WILEY-VCH

Exploring the Synthetic Potential of Dihydroxyacetone-Aldolases from Acidophilic Organisms

Léo Paulat,[a] Cédric Gastaldi,[a] Lionel Nauton,[a] Mariline Theveniot,[a] Muriel Joly,[a] Jean-Louis Petit,[b] Véronique de Berardinis,[b] Virgil Hélaine*[a] and Christine Guérard-Hélaine*[a]

[a] Université Clermont Auvergne, Clermont Auvergne INP, CNRS, Institut de Chimie de Clermont-Ferrand, 63000 Clermont-Ferrand, France ; e-mail: christine.helaine@uca.fr
[b] Génomique Métabolique, Génoscope, Institut François Jacob, CEA, CNRS, Univ Evry, Université Paris-Saclay, Evry-Courcouronnes, 91057 Evry, France.

Material

D-fructose-6-phosphate dipotassium salt, α-glycerophosphate dehydrogenase, D-arabinose-5-phosphate, alcohol dehydrogenase from *Saccharomyces cerevisiae*, Trizma hydrochloride, sodium phosphate monobasic dihydrate, hydroxyacetone, 1,3-dihydroxyacetone dimer, glycolaldehyde dimer, 3,3-diethoxy-1-propanol, acetaldehyde, 5-hydroxypentan-2-one and chloroacetaldehyde 55% in water were purchased from Sigma-Aldrich whereas β-nicotinamide adenine dinucleotide reduced disodium salt hydrate (NADH) and glycylglycine were purchased from Alfa Aesar. D-arabinose was purchased from Fluka and acetone from Carlo Erba. The Coomassie protein assay reagent, used for Bradford tests, was purchased from Thermo Scientific. Protein desalting columns were purchased from GE Healthcare and Ni-NTA-agarose was purchased from Qiagen. VWR centrifugal filters were purchased from VWR North America Cat. Glycerol dehydrogenase was obtained as described by A.K. Samland.[1]

Nuclear magnetic resonance (NMR) spectra were recorded using D_2O as solvent on a Bruker AC-400 spectrometer, operating at 400 MHz for ^{1}H and 100 MHz for ^{13}C. Chemical shifts (δ) are reported in ppm relative to TMS signal. Coupling constant values (*J*) are given in Hertz. Electrospray ionization mass spectra (ESI-MS) were recorded on a micro Orbitrap Q-Exactive (70000 V) and high-resolution mass spectra (HR-MS) were recorded on the same instrument with an internal lock mass (H_3PO_4) and an external lock mass (Leu-enkephalin).

Methods

1. Database exploration

Uniprot ID	Organism
P78055	*Escherichia coli (strain K12)*
P32669	*Escherichia coli (strain K12)*
A0A0E4G3U3	*Listeria monocytogenes serovar 1/2a (strain ATCC BAA-679 / EGD-e)*
A0A0D6J3Z8	*Streptococcus pneumoniae*
Q8DVJ4	*Streptococcus mutans serotype c (strain ATCC 700610 / UA159)*
A0A0D6H018	*Streptococcus pyogenes*
A0A0E4C363	*Streptococcus gordonii*
Q8E738	*Streptococcus agalactiae serotype III (strain NEM316)*
A0A0E4C393	*Streptococcus suis*

Table S1. Reference set composed of already identified FSA

	A0A399XY01 (FSA_{AciBac})	A0A0E4G3U3	A0A0Q0RVA3 (FSA_{AcAeo})	P78055	P32669	Q8DVJ4	Q8E738	A0A0D6J3Z8	A0A0E4C393	A0A0D6H018	A0A0E4C363
A0A399XY01 (FSA_{AciBac})	100	25	24	27	31	25	22	24	24	25	25
A0A0E4G3U3	25	100	31	30	32	34	32	35	34	32	34
A0A0Q0RVA3 (FSA_{AcAeo})	24	31	100	27	28	29	28	30	32	31	32
P78055 (FSA_{Ecoli})	27	30	27	100	70	38	40	38	37	39	36
P32669	31	32	28	70	100	38	37	35	37	35	34
Q8DVJ4	25	34	29	38	38	100	64	69	66	69	69
Q8E738	22	32	28	40	37	64	100	69	70	71	72
A0A0D6J3Z8	24	35	30	38	35	69	69	100	70	73	75
A0A0E4C393	24	34	32	37	37	66	70	70	100	72	76
A0A0D6H018	25	32	31	39	35	69	71	73	72	100	77
A0A0E4C363	25	34	32	36	34	69	72	75	76	77	100

Figure S1. Matrix of identity percentage between FSA from the reference set and the two FSA candidates from acidophilic bacteria (written in blue)

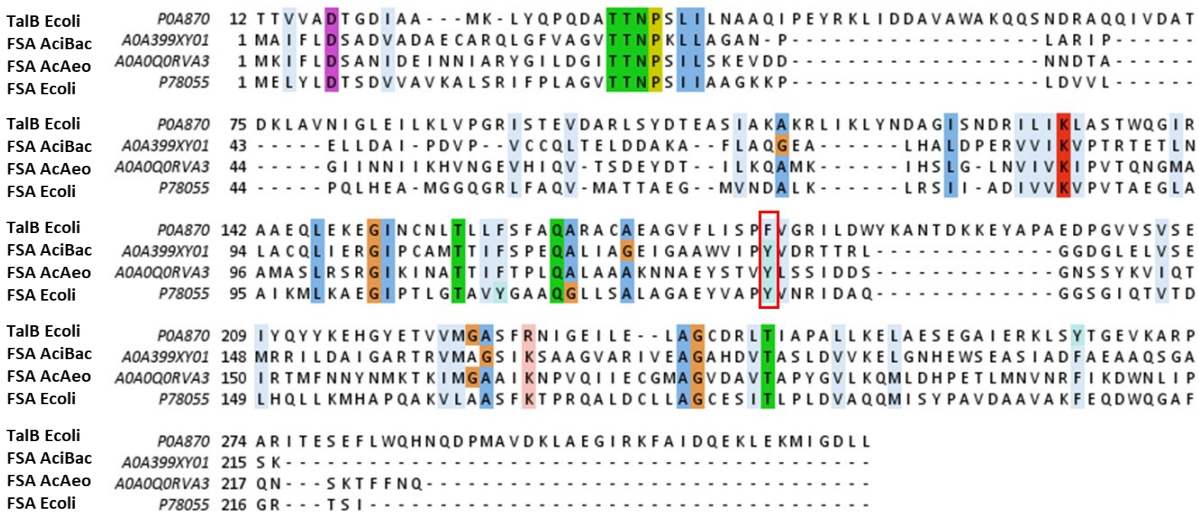


Figure S2. Alignment of the amino acid sequences of *E. coli* FSAA, *E. coli* transaldolase and the two FSA candidates from acid-loving organisms.

