1 **Zymomonas mobilis ZM4 utilizes an acetaldehyde dehydrogenase to produce**
2 **acetate**

3 Magdalena M. Felczak[1] and Michaela A. TerAvest[1]

4 [1]Department of Biochemistry and Molecular Biology, Michigan State University, East
5 Lansing, MI, USA

6 Abstract

7 *Zymomonas mobilis* is a promising bacterial host for biofuel production but further

8 improvement has been hindered because some aspects of its metabolism remain poorly

9 understood. For example, one of the main byproducts generated by *Z. mobilis* is acetate

10 but the pathway for acetate production is unknown. Acetaldehyde oxidation has been

11 proposed as the major source of acetate and an acetaldehyde dehydrogenase was

12 previously isolated from *Z. mobilis via* activity guided fractionation, but the

13 corresponding gene has never been identified. We determined that the locus ZMO1754

14 (also known as ZMO_RS07890) encodes an NADP$^+$-dependent acetaldehyde

15 dehydrogenase that is responsible for acetate production by *Z. mobilis*. Deletion of this

16 gene from the chromosome resulted in a growth defect in oxic conditions, suggesting

17 that acetaldehyde detoxification is an important role of acetaldehyde dehydrogenase.

18 The deletion strain also exhibited a near complete abolition of acetate production, both

19 in typical laboratory conditions and during lignocellulosic hydrolysate fermentation. Our

20 results show that ZMO1754 encodes the major acetaldehyde dehydrogenase in *Z.*

21 *mobilis* and we therefore rename the gene *aldB* based on functional similarity to the

22 *Escherichia coli* acetaldehyde dehydrogenase.

23

24

25

Importance

Biofuel production from non-food crops is an important strategy for reducing carbon emissions from the transportation industry but it has not yet become commercially viable. An important avenue to improve biofuel production is to enhance the characteristics of fermentation organisms by genetic engineering. To make genetic modifications successful, we must gain sufficient information about the genome and metabolism of the organism to enable rational design and engineering. Here, we improved understanding of *Zymomonas mobilis*, a promising biofuel producing bacterium, by identifying a metabolic pathway and associated gene that lead to byproduct formation. This information may be used in the future for genetic engineering to reduce byproduct formation during biofuel production.

Introduction

Zymomonas mobilis is an α-proteobacterium that produces ethanol from glucose with a very high specificity and may be useful for biofuel production (1). Under optimal conditions, *Z. mobilis* converts glucose to ethanol at up to 97% of the theoretical yield, converting only a small fraction of the carbon into byproducts or biomass (2, 3). However, under less favorable conditions, *Z. mobilis* generates byproducts, including acetaldehyde and acetate, especially when exposed to oxygen (4–6). Stressors present in lignocellulosic hydrolysates may also lead to byproduct formation. We previously observed that *Z. mobilis* produces 5-10 mM acetate when grown in anaerobic lignocellulosic hydrolysate (7). Acetate production during lignocellulose fermentation is problematic because it reduces the efficiency of biofuel production and further increases the acetate concentration. Acetate is formed by many biomass hydrolysis methods, leading to high acetate concentrations in lignocellulosic hydrolysates, which are often

50 inhibitory to the fermentation organism (8, 9). Because of the already inhibitory

51 concentrations and low value of acetate, generation of this molecule during hydrolysate

52 fermentations is highly undesirable.

53   Although acetate production by *Z. mobilis* has been well-documented, the

54 metabolic pathway that leads to significant acetate secretion is unknown (10). One

55 possible source of acetate secreted by *Z. mobilis* is oxidation of acetaldehyde, which is

56 an intermediate in the ethanol production pathway. Acetaldehyde is typically

57 undetectable in anoxic cultures of *Z. mobilis* but accumulates in oxic cultures because

58 of a redox imbalance. NAD(P)H that is oxidized by the respiratory chain is not available

59 to reduce acetaldehyde to ethanol, leading to a buildup of acetaldehyde (**Figure 1**).

60 Under oxic conditions, oxidation of acetaldehyde to acetate provides a possible route for

61 detoxification, with additional NAD(P)H being recycled by the respiratory chain.

62   An $NADP^+$-dependent acetaldehyde dehydrogenase was discovered in *Z. mobilis*

63 strain Z6 in 1989 *via* activity guided fractionation, suggesting that acetaldehyde

64 oxidation could be a major source of acetate (11). However, there is no acetaldehyde

65 dehydrogenase annotated in the *Z. mobilis* Z6 genome (12). Further, we have not found

66 an annotated acetaldehyde dehydrogenase in any other *Z. mobilis* genome (7). As a

67 result, most genome scale metabolic models for *Z. mobilis* only include acetate as a

68 byproduct of less central reactions, such as amino acid biosynthesis (10, 13, 14). Only

69 one model includes acetaldehyde dehydrogenase, although no gene evidence was

70 provided for the reaction (15). Metabolic models of *Z. mobilis* tend to underestimate

71 acetate production compared with experimental results (10), suggesting that a major

72 source of acetate is missing. To address this gap in our knowledge of *Z. mobilis*

73 metabolism, we identified a locus in the *Z. mobilis* ZM4 genome encoding an

74 acetaldehyde dehydrogenase and determined that it is the main source of acetate

75 production under laboratory growth conditions and during lignocellulosic hydrolysate

76 fermentation. Here, we refer to the locus ZMO1754 as *aldB* due to the functional

77 similarity to the *E. coli* gene of the same name. The ZM4 strain was used because it is

78 the most promising strain for hydrolysate fermentation (7) and recently developed

79 genome modification tools facilitated generation of a markerless deletion strain in this

80 background (16).

81 <u>Materials and Methods</u>

82 *Media and Chemicals*

83 *Escherichia coli* strains were grown in Luria-Bertani medium (Miller, Accumedia).

84 *Zymomonas mobilis* ZM4 was grown in ZRMG (1% yeast extract, 2% glucose, and 15

85 mM KH_2PO_4) or ZMMG (1.0 g KH_2PO_4, 1.0 g K_2HPO_4, 0.5 g NaCl, 1.0 g $(NH_4)_2SO_4$, 0.2

86 g $MgSO_4$ x 7 H_2O, 0.025 g Na_2MoO_4 x 2 H_2O, 0.025 g $FeSO_4$ x 7 H_2O, 0.010 g $CaCl_2$ x

87 2 H_2O, 0.001 g calcium pantothenate, 20.0 g D-glucose per 1 liter). Solid media

88 contained 1.5% agar (Difco). Chloramphenicol was added to 100 μg/ml or 35 μg/ml, and

89 spectinomycin to 100 μg/ml or 50 μg/ml for *Z. mobilis* and *E. coli*, respectively. Ampicillin

90 was used at 100 μg/ml for *E. coli.* 2,6-Diaminopimelic acid (DAP, Aldrich) was added to

91 a final concentration of 0.3 mM when needed. B-PER Reagent was from Thermo

92 Scientific, Profinity immobilized metal affinity chromatography (IMAC) Ni-charged Resin

93 was from Bio-Rad; all IMAC buffers contained 20 mM Tris HCl pH 8, 0.3 M NaCl, 10%

94 glycerol and imidazole at 5 mM, 20 mM or 500 mM for IMAC A, B and C, respectively.

