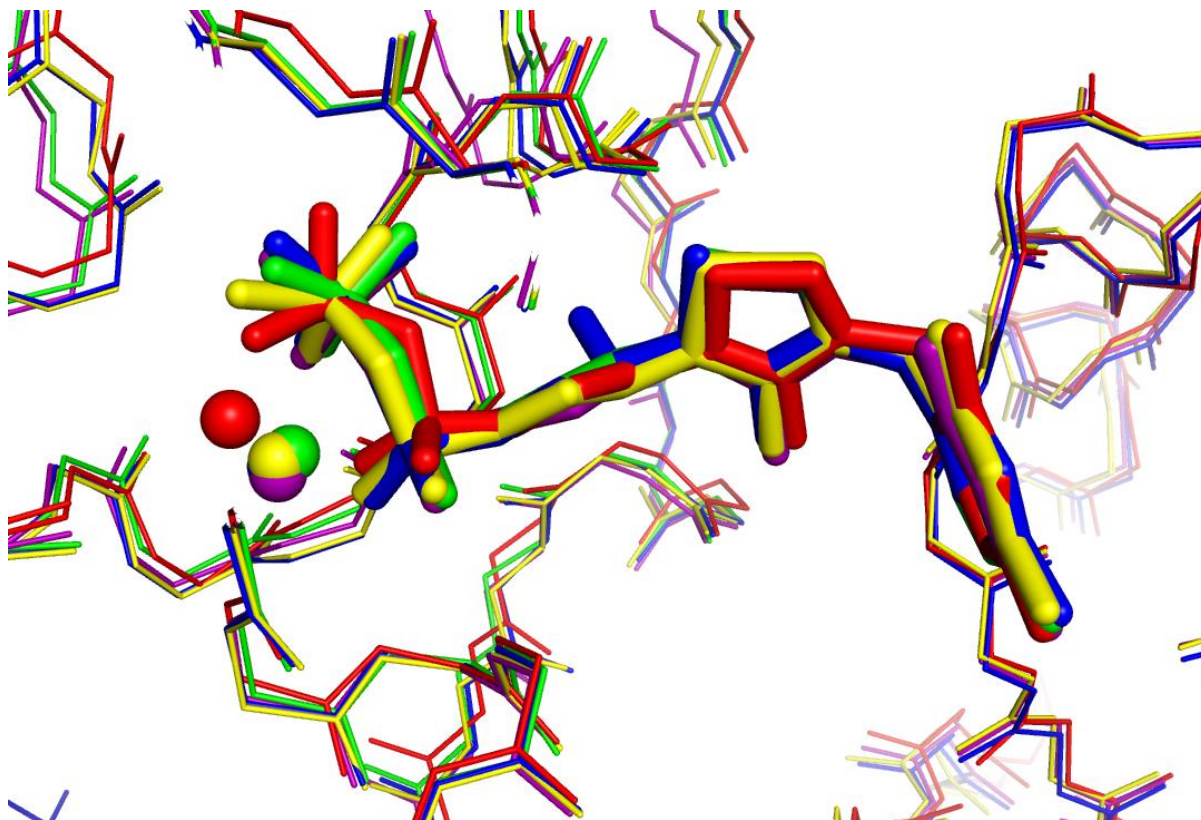


Figure S4 Comparison of ligand geometry. Superposition of bacterial PDCs ANC27 5npu (red), ApPDC 2vbi (purple), GdPDC 4cok (yellow), ZmPDC 1zpd (blue) and ZpPDC 5euj (green). The main chains are displayed as lines, TPP molecules as sticks and magnesium ions as spheres.