2. Purification and biochemical characterization of the enzymes

2.1. Cloning, expression and purification of the recombinant enzymes

The genes were synthetized and optimized for expression in *E. coli* by Twist bioscience. The primers for amplification (Table S2) were designed for each gene in order to incorporate a 6-histidines tag in C-terminal position of the proteins. Genes were cloned in the pET22b(+) vector (Novagen), modified for ligation-independent cloning as described.[2] The recombinant plasmid was transformed in a BL21-codonPlus (DE3)-RIPL competent cell which have been deleted of its triosephosphate isomerase gene.[3] 500 mL of TB medium complemented with 500 mM sorbitol, 5 mM betaine and 100 µg/mL carbenicillin were inoculated with an overnight pre-culture. After growing to an O.D. of 1.6 the expression was inducted with 0.5 mM of IPTG and cells were grown overnight at 20°C. After centrifugation, the pellet was lysed with Bug buster reagent in 32 mL of buffer (50 mM Phosphate (Na/K) pH 8, 500 mM NaCl, 30 mM imidazole, 15% glycerol, 1 mM DTT, lysonase and pefabloc). The proteins were purified on a chromatography system (Akta pure Cytiva). First, the lysate was loaded on an affinity column (His-trap FF 5 ml) and eluted with buffer (50 mM Phosphate (Na/K) pH 8, 500 mM NaCl, 250 mM imidazole, 15% glycérol and 1 mM DTT). Then the eluted peak was loaded on a Gel filtration column (Hi load 16/600 Superdex 200) and collected in buffer (50 mM NaCl, 50 mM Tris pH 7,5, glycerol 10% and 1 mM DTT). The concentration of protein was estimated by Bradford assay with a BSA curve as standard. Purity was controlled by migration on an SDS-PAGE.

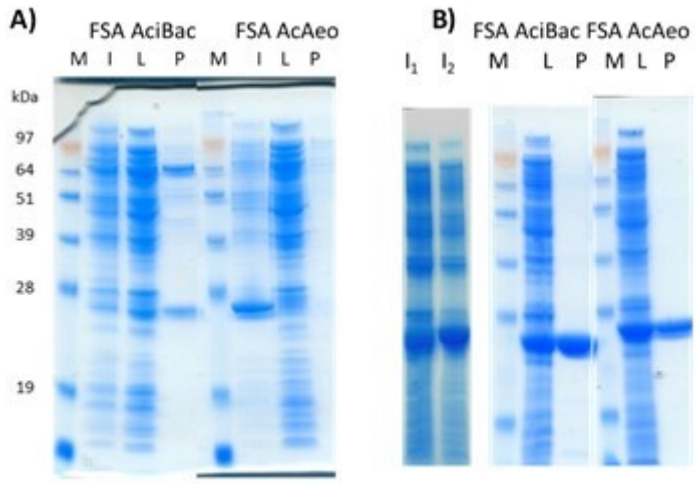


Figure S3 : SDS-gels of A) N-terminally tagged or B) C-terminally tagged studied proteins throughout the various stages of production and purification of the 2 enzymes. I: induction culture (I1: related to FSA AciBac, I2: to FSA AcAeo), L: clear lysate (obtained after sonication and centrifugation), P: purified lysate (obtained after purification by IMAC technique) and M: Molecular mass marker.

	Primer-5Prim	Primer-3Prim	optimized sequence
FSA AciBac	AAAGAA GGAGAT AGGATC ATGGCC ATCTTTC TTGACT CAGCC	GTGTAA TGGATA GTGATC TTAATG GTGATG GTGATG ATGCTT TGAAGC ACCGGA TTGGG	ATGGCCATCTTTCTTGACTCAGCCGATGTTGCGGATGCGGAATGTGCACGTCAGCTGGGTTTCGT GGCTGGTGTTACGACCAACCCCAAACTGCTGGCGGGAGCGAATCCTTTGGCACGTATTCCGGAAC TGCTGGATGCCATTCCTGACGTTCCAGTATGTTGTCAATTAACAGAGTTGGATGATGCGAAAGCG TTTCTCGCACAAGGAGAGGCGCTGCACGCCTTAGACCCAGAACGCGTGGTTATTAAAGTTCCGAC GCGCACCGAAACCCTTAACTTAGCGTGCCAGCTGATTGAACGCGGTATCCCATGTGCCATGACCA CTATTTTCTCTCCGGAGCAGGCATTAATCGCAGGCGAAATTGGCGCGGCCTGGGTTATTCCGTAC GTGGATCGTACTACGCGCCTCGGCGGTGATGGCCTGGAACTGGTATCTGAAATGCGCCGCATTTT GGACGCCATTGGTGCGCGCACGCGTGTGATGGCGGGGTCCATTAAAAGCGCTGCCGGTGTGGCAC GCATCGTGGAAGCGGGTGCACATGACGTAACCGCCTCGCTGGATGTTGTGAAAGAATTAGGTAAC CACGAATGGAGCGAGGCGAGCATCGCTGATTTCGCGGAAGCCGCCCAATCCGGTGCTTCAAAGTAA
FSA AcAeo	AAAGAA GGAGAT AGGATC ATGAAG ATTTTCC TTGACT CGGCG	GTGTAA TGGATA GTGATC TTAATG GTGATG GTGATG ATGCTG ATTGAA GAACGT CTTC	ATGAAGATTTTCCTTGACTCGGCGAACATCGACGAGATTAACAACATCGCGCGCTATGGTATCTTGG ACGGTATCACTACTAACCCAAGCATTCTGTCCAAAGAGGTTGACGATAACAACGACACGGCAGGCA TTATCAATAACATCATTAAGCACGTAAACGGAGAAGTTCATATTCAAGTCACGAGTGACGAGTATGA TACGATCCTGAAGCAAGCAATGAAGATTCACTCGCTCGGCTTGAACGTCATCGTCAAAATCCCGGTA ACCCAGAACGGTATGGCAGCCATGGCGAGCCTTCGCTCGCGCGGTATCAAGATCAATGCGACAACT ATTTTCACGCCATTACAAGCCCTGGCGGCGGCGAAGAACAACGCGGAGTACAGCACCGTGTACCTG AGTTCCATTGACGACAGTGGGAATAGCAGCTACAAAGTGATCCAGACCATCCGCACTATGTTCAACA ACTACAACATGAAGACTAAGATTATGGGTGCGGCCATCAAGAACCCGGTCCAAATCATTGAATGTG GCATGGCGGGCGTGGATGCAGTGACCGCTCCGTACGGGGTGCTCAAACAAATGCTTGATCATCCCGA GACTCTTATGAACGTGAATCGCTTCATTAAAGATTGGAACCTCATCCCGCAAAATTCGAAGACGTTC TTCAATCAGTAA