95 Complete EDTA-free Protease Inhibitor tablets were from Sigma-Aldrich/Roche.

96   Bradford Protein Assay Reagents 1 and 2 and Precision Plus Protein All Blue

97   Prestained Protein Standard were from Bio-Rad. The anti His-tag mouse antibody was

98   from GenScript (cat no. A00186). Horseradish peroxidase-conjugated rabbit anti-mouse

99   antibody was from Sigma-Aldrich. TBST buffer is 20 mM Tris-HCl pH 7.5, 150 mM

100   NaCl, 0.01% Tween-20. Clarity Western ECL was from BioRad. $NADP^+$ sodium salt was

101   from Sigma-Aldrich. NAD^+ was from Amresco. Acetaldehyde, propionaldehyde,

102   butyraldehyde, valeraldehyde, formaldehyde and glyceraldehyde were from Sigma-

103   Aldrich/Millipore. Restriction enzymes, T4 ligase, Q5DNA Polymerase and HiFi DNA

104   Assembly Master Mix were from New England Biolabs. Egg lysozyme was from Sigma-

105   Aldrich and IPTG was from Goldbio. Oligonucleotides were synthesized by Integrated

106   DNA Technologies (IDT).

107   *Bacterial strains and plasmids*

108   *Zymomonas mobilis* ATCC 31821 (ZM4) was obtained from Dr. Robert Landick

109   (University of Wisconsin, Madison). *E. coli* strains Mach 1 and BL21(DE3) pLysS, were

110   from Invitrogen. *E. coli* WM6026 (*lacIq rrnB3* ΔlacZ4787 *hsdR*514Δ*araBAD*567

111   Δ*rhaBAD568* rph-1 attλ::pAE12(ΔoriR6K-cat:: Frt5), ΔendA::Frt uidA(ΔMluI)::pir

112   attHK::pJK1006(ΔoriR6K-cat::Frt5; trfA::Frt) ΔdapA::Frt , (17)) was obtained from Dr.

113   Patricia Kiley (University of Wisconsin, Madison). Strain ZM4 Δ*aldB* was constructed

114   using the method described in Lal et al. (16). Briefly, 500-bp fragments directly

115   upstream and downstream of *aldB* were PCR amplified with Q5 DNA Polymerase using

116   the following pairs of primers: "ZMO1754 up" for upstream and "ZMO1754 dn" for

117   downstream fragments (Table 2). The fragments were cloned into the SpeI restriction

118   site of a non-replicating plasmid pPK15534 (16), digested with SpeI, using NEB Builder

119   HiFi DNA Assembly Master Mix to get pPKΔZMO1754 (Fig S1). The insertion was

120   confirmed by PCR. pPKΔZMO1754 was introduced into ZM4 by conjugation from

121   WM6026/ pPKΔZMO1754, and subsequent selection for chloramphenicol resistance

122   and DAP independence, as described in Lal et al. (16). Integration of pPKΔZMO1754

123   into the chromosome by homologous recombination in chloramphenicol resistant cells

124   was confirmed by colony PCR. One colony with inserted plasmid was grown for about

125   10 generations in liquid ZRMG medium without chloramphenicol and 10,000 CFU were

126   plated from an appropriate dilution onto ZRMG plates to get around 100 colonies per

127   plate. Plates were incubated for at least 48 hours in aerobic conditions at 30°C and

128   colonies were screened for loss of green fluorescence under a Blue Light Illuminator.

129   Non-fluorescent colonies were checked for loss of *aldB* by PCR using primers *aldB* up F

130   and *aldB* dn R (Table 2), followed by Sanger sequencing.

131   Plasmids and oligonucleotides are listed in Table 1 and Table 2, respectively.

132   pET16b*aldB* was constructed by amplification of *aldB* from ZM4 genomic DNA using

133   primers MF_17 and MF_18, and Gibson assembly into the pET16b amplified from

134   primers MF_19 and MF_20. This assembly places *aldB* between BamHI and NdeI of

135   pET16b under control of the T7 promoter and downstream of a 10-histidine tag (Figure

136   S2). After transformation into Mach 1, plasmid was isolated from ampicillin resistant

137   colonies and the sequence was confirmed by Sanger sequencing. Next, pET16b*aldB*

138   was transformed to BL21(DE3) pLysS and IPTG-induced overexpression of His-tagged

139   AldB was confirmed by Western blotting as described in His-AldB purification section.

140   pRL*aldB* was constructed by amplification of *aldB* from ZM4 using primers MF_9 and

141   MF_10 and assembled with pRL814 amplified with primers MF_11 and MF_12 using

142   NEB HiFi DNA Assembly Mastermix. This assembly replaces the GFP gene in pRL814

143   with *aldB*. pRLHis-*aldB* was constructed by amplification of His-*aldB* from pET16b*aldB*

144   using primers MF_23 and MF_24 and Gibson assembly with the pRL814 backbone

145   amplified with primers MF_21 and MF_22. After transformation into Mach 1, isolated

146   plasmids were sequenced by the Sanger method.

147   *Bacterial growth*

148   ZM4 or Δ*aldB* were grown in ZRMG or ZMMG at 30°C, with shaking at 250 rpm. For

149   anaerobic conditions bacteria were grown in Hungate tubes: 5 ml of media stored in an

150   anaerobic chamber (Coy Lab Products) were inoculated from single colonies in the

151   chamber and the tubes were closed with butyl-rubber stoppers and secured with

152   aluminum crimps before removing from the chamber. For aerobic growth Hungate tubes

153   were loosely covered with aluminum foil. Samples were removed periodically for OD_{600}

154   measurements and HPLC analysis. Spectinomycin was added to cultures bearing

155   pRL814, pRL*aldB* or pRLHis-*aldB* and IPTG was added as indicated.

156   *HPLC analysis*

157   Samples were prepared by centrifugation at 14,000 rpm for 10 min in microfuge tubes at

158   4°C to remove cells and supernatants were transferred into 2 ml glass vials. Analysis

159   was performed on Shimadzu 20A HPLC using Aminex HPX-87H (BioRad) column at

160   50°C. Separation of compounds was with 5 mM sulfuric acid at flow rate of 0.6

161   ml/minute. Compounds were identified by refractive index and concentrations

162   determined from standard curves for external standards run in the same batch.

163   *Soluble protein fraction from* Z. mobilis

164 50 ml ZRMG in 250-ml flasks were inoculated from fresh colonies of ZM4 or Δ*aldB* and

165 grown aerobically with shaking at 250 rpm at 30°C to OD_{600} 1.0. Cells were harvested

166 by centrifugation at 4°C at 8,000 rpm for 10 minutes (Sorval ST 8R, 75005709 rotor).

167 Pellets were washed twice with 2 ml ice-cold 0.1 M Tris HCl pH 8.0 and resuspended in

168 0.25 ml of the same buffer. Lysis components were added to final concentrations: 20

169 mM EDTA, 1 mg/ml lysozyme, 4 mM DTT and Protease Inhibitor as recommended by

170 manufacturer. Lysis was performed at 37°C for 20 minutes. Samples were spun in

171 microfuge tubes at 17,000 rpm (Sorval ST 8R 75005715 rotor), for 30 minutes at 4°C.