Table S2. Primers and optimized sequences for FSA_{AciBac} and FSA_{AcAeo}

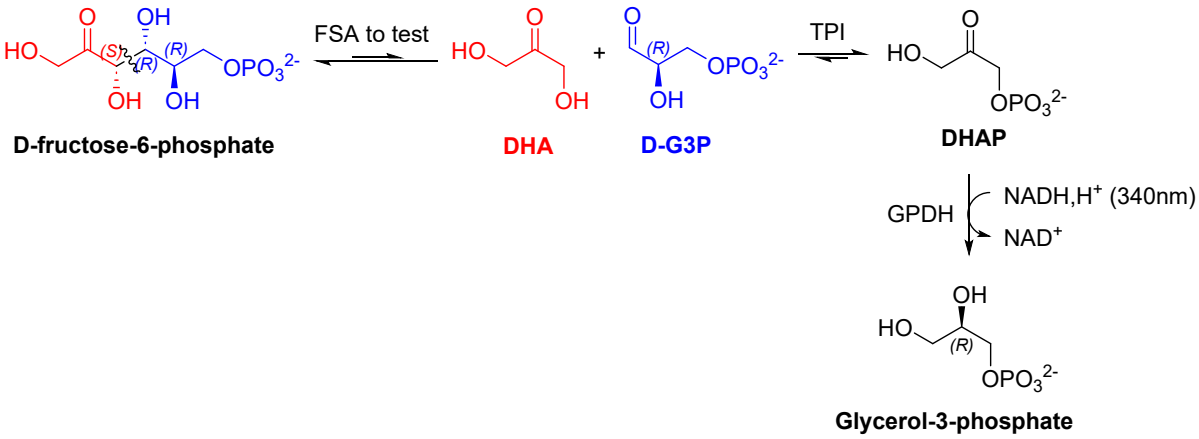
Protocol for 1 L culture*: E. coli* colonies (expression strain) containing the plasmids were cultured in 1 L of Luria-Bertani (LB) broth in the presence of a selection antibiotic (ampicillin) at 37ºC with shaking. When the culture reached an OD_{600nm} of 0.5, protein expression was induced with IPTG (0.5 mM) and the temperature was dropped to 30ºC. The culture was incubated for a further period of 12 h. Cells were harvested by centrifugation, washed twice and resuspended in buffer A (50 mM NaH_2PO_4, 300 mM NaCl, pH 8.0). Cell

suspension was disrupted by sonication and the cell lysate was centrifuged at 10000 x g for 20 min. Clear supernatant was loaded onto a Ni^{2+}-NTA-agarose column (Qiagen, h=3 cm; Ø=2.5 cm) pre-equilibrated with buffer A containing 10 mM of imidazole (buffer B). The column was washed with buffer B, and the retained proteins were eluted with the same buffer containing 500 mM of imidazole. Eluted fraction containing pure protein was dialyzed against water for desalting and imidazole removal before the lyophilization process. For FSA_{AciBac} 373 mg of enzyme were obtained. For FSA_{AcAeo} during the dialysis step, a 3M ammonium sulfate solution was added. The protein concentration was then determined by Bradford, resulting in 290 mg of enzyme obtained as a 29 mg/mL solution.

2.2. Activity assays

Specific activity towards D-fructose-6-phosphate :

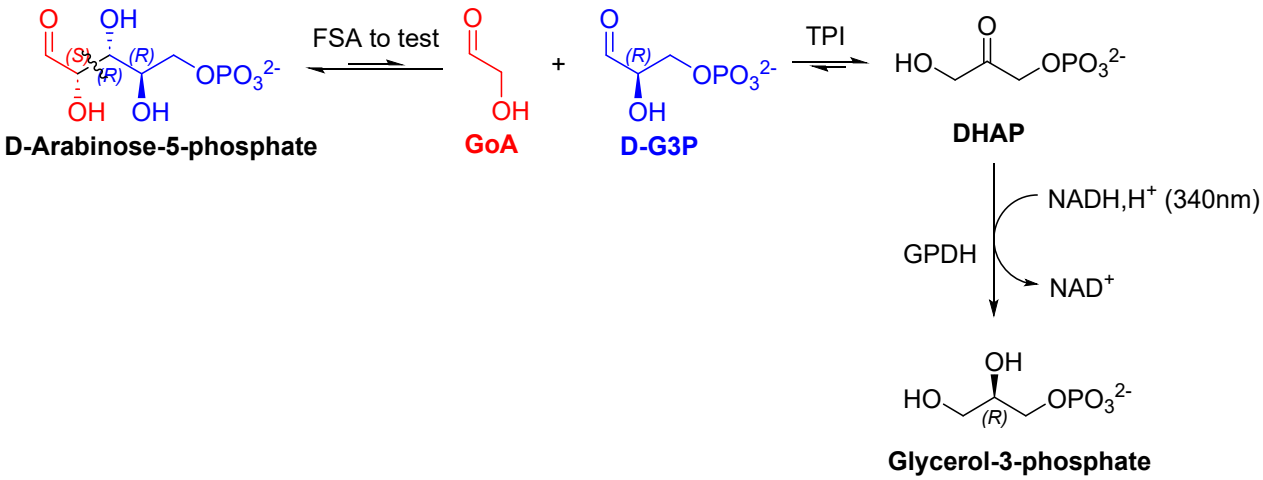
To a solution of D-fructose-6-phosphate (10 mM), DHA aldolase (96 µg of FSA_{AciBac} or 47 µg of FSA_{AcAeo}) in 50 mM glycyl-glycine buffer pH 7.5, was added at 25°C. NADH (0.5 mM), the auxiliary enzymes triose-phosphate isomerase (TPI) and glycerol-3-phosphate dehydrogenase (GPDH) (10 U and 1 U respectively) were added to isomerise D-glyceraldehyde-3-phosphate formed and reduce subsequent dihydroxyacetone phosphate (DHAP). The final volume was 1 mL. The reaction was monitored by spectrophotometry at 340 nm following the consumption of NADH. One mmol of NADH oxidized was equivalent to 1 mmol of D-fructose-6-phosphate cleaved. One unit (U) of DHA aldolase is defined as the amount of enzyme able to cleave 1 µmol of D-fructose-6-phosphate to afford D-glyceraldehyde-3-phosphate and dihydroxyacetone per minute. A specific activity of 11 mU/mg for FSA_{AciBac} and 151 mU/mg for FSA_{AcAeo} were found.



Specific activity towards D-arabinose-5-phosphate:

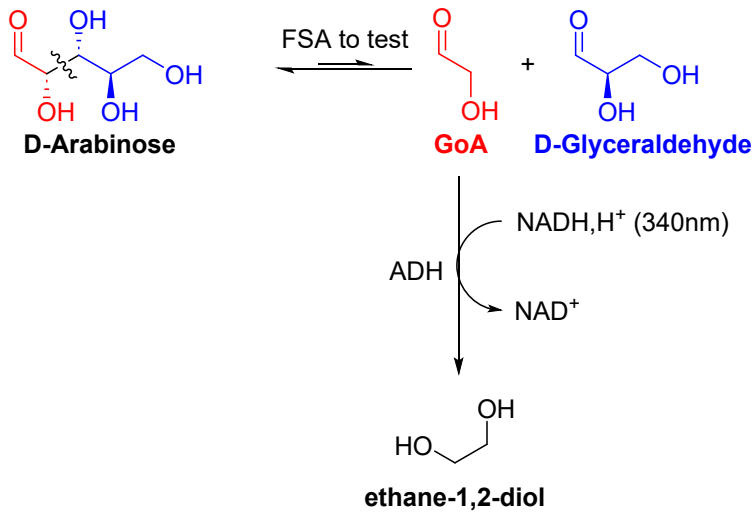
The spectrophotometric assays were recorded on a Safas UVMC2 (Safas, Monaco) using microcells high-precision cell quartz with 10 mm light path (Hellma Analytics,

Müllheim, Germany) in a volume of 100 µl. To a solution of 5 mM D-arabinose-5-phosphate, 0.3 mM NADH in 50 mM glycyl-glycine buffer pH 7.5 with the auxiliary enzymes triose-phosphate isomerase and glycerol-3-phosphate dehydrogenase (TPI and GPDH: 10 U and 1 U respectively), DHA aldolase (0.25 µg of FSA_{AciBac} or 0.25 µg of FSA_{AcAeo}) was added at 25°C. The reaction was monitored at 340 nm following the consumption of NADH. One mmol of NADH oxidized was equivalent to 1 mmol of D-arabinose-5-phosphate cleaved. One unit (U) of DHA aldolase is defined as the amount of enzyme able to cleave 1 µmol of D-arabinose-5-phosphate to afford D-glyceraldehyde-3-phosphate per minute. A specific activity of 3.7 U/mg for FSA_{AciBac} and 0.03 U/mg for FSA_{AcAeo} were found.



Specific activity towards D-arabinose :

To a solution of D-arabinose (300 mM) in 50 mM glycyl-glycine buffer pH 7.5, DHA aldolase (10 µg of FSA_{AciBac} or 5 µg of FSA_{AcAeo}) and NADH (0.5 mM) were added at 25°C. The auxiliary enzyme alcohol dehydrogenase 20 µL (560 U) was added to reduce glycolaldehyde (GoA). The final volume was 1 mL. The reaction was monitored by spectrophotometry at 340 nm following the consumption of NADH. One mmol of NADH oxidized was equivalent to 1 mmol of D-arabinose cleaved. One unit (U) of DHA aldolase is defined as the amount of enzyme able to cleave 1 µmol of D-arabinose to afford glycolaldehyde per minute. A specific activity of 640 mU/mg for FSA_{AciBac} and 1 U/mg for FSA_{AcAeo} were found.



2.3. Optimum pH determination

FSA optimum pH was calculated assaying its retroaldol activity at pH values between 6.0 and 9.0 in 50 mM of glycylglycine, Tris-HCl and phosphate buffer. Activities were determined by measuring the formation of D-glyceraldehyde-3-phosphate (D-G3P) from 5 mM of D-arabinose-5-phosphate, with 0,3 mM NADH and the auxiliary enzymes triose-phosphate isomerase and glycerol-3-phosphate dehydrogenase (10 U and 1 U respectively), for FSA_{AciBac} (2.5 µg/ml). Concerning FSA_{AcAeo}, formation of glycolaldehyde was followed from 25 mM of D-arabinose with 25 µg/ml FSA_{AcAeo} for glycylglycine and Tris-HCl buffer or 50 µg/ml for phosphate buffer, with 0.4 mM NADH and the auxiliary enzyme alcohol dehydrogenase (371 U). The reaction was monitored at 340 nm following the consumption of NADH. Variations of A_{340nm} were proportional to formation of D-G3P for FSA_{AciBac} or formation of glycolaldehyde for FSA_{AcAeo} in reaction (ε_{NADH} = 6220 $cm^{-1}M^{-1}$).