172 Clear supernatant (soluble protein fraction, FI) was gently transferred to new microfuge

173 tubes and flash frozen in liquid nitrogen. FI was stored at -80°C.

174 *Purification of His-AldB by Ni-affinity chromatography*

175 100 ml of LB containing ampicillin was inoculated from a single colony of *E. coli*

176 BL21(DE3) pLysS bearing pET16b*aldB* and grown at 30°C with shaking at 250 rpm

177 overnight. An overnight culture was diluted to $OD_{600} = 0.1$ in two liters of LB media with

178 ampicillin in a 6-L flask and grown with shaking at 30°C to $OD_{600} = 0.6$. At this point,

179 IPTG was added to a final concentration of 0.1 mM and growth was continued for two

180 hours to $OD_{600} = 2.5$. Cells were harvested by centrifugation at 6,000 rpm for 10 minutes

181 (GC-3 Sorval) at 4°C. Pellets were resuspended gently in 20 ml of ice-cold B-PER

182 Reagent containing two tablets of Complete Protease Inhibitor and flash freeze in liquid

183 nitrogen. On the next day, cells were thawed on ice and solid NaCl was added to a final

184 concentration of 0.3 M and 100 mM imidazole to 5 mM. At this point cells were

185 completely lysed. The soluble fraction was separated from membranes and cell debris

186 by ultracentrifugation (Beckman, Ti 45) at 22,000 rpm for 40 minutes. Clear

187   supernatant, containing soluble protein fraction (FI), was analyzed by SDS PAGE for

188   the presence of His-AldB and the total protein concentration was determined using a

189   Bradford assay. Activity of His-AldB was measured by an acetaldehyde dehydrogenase

190   assay as described below. 11 ml of FI containing 14 mg/ml of total protein was mixed

191   with 2.5 ml of Ni-charged IMAC resin equilibrated with IMAC buffer A and incubated on

192   a rocker for 1 hour in 4°C. The entire slurry was loaded onto a 10-ml ThermoFisher

193   disposable column under gravity flow at 4°C. The column was washed with 10 column

194   volumes (CV) of IMAC A and 10 CV of IMAC B under gravity flow. His-AldB was eluted

195   with 5 CV of IMAC C; 0.25-ml fractions were collected. Protein concentration in fractions

196   was determined by the Bradford method and the purity of His-AldB was estimated by

197   SDS PAGE followed by Coomassie staining and Western blotting with anti-His antibody.

198   Dithiotreitol (DTT) was added to a final concentration of 10 mM to the peak fractions

199   before freezing in liquid nitrogen. Fractions were stored at -80°C.

200   *SDS PAGE and Western blotting*

201   The indicated amounts of Ni-chromatography fraction of His-*aldB* or soluble protein

202   fraction was added to gel loading buffer (0.4 M Tris base, 30% glycerol, 1% SDS, 0.5%

203   bromophenol Blue), containing 10 mM DTT to get final volume of 10 µl and boiled for 5

204   minutes. To get whole cell lysate, a pellet from 1 ml of cells at $OD_{600} = 1.0$ was

205   resuspended in 0.1 ml of loading buffer and boiled for 10 minutes. Samples were loaded

206   on 4–20% gradient Mini-PROTEAN TGX stain free gels (BioRad) and run in

207   Tris/Glycine/SDS running buffer (BioRad). Gels were stained with Coomassie Brilliant

208   Blue or proteins were transferred into nitrocellulose membranes in Trans-Blot-Turbo

209   Transfer Buffer (BioRad) for 7 min in a Trans-Blot Turbo Transfer System (BioRad).

210   Membranes were blocked with 3% BSA in TBST and incubated with mouse anti-His

211   monoclonal antibody overnight in a cold room. Membranes were washed in TBST buffer

212   for two hours and incubated with horseradish peroxidase-conjugated rabbit anti-mouse

213   antibody for two hours. The membranes were washed and chemiluminescent detection

214   was performed with Clarity Western ECL (BioRad) for 5 min.

215   *Standard acetaldehyde dehydrogenase assay*

216   Reactions were performed in vitro in 1 ml of reaction buffer in 24-well plate using a

217   Synergy H1 BioTek plate reader. The standard reaction buffer contained 50 mM Tris pH

218   8.0, 10 mM DTT, 22 mM acetaldehyde, 2 mM $NADP^+$ or NAD^+ and 50 mU of purified

219   His-AldB or indicated volumes of soluble protein fraction from *Z. mobilis* or *E. coli*

220   expressing *aldB* from plasmid. The reaction was performed at 30°C for 30 minutes.

221   Acetaldehyde dehydrogenase activity was measured as reduction of $NAD^+/NADP^+$ to

222   NADH/NADPH by monitoring absorbance at 340 nm.

223   Results

224   To determine which gene in the *Z. mobilis* ZM4 genome encodes an ortholog of the

225   previously purified acetaldehyde dehydrogenase from Z6, we searched for a locus with

226   the following characteristics: i) the gene should encode a protein with an approximate

227   size of 55 kDa, ii) the gene should be induced under aerobic conditions (11). Based on

228   a recent multi-omic analysis of oxygen exposure in *Z. mobilis* ZM4 by Martien et al. (18),

229   we hypothesized that locus ZMO1754 (also known as ZMO_RS07890) encodes the

230   previously discovered acetaldehyde dehydrogenase. ZMO1754 encodes a protein with

231   an expected size of 49.6 kDa, and both its transcript and protein abundance increased

232   upon oxygen exposure (18). ZMO1754 is annotated as an NAD^+-dependent succinate-

233   semialdehyde dehydrogenase, suggesting a role as a redox enzyme. Further, an

234   ortholog of ZMO1754 is present in the Z6 genome, indicating that it could encode the

235   previously identified acetaldehyde dehydrogenase. The results below show that the

236   protein encoded by ZMO1754 shares strong functional similarity to the *E. coli* gene

237   *aldB*, which is also an NADP⁺-dependent acetaldehyde dehydrogenase (19). Due to the

238   functional similarity, we propose to rename the locus ZMO1754 to *aldB* in *Z. mobilis* and

239   refer to the gene as *aldB* throughout the manuscript*.*

240   To determine whether *aldB* encodes an acetaldehyde dehydrogenase, we first

241   deleted it from the *Z. mobilis* ZM4 genome and observed the effect on growth in oxic

242   and anoxic conditions in rich and minimal media (**Figure 2**). We observed that the

243   deletion had little impact on growth in anoxic conditions but reduced the final culture

244   density by 19% and 30% in oxic rich and minimal media, respectively. We also

245   measured acetate production by both strains under each condition. The wild-type (WT)

246   strain produced acetate in all culture conditions, while acetate production by Δ*aldB* was

247   undetectable except in oxic rich medium (**Figure 3**). In oxic rich medium, Δ*aldB*

248   produced a small amount of acetate (1.37 mM), but this was a 96% decrease compared

249   with the 30.87 mM produced by WT in the same conditions.