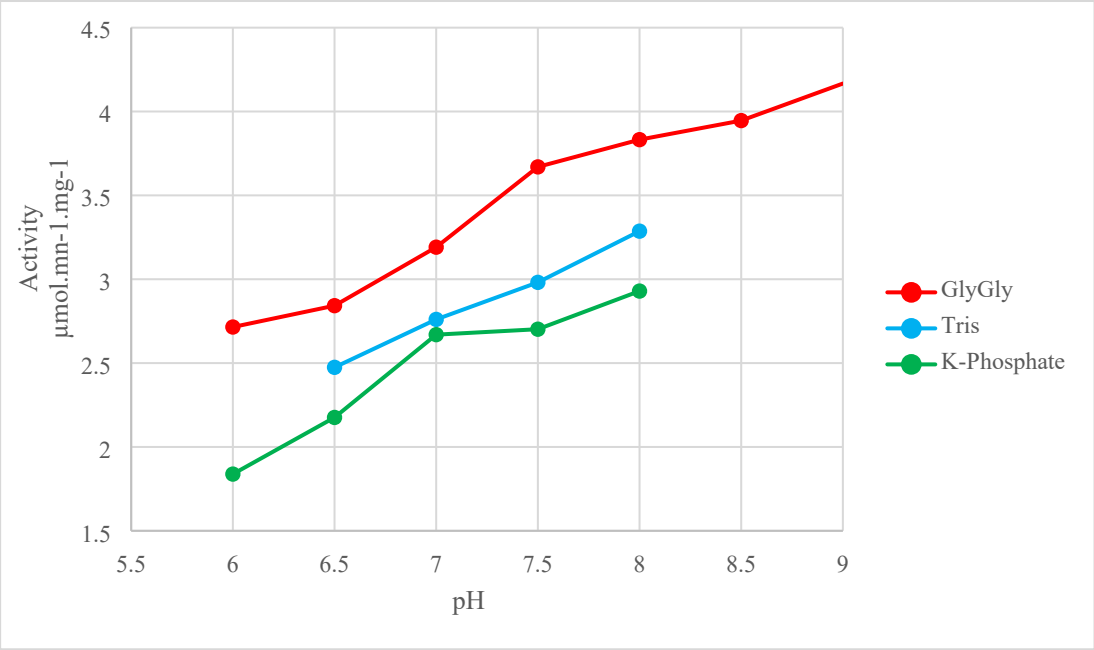


Figure S4 : pH study for FSA_{AciBac}

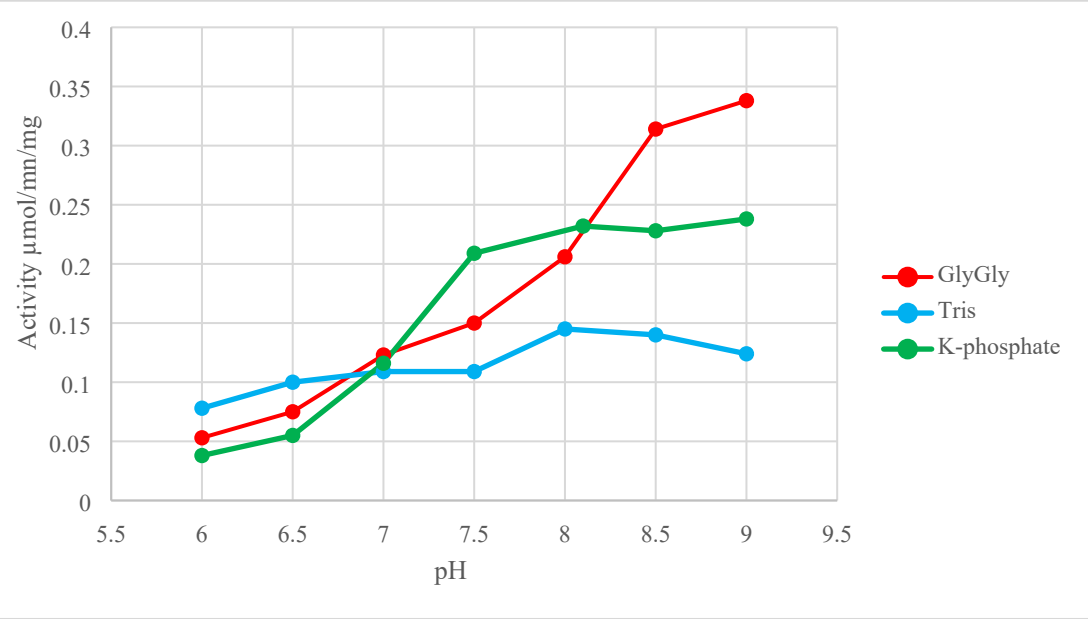


Figure S5 : pH study for FSA_{AcAeo}

2.4. Enzyme thermostability

Melting temperatures of FSA_{AcAeo}, FSA_{AciBac} and FSA_{Ecoli} were determined by fluorescence spectroscopy using the protocol described by Life Technologies®. A 125-fold dilution of Protein Thermal Shift^{Tm} Dye (fluorophore) was carried out, and Protein Thermal Shift^{Tm} Buffer (5µL), then the enzyme solution (12.5 µL *i.e.* 5 µg of FSA_{AcAeo} or 62.5µg for FSA_{AciBac} and $FSA_{E.coli}$) and Protein Thermal Shift^{Tm} Dye (2.5 µL) were added successively to each well in a 96-well plate. The plate was then sealed with MicroAmp® Optical Adhesive Film and shaken at 1000 rpm for 1 min. The plate was then inserted into the Applied Biosystems Real-

Time PCR System. The plate was incubated at 25°C for 2 minutes, then a gradient of 0.05°C/s is applied until 99°C was reached. Fluorescence was measured in real time on a channel with an x4 excitation filter set on λ= 580±10 nm and an m4 emission filter set on λ=623±14 nm. Melting curves were obtained and once derived, the minimum value gave the Tm.

For FSA_{AcAeo}, Tm was found at 73°C; for FSA_{AciBac}, Tm of the enzyme was found at 62°C: this first destructuration was associated with a loss of enzymatic activity (measured during an additional manipulation) the enzyme then underwent a more significant destructuration around 81°C. For FSA_{Ecoli} the melting temperature was measured at 70°C by this method, in accordance with the literature (established by other methods).