250   We complemented Δ*aldB* by expressing *aldB* from an IPTG inducible plasmid

251   (pRL*aldB*) and measuring growth and acetate production in oxic rich medium (**Figure

252   4**). We used pRL814, which expresses GFP under an IPTG-inducible promoter, as a

253   control for the effect of the plasmid and heterologous protein expression. We observed

254   that expressing GFP slightly reduced growth in both WT and Δ*aldB*, while expressing

255   *aldB* dramatically improved growth for both strains. Similarly, GFP expression had no

256 effect or reduced acetate production, while *aldB* expression increased acetate

257 production by nearly 3-fold for ZM4 and over 16-fold for Δ*aldB*. Overall, we found that

258 expression of *aldB in trans* complemented the growth and acetate production

259 phenotypes of Δ*aldB*. Further, expression *in trans* increased growth and acetate

260 production in WT under the conditions tested. We also performed complementation with

261 His-tagged AldB and observed the same results, indicating that the tagged version of

262 the protein is active (Figure S4).

263   We confirmed that the Δ*aldB* phenotype was caused by loss of acetaldehyde

264 dehydrogenase activity by using an activity assay to compare soluble protein fractions

265 of WT and Δ*aldB*. We tested activity of the soluble protein fraction to convert

266 acetaldehyde to acetate using $NADP^+$ or NAD^+ as the cofactor. We observed that

267 deletion of *aldB* completely abolished the $NADP^+$-dependent activity. In WT, the NAD^+-

268 dependent acetaldehyde dehydrogenase activity was <20% as high as the $NADP^+$-

269 dependent activity. The NAD^+-dependent activity was reduced slightly by deletion of

270 *aldB* (**Figure 5**).

271   We next overexpressed a His-tagged version of AldB in *E. coli* and confirmed

272 acetaldehyde dehydrogenase activity of the tagged protein in the soluble protein fraction

273 from cells induced with IPTG (**Figure 6**). We purified the protein using Ni-affinity

274 chromatography as described in Materials and Methods (Figure S3). We pooled

275 fractions based on their specific activity and used fractions 9-11 to measure substrate

276 specificity using an acetaldehyde dehydrogenase assay as described in Materials and

277 Methods. We observed that the activity was specific for acetaldehyde with moderate

278 activity toward other short-chain fatty aldehydes, propionaldehyde (3C), butyraldehyde

279   (4C), and valeraldehyde (5C) (**Table 3**). The activity toward formaldehyde and

280   glyceraldehyde was low.

281   Because we previously observed some acetate secretion by ZM4 grown in

282   lignocellulosic hydrolysate, we also measured growth and acetate concentration in

283   fermentations using ZM4 or Δ*aldB*. Growth and substrate utilization by the two strains

284   was not significantly different; both strains reached a final OD$_{600}$ of 5.4 after 48 hours

285   and consumed all available glucose. However, the acetate and ethanol concentrations

286   were significantly different between the two strains. The initial acetate concentration in

287   the hydrolysate was 38.48 mM and after 48 hours of fermentation by ZM4, the

288   concentration rose to 41.31±0.06 mM. In contrast, with Δ*aldB*, the acetate concentration

289   decreased slightly to 37.76±0.04 mM. Ethanol production by Δ*aldB* was slightly

290   decreased, at 750.01±0.28 mM versus 752.7±0.21 mM for ZM4.

291   Discussion

292   We determined that the locus ZMO1754 (also known as ZMO_RS07890) in the *Z.*

293   *mobilis* ZM4 genome encodes an acetaldehyde dehydrogenase that is the main source

294   of acetate production. Here, we refer to the gene encoded at that locus as *aldB* based

295   on functional similarity to *E. coli aldB*. Although acetate production by *Z. mobilis* is well-

296   documented and an acetaldehyde dehydrogenase was previously purified from strain

297   Z6, the genes necessary for acetate production were previously unknown. Yang et al.

298   previously speculated that ZMO1754 could encode an acetaldehyde dehydrogenase but

299   did not test the hypothesis because they found no evidence of upregulation in response

300   to oxygen (5). However, a more recent transcriptomic and proteomic characterization of

301   oxygen response in *Z. mobilis* did observe that the transcript and protein encoded by

302   ZMO1754 were upregulated in response to oxygen (18). The newer multi-omic

303   characterization led us to hypothesize that ZMO1754 encodes an acetaldehyde

304   dehydrogenase. Our deletion studies show that ZMO1754 encodes the enzyme

305   responsible for acetate production in *Z. mobilis* ZM4.

306   We also characterized AldB from *Z. mobilis* ZM4 to determine whether it is an

307   ortholog of the NADP$^+$-dependent acetaldehyde dehydrogenase purified from *Z. mobilis*

308   Z6 by Barthel et al. (11). We found that both are aldehyde dehydrogenases with

309   significant specificity toward acetaldehyde, are upregulated in response to oxygen, have

310   approximately the same molecular weight, and use NADP$^+$ rather than NAD$^+$ as a

311   cofactor. Although we observed a minor NAD$^+$-dependent acetaldehyde dehydrogenase

312   activity in the protein fraction extracted from Δ*aldB*, both acetate concentrations in

313   culture and activity assays confirm that *aldB* encodes the major acetaldehyde

314   dehydrogenase.

315   The presence of a growth defect in Δ*aldB* only under oxic conditions suggests

316   that acetate production is beneficial only when oxygen is present. This finding aligns

317   with previous research indicating that acetaldehyde accumulation causes poor growth of

318   *Z. mobilis* when oxygen is present (**Figure 1**) (20, 21). We observed that although there

319   was some acetate production under anoxic conditions, the concentration was >10-fold

320   lower than in oxic conditions, suggesting that there is little acetaldehyde available as a

321   substrate when oxygen is absent. In context with previous work, our findings suggest

322   that the main role of AldB is to detoxify acetaldehyde that is generated under oxic

323   conditions.

324   From a biotechnological perspective, identifying the locus encoding acetaldehyde

325   dehydrogenase is useful because a deletion strain could be used to eliminate acetate

326   production in biofuel fermentations. On the other hand, we observed decreased ethanol

327   production by Δ*aldB*, suggesting that acetate production may be involved in a stress-

328   response pathway in *Z. mobilis* ZM4. We speculate that reducing equivalents may be

329   diverted from central metabolism to detoxification pathways that process toxins found in

330   lignocellulosic hydrolysates. The diversion of reducing equivalents would necessitate

331   acetaldehyde detoxification as observed in oxic conditions. Further, *Z. mobilis* has been

332   proposed as a host for acetaldehyde production and deletion of *aldB* is likely to

333   dramatically increase acetaldehyde yield (22). (Note that we do not report acetaldehyde

334   concentrations here because our culture method resulted in significant evaporative loss

335   of acetaldehyde.) Further work is necessary to determine whether deletion or

336   upregulation of *aldB* from *Z. mobilis* would be beneficial for biotechnological processes

337   overall.

338   We expect our findings to be generalizable across *Zymomonas* because the

339   acetaldehyde dehydrogenase encoded by *aldB* in strain ZM4 appears to be conserved

340   across the entire *Zymomonas* genus. We recently conducted a genome comparison

341   analysis for all sequenced *Zymomonas* genomes and used our previous ortholog

342   analysis to determine that an *aldB* ortholog is present in all analyzed genomes. The

343   *aldB* gene shares 99.1% identity across *Z. mobilis* subspecies *mobilis* and 92% and

344   84% between *francensis* and *mobilis* and *pomaceae* and *mobilis*, respectively (7).