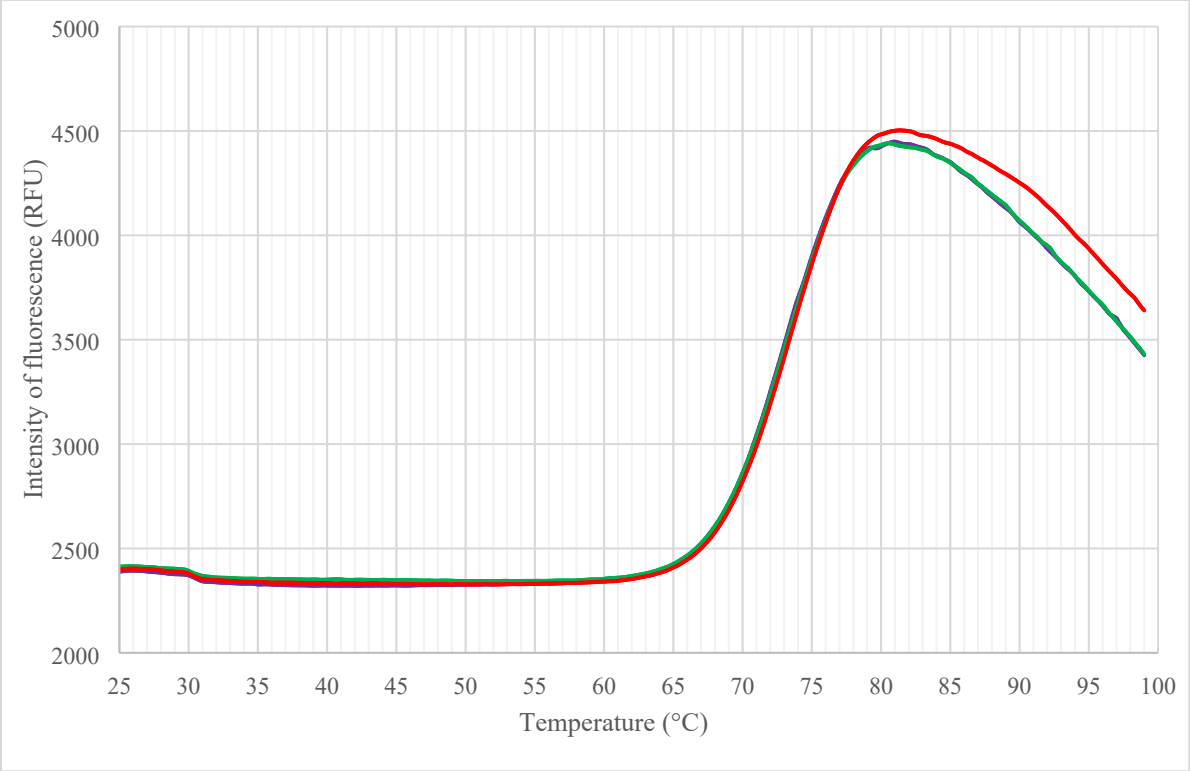


Figure S6 : Melting curve of FSA_{AcAeo} by measuring fluorescence intensity as a function of temperature, in triplicate

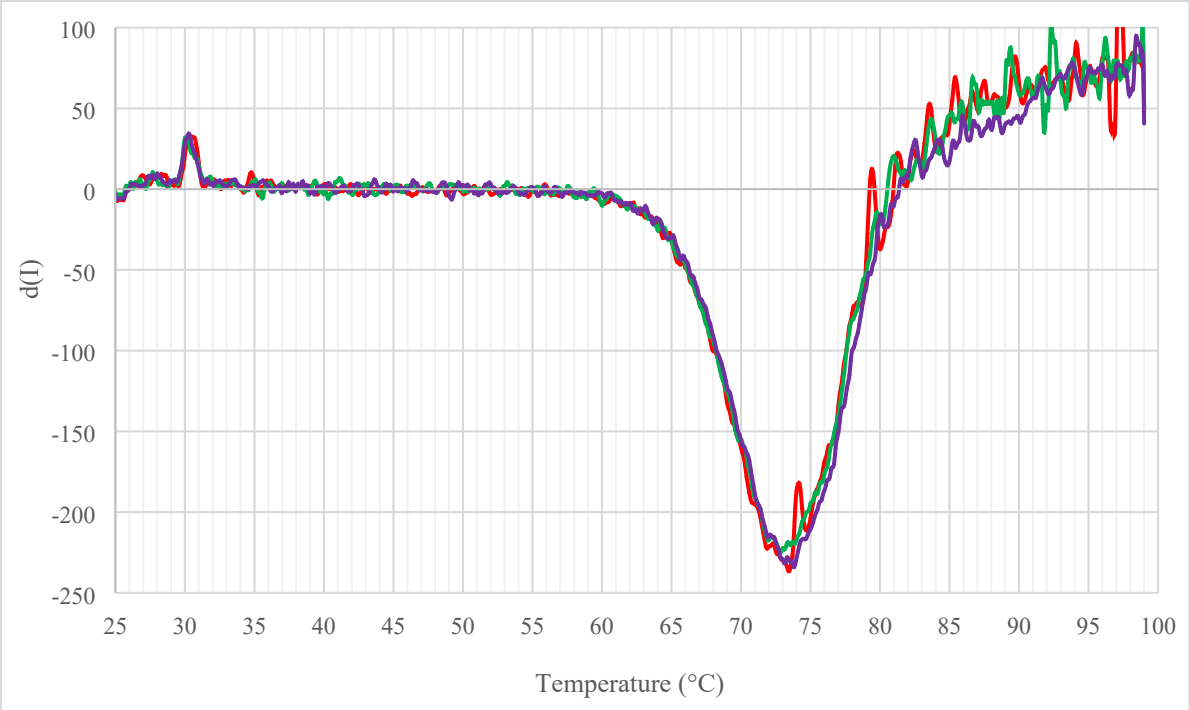


Figure S7 : Derivative of fluorescence intensity for FSA_{AcAeo} as a function of temperature, in triplicate