345   Acknowledgements

351   References

352   1.   Yang S, Fei Q, Zhang Y, Contreras LM, Utturkar SM, Brown SD, Himmel ME,
353   Zhang M. 2016. *Zymomonas mobilis* as a model system for production of biofuels
354   and biochemicals. Microb Biotechnol2016/09/15. 9:699–717.

355   2.   Rogers PL, Lee KJ, Tribe DE. 1979. Kinetics of alcohol production by
356   *Zymomonas mobilis* at high sugar concentrations. Biotechnol Lett 1:165–170.

357   3.   Kremer TA, LaSarre B, Posto AL, McKinlay JB. 2015. N_2 gas is an effective
358   fertilizer for bioethanol production by *Zymomonas mobilis*. Proc Natl Acad Sci U S
359   A 112:2222–6.

360   4.   Bringer S, Finn RK, Sahm H. 1984. Effect of oxygen on the metabolism of
361   *Zymomonas mobilis*. Arch Microbiol 139:376–381.

362   5.   Yang S, Tschaplinski TJ, Engle NL, Carroll SL, Martin SL, Davison BH, Palumbo
363   A V, Rodriguez M, Brown SD. 2009. Transcriptomic and metabolomic profiling of
364   *Zymomonas mobilis* during aerobic and anaerobic fermentations. BMC Genomics
365   10:34.

366   6.   Swings J, De Ley J. 1977. The biology of *Zymomonas*. Bacteriol Rev 41:1–46.

367   7.   Felczak MM, Bowers RM, Woyke T, TerAvest MA. 2021. Zymomonas diversity
368   and potential for biofuel production. Biotechnol Biofuels 14:112.

369   8.   Franden MA, Pilath HM, Mohagheghi A, Pienkos PT, Zhang M. 2013. Inhibition of
370   growth of *Zymomonas mobilis* by model compounds found in lignocellulosic
371   hydrolysates. Biotechnol Biofuels 6.

372   9.   Yang S, Franden MA, Brown SD, Chou YC, Pienkos PT, Zhang M. 2014. Insights
373   into acetate toxicity in *Zymomonas mobilis* 8b using different substrates.
374   Biotechnol Biofuels 7.

375   10.   Ong WK, Courtney DK, Pan S, Andrade RB, Kiley PJ, Pfleger BF, Reed JL. 2020.
376   Model-driven analysis of mutant fitness experiments improves genome-scale
377   metabolic models of *Zymomonas mobilis* ZM4. PLOS Comput Biol 16:e1008137.

378   11.   Barthel T, Jonas R, Sahm H. 1989. $NADP^+$-dependent acetaldehyde
379   dehydrogenase from *Zymomonas mobilis*. Arch Microbiol 153:95–100.

380   12.   Desiniotis A, Kouvelis VN, Davenport K, Bruce D, Detter C, Tapia R, Han C,
381   Goodwin LA, Woyke T, Kyrpides NC, Typas MA, Pappas KM. 2012. Complete
382   genome sequence of the ethanol-producing *Zymomonas mobilis* subsp. mobilis
383   centrotype ATCC 29191. J Bacteriol 194:5966–5967.

384   13.   Lee KY, Park JM, Kim TY, Yun H, Lee SY. 2010. The genome-scale metabolic
385   network analysis of *Zymomonas mobilis* ZM4 explains physiological features and
386   suggests ethanol and succinic acid production strategies. Microb Cell Fact 9:94.

387   14.   Widiastuti H, Kim JY, Selvarasu S, Karimi IA, Kim H, Seo JS, Lee DY. 2011.
388   Genome-scale modeling and in silico analysis of ethanologenic bacteria
389   *Zymomonas mobilis*. Biotechnol Bioeng 108.

390   15.   Nouri H, Fouladiha H, Moghimi H, Marashi S-A. 2020. A reconciliation of genome-
391   scale metabolic network model of *Zymomonas mobilis* ZM4. Sci Rep 10:7782.

392   16.   Lal PB, Wells FM, Lyu Y, Ghosh IN, Landick R, Kiley PJ. 2019. A markerless
393   method for genome engineering in *Zymomonas mobilis* ZM4. Front Microbiol.

394   17.   Blodgett JA V, Thomas PM, Li G, Velasquez JE, van der Donk WA, Kelleher NL,
395   Metcalf WW. 2007. Unusual transformations in the biosynthesis of the antibiotic
396   phosphinothricin tripeptide. Nat Chem Biol2007/07/15. 3:480–485.

397   18.   Martien JI, Hebert AS, Stevenson DM, Regner MR, Khana DB, Coon JJ, Amador-
398   Noguez D. 2019. Systems-level analysis of oxygen exposure in *Zymomonas
399   mobilis*: implications for isoprenoid production. mSystems 4:e00284-18.

400   19.   Ki HK, Henry W. 2005. Isolation and characterization of an aldehyde
401   dehydrogenase encoded by the *aldB* gene of *Escherichia coli*. J Bacteriol
402   187:1067–1073.

403   20.   Kalnenieks U, Galinina N, Strazdina I, Kravale Z, Pickford JL, Rutkis R, Poole RK.
404   2008. NADH dehydrogenase deficiency results in low respiration rate and
405   improved aerobic growth of *Zymomonas mobilis*. Microbiology 154:989–994.

406   21.   Kalnenieks U, Galinina N, Toma MM, Poole RK. 2000. Cyanide inhibits respiration
407   yet stimulates aerobic growth of *Zymomonas mobilis*. Microbiology 146:1259–66.

408   22.   Kalnenieks U, Balodite E, Strähler S, Strazdina I, Rex J, Pentjuss A, Fuchino K,
409   Bruheim P, Rutkis R, Pappas KM, Poole RK, Sawodny O, Bettenbrock K. 2019.
410   Improvement of acetaldehyde production in *Zymomonas mobilis* by engineering of
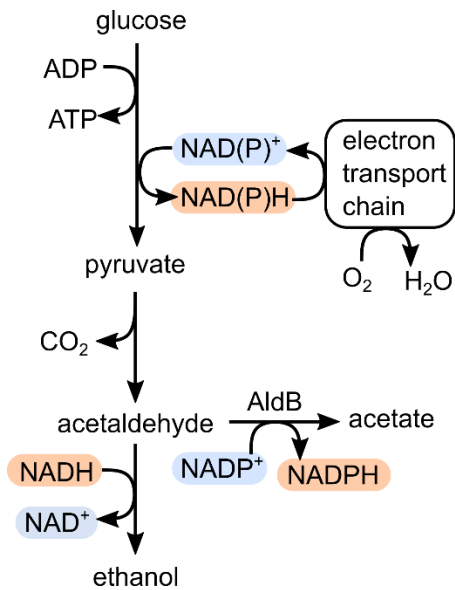411   its aerobic metabolism. Front Microbiol 10:2533.

412   Figures

413

Figure 1. Central metabolism of *Z. mobilis,* including the acetaldehyde

dehydrogenase encoded by *aldB*. Oxidized redox cofactors are highlighted in blue

and reduced redox cofactors are highlighted in red to indicate the connections formed

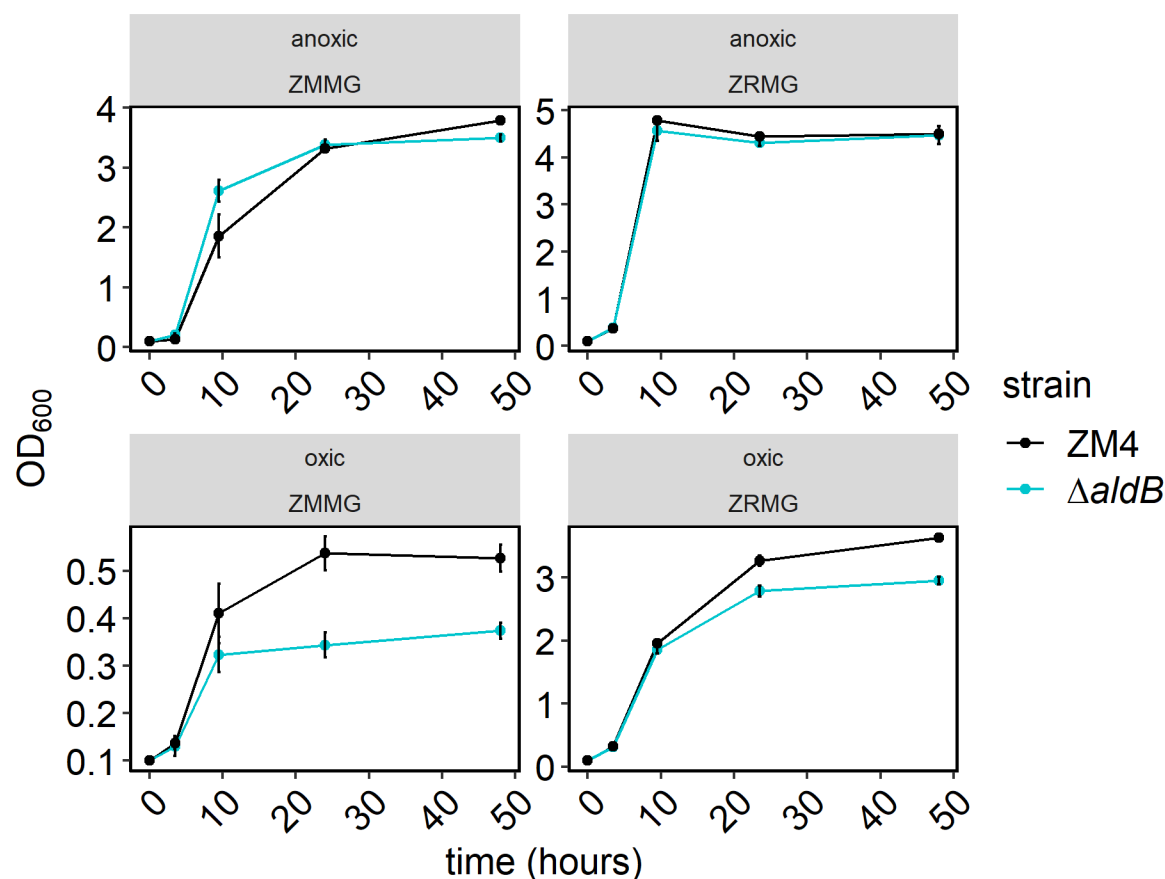between oxygen and ethanol, acetaldehyde, and acetate formation.