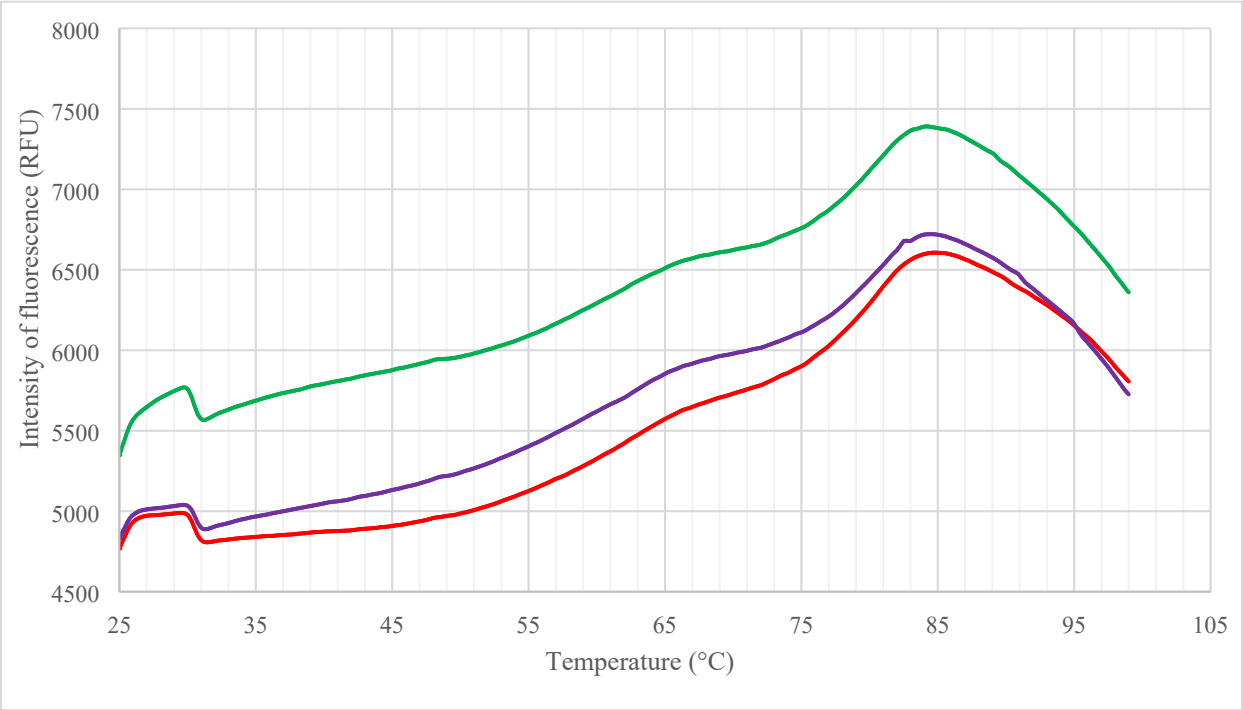


Figure S8 : Melting curve of FSA_{AciBac} by measuring fluorescence intensity as a function of temperature, in triplicate

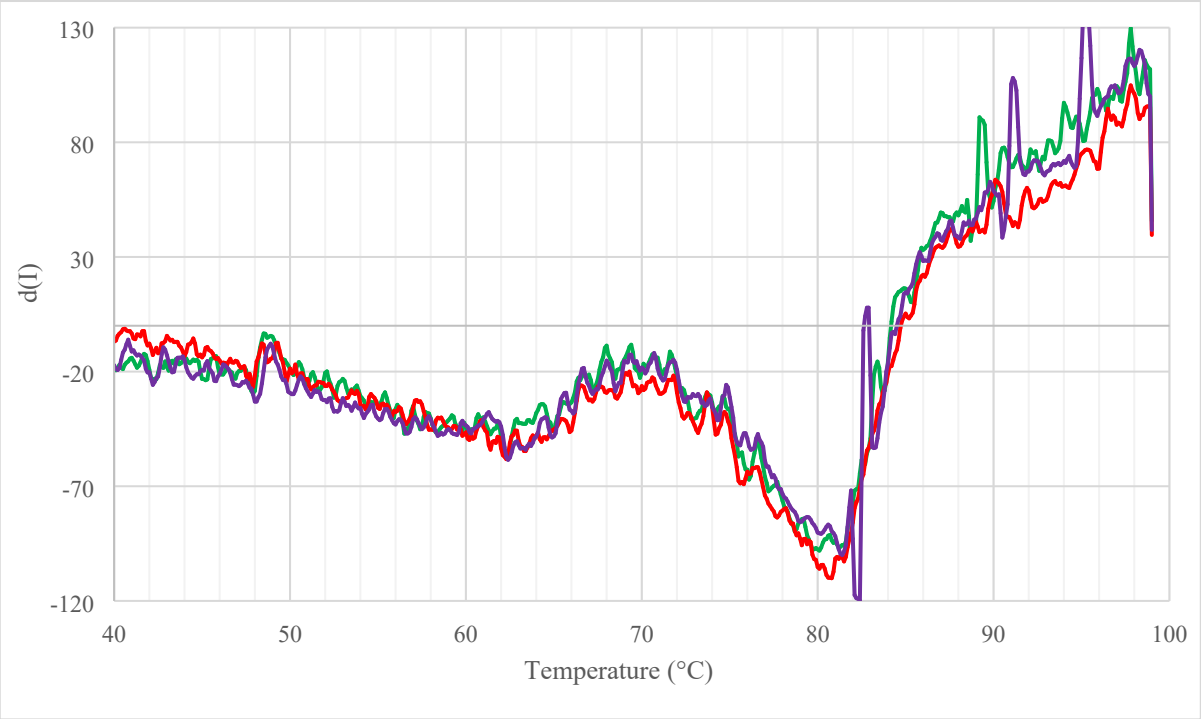


Figure S9 : Derivative of fluorescence intensity for FSA_{AciBac} as a function of temperature, in triplicate