418

Figure 2. Growth of *Z. mobilis* ZM4 WT and Δ*aldB*. Overnight cultures were inoculated from single colonies and grown overnight in rich or minimal media in static cultures in oxic or anoxic conditions. For subsequent anoxic growth, the cultures were diluted to OD_{600} of 0.1 in 5 ml of fresh rich or minimal media in Hungate tubes in an anaerobic chamber and tubes were closed with stoppers and secured with crimps. For oxic growth cultures were started similarly but Hungate tubes were covered with aluminum foil. All cultures were grown outside anaerobic chamber in 30°C with shaking at 250 rpm for time indicated. Samples were taken periodically to monitor growth. Points represent the average of 3 biological replicates and error bars represent standard error. Note that y-axis scales differ across panels.
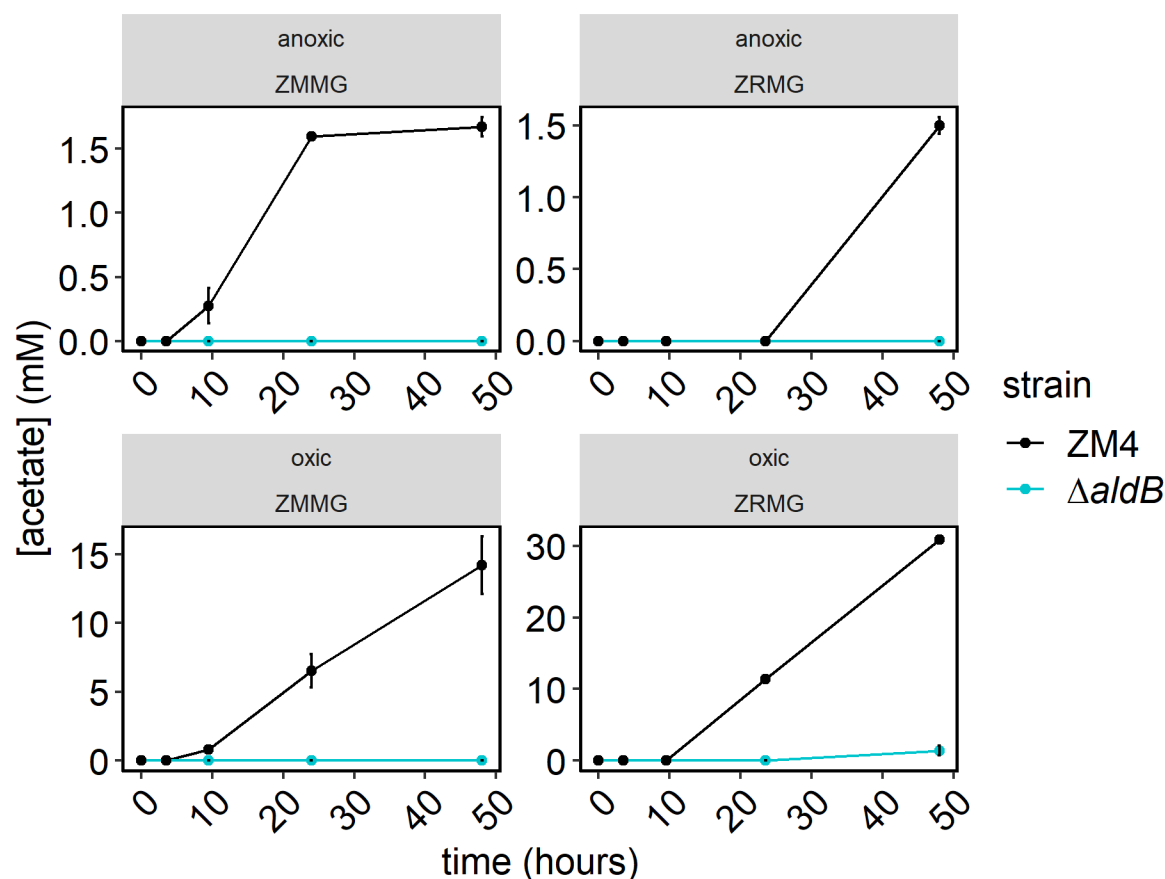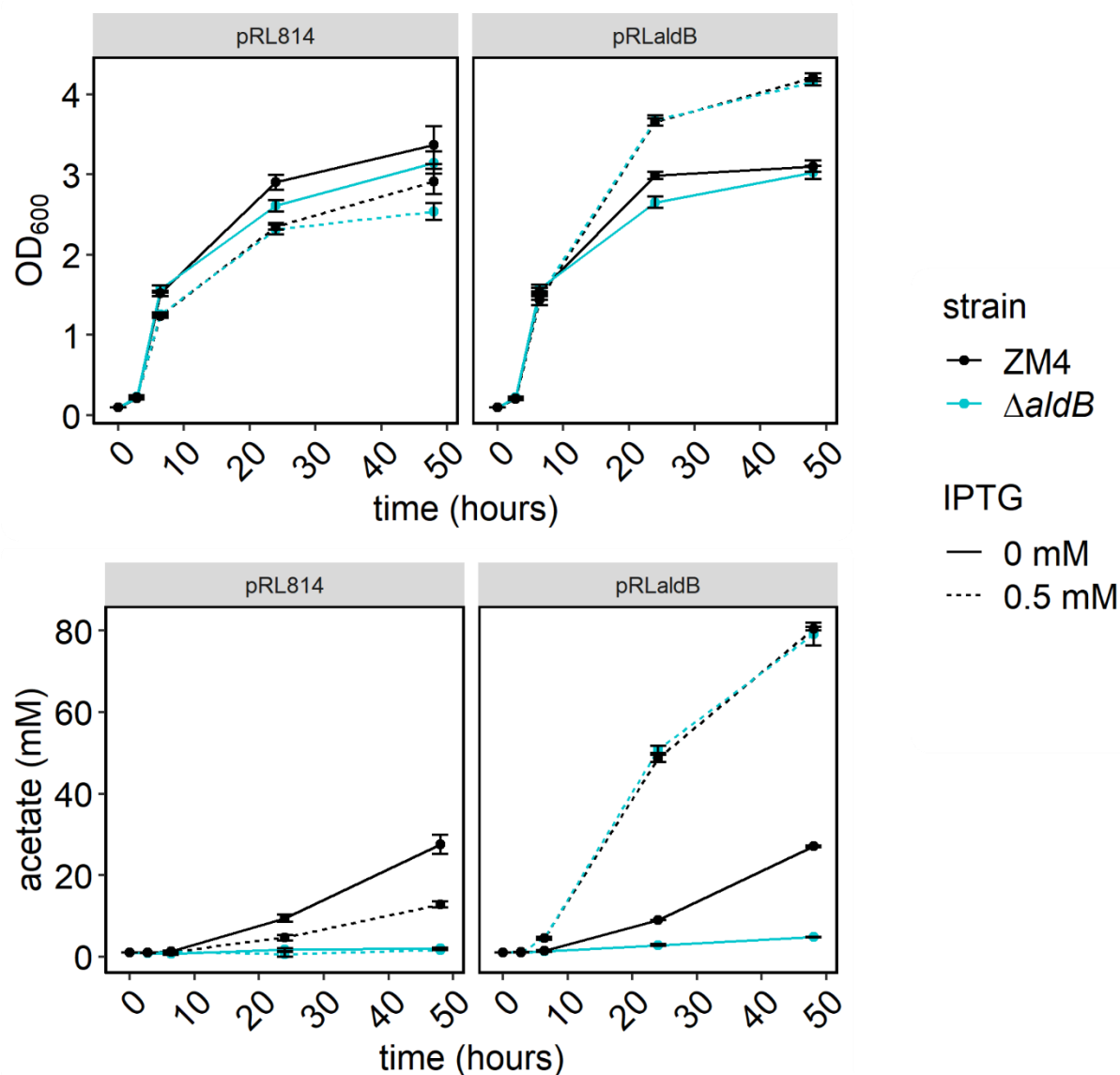
Figure 3. Acetate production by _Z. mobilis_ ZM4 WT and Δ_aldB_. Samples from cultures used in Figure 1 were centrifuged and clear supernatants were analyzed by HPLC as described in Materials and Methods. Acetate concentrations were calculated from standard curves. Points represent the average of 3 biological replicates and error bars represent standard error. Note that y-axis scales differ across panels.

437

Figure 4. Growth and acetate production by _Z. mobilis_ ZM4 WT and Δ_aldB_

expressing _aldB_ from inducible plasmid. Strains bearing pRL_aldB_ or pRL814, were

grown as in **Figure 1** but media were supplemented with 100 μg/ml of spectinomycin

and IPTG was added to 0.5 mM, at time of dilution, when indicated. HPLC analysis was

performed as in **Figure 2**. Points represent the average of three biological replicates

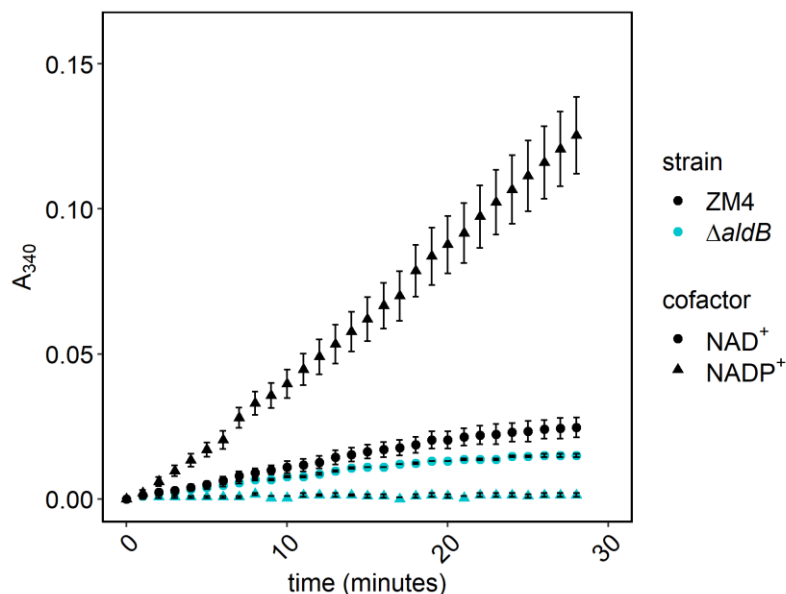and error bars represent standard error.

444

Figure 5. Acetaldehyde dehydrogenase activity in soluble protein fractions from ZM4 WT and ΔaldB. Soluble protein fraction (FI) was obtained as described in "Materials and Methods". Average, total protein concentration in FI was 5.7 +/- 0.5 mg/ml and 5.2 +/- 0.7 mg/ml for ZM4 and ΔaldB, respectively. Each enzymatic reaction contained: 0.1 M Tris HCl pH 8.0, 0.1 M KCl, 10 mM β-mercaptoethanol, 2 mM acetaldehyde and 0.67 mM NAD or NADP (protocol for yeast acetaldehyde dehydrogenase from Sigma-Aldrich). Reaction was started by adding 33 μl of FI and measured for 30 minutes at 25°C in 24-well microtiter plate as described in "Materials and Methods". Absorbance at 340 nm in control without FI was subtracted from the reactions. Points represent the average of three independent experiments with three technical repeats and error bars represent standard error.
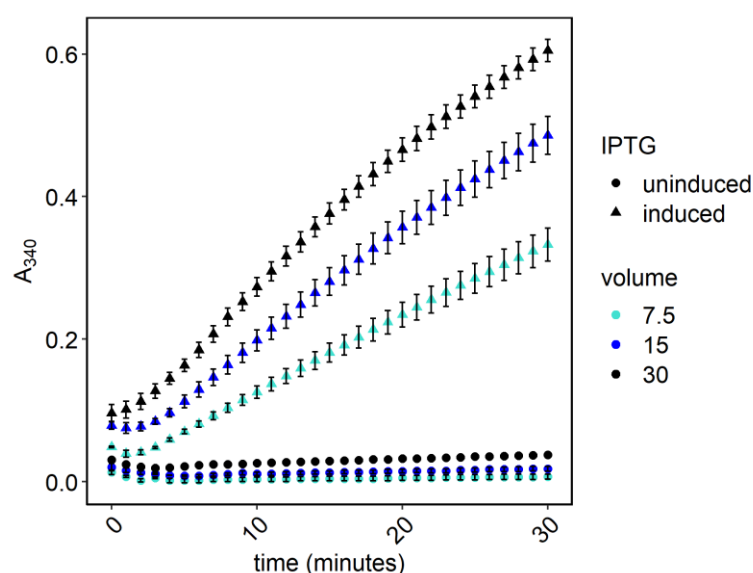
457

**Figure 6. Acetaldehyde dehydrogenase activity in soluble protein fraction from *E.*

coli BL21 carrying a plasmid expressing *aldB*.** 50-mL cultures were grown in LB

supplemented with ampicillin at 30ºC to OD_{600} 0.4. At this point, IPTG was added to a

final concentration of 0.1 mM, where indicated, and growth continued for 1.5 hours. FI

was obtained as described in "Materials and Methods" for *Z. mobilis*. Average total

protein concentration in three independent FI was 11±1.2 mg/ml and 8.4±1.1 mg/ml for

uninduced and induced cultures, respectively. An enzymatic assay was performed as

described in **Figure 4**, with NADP as a cofactor. The reaction was started by adding

variable volumes of FI as indicated. Points represent the average of three independent

FI and error bars represent standard error.

468

469

470 Tables

471 **Table 1. Plasmids**

Plasmid	Parent	Description	Relevant genotype	reference
pPK15534		Suicide vector	GFP, Cm[R], SpeI site	P. Lal
pPKΔZMO1754	pPK15534	ZMO1754 (*aldB*) knockout plasmid	GFP, Cm[R], 1000 bp insert up/down *aldB*	This work
pRL814		broad host range plasmid	T7 lac GFP, Sp[R]	I. Gosh
pRL*aldB*	pRL814	complementing plasmid	T7 lac *aldB*, Sp[R]	This work
pRLHis-*aldB*	pRL814	complementing plasmid	T7 lac His-*aldB*, Sp[R]	This work
pET16b		*E. coli* expression vector	His-tag, Amp[R]	Novagen
pET16b*aldB*	pET16b	*aldB* expression in *E. coli*	His-*aldB*, Amp[R]	This work

472

473 **Table 2. Oligonucleotides**

Name	Sequence	Description
ZMO1754 up F	TCATGTTTGACAGCTTATCATTCTGGTTCACAGCGTCC	F[1] primer to amplify 500 bp upstream of *aldB*
ZMO1754 up R	AGGCTCTTTTAAATTTCTTAAAATCCGTTTTAAGCG	R[1] primer to amplify 500 bp upstream of *aldB*
ZMO1754 dn F	TAAGAAATTTAAAAGAGCCTACCTTTCTTATTACTTCAGCTTAGGAC	F primer to amplify 500 bp downstream of *aldB*
ZMO1754 dn R	CAAGGCAAGACCGAGCGCTAGGTGTGGACGGGCGGTGA	R primer to amplify 500 bp downstream of *aldB*
MF_9	GGAGATATACATATGGCATATGAATCTGTCAATCCCG	F primer; amplifies *aldB* with overlaps to pRL814
MF_10	ATATCAAGCTTAGGCAGCAGAAGCGTCTATCAA	R primer; amplifies *aldB* with overlaps to pRL814
MF_11	GCTGCCTAAGCTTGATATCGAATTCCTGCAGC	F primer; amplifies pRL814 with overlaps to *aldB*
MF_12	TTCATATGCCATATGTATATCTCCTTCTTAAAGTTAAACTAATTCTAGATGTG	R primer; amplifies pRL814 with overlaps to *aldB*
MF_17	AGGTCGTCATATGGCATATGAATCTGTCAATCCCG	F primer; amplifies *aldB*[1] with overlaps to pET16B
MF_18	CCGGATCCTTAGGCAGCAGAAGCGTCTATCAATT	R primer; amplifies *aldB*[1] with overlaps to pET16B
MF_19	ATATGCCATATGACGACCTTCGATATGGCC	F primer; amplifies pET16b with overlaps for *aldB*
MF_20	CTGCCTAAGGATCCGGCTGCTAACAAAG	R primer; amplifies pET16b with overlaps for *aldB*
MF_23	ATATCAAGCTTAGGCAGCAGAAGCGTCTATCA	F primer; amplifies His-*aldB* on pET16b*aldB*
MF_24	AGATATACATATGGGCCATCATCATCATCATCATCATCATCAC	R primer; amplifies His-*aldB* on pET16b*aldB*
MF_22	GCTGCCTAAGCTTGATATCGAATTCCTGCAGCC	F primer; amplifies pRL814 for GA[1] of His-AldB

| MF_21 | TGATGGCCCATATGTATATCTCCTTCTTAAAGT TAAACTAATTCTAGATGTGT | R primer; amplifies pRL814 for GA[1] of His-AldB |

474   [1.]   Abbreviations: F, forward; R, reverse; GA, Gibson Assembly

475

476 **Table 3. Substrate specificity of ZM4 acetaldehyde dehydrogenase.** Enzymatic

477 reactions were performed in standard assay conditions (22 mM aldehyde, 2 mM

478 NADP$^+$, 10 mM DTT, 50 mM Tris HCl pH 8.0 and 50 mU of enzyme) at 30°C as

479 described in Materials and Methods for 10 minutes. Aldehydes were added as 0.23 M

480 solutions in DMSO. Activity was normalized by setting activity with acetaldehyde at

481 100%. Data are averages of 5 independent experiments with three technical replicates

482 for each substrate.

Substrate	Relative Activity
acetaldehyde	100
propionaldehyde	64 ± 15
butyraldehyde	83 ± 10
valeraldehyde	47 ± 10
formaldehyde	4 ± 3
glyceraldehyde	8 ± 3

483

484