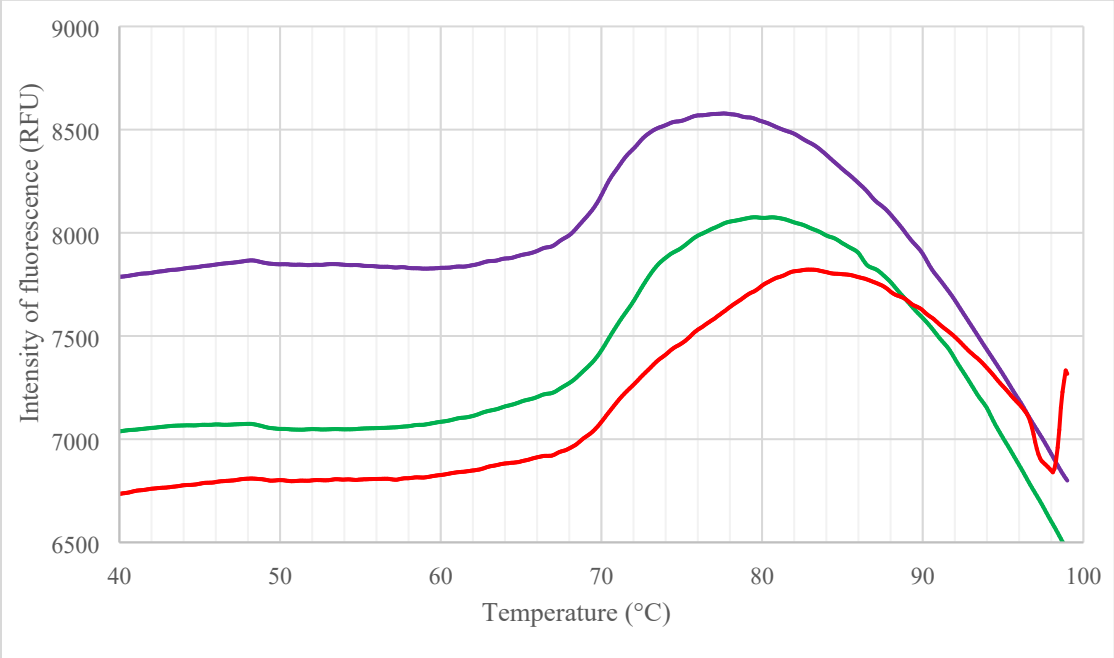


Figure S10 : Melting curve of FSA_{Ecoli} by measuring fluorescence intensity as a function of temperature, in triplicate.

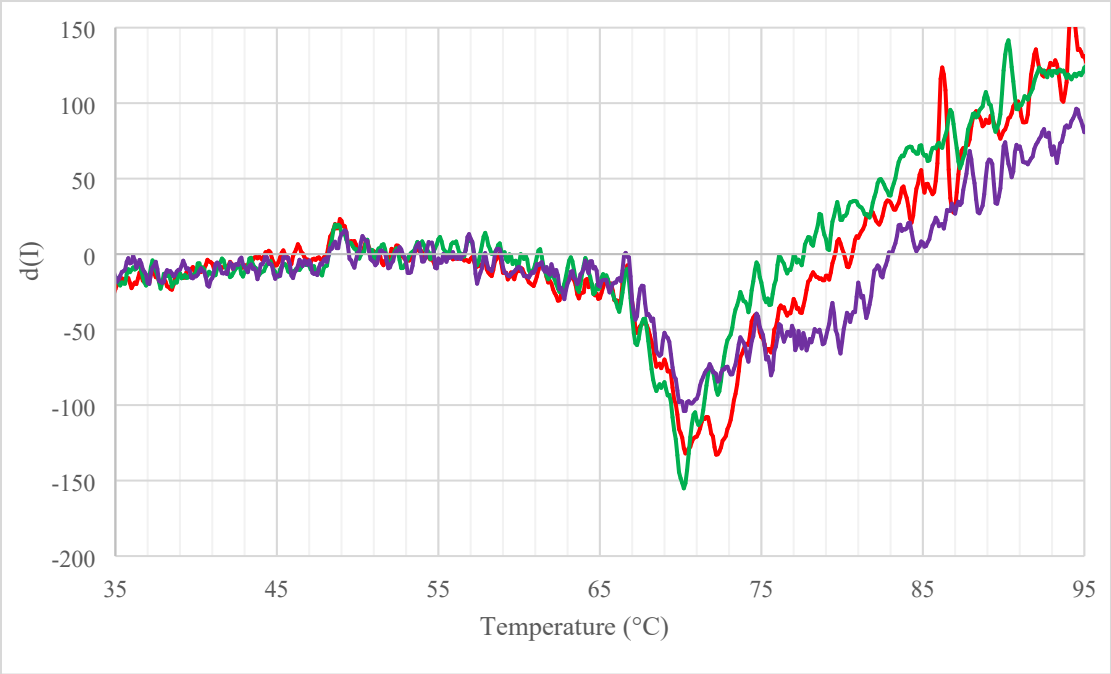


Figure S11 : Derivative of fluorescence intensity for FSA_{Ecoli} as a function of temperature, in triplicate.

2.5. Kinetic constants

FSA kinetics towards D-arabinose-5-phosphate (A5P) or D-arabinose (A) were determined using the same assays as previously described. Each kinetic assay was implemented using 5-300 mM of D-arabinose or 0.2-8 mM of D-arabinose-5-phosphate in glycylglycine buffer 50 mM pH 7.5. with 0.5 µg of FSA_{AciBac}, 2.5 µg -100 µg of FSA_{AcAeo} or 0.4 µg -2.0 mg of FSA_{Ecoli}, the amount of enzyme varying in the last two cases to maintain an enzymatic rate that can be accurately measured with the spectrophotometer.

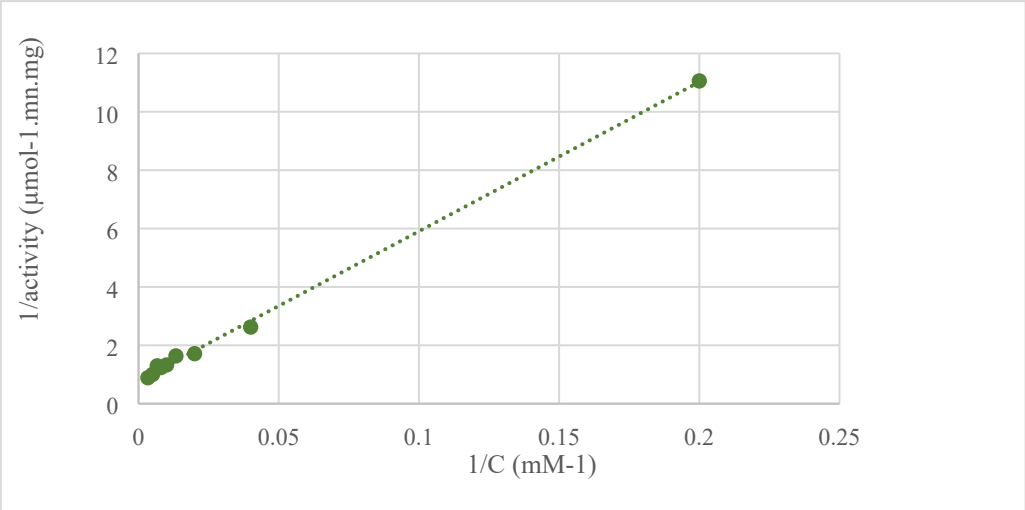


Figure S12 : Lineweaver-Burk plot for FSA_{AcAeo} with D-arabinose in glycylglycine buffer (pH 7.5, 50 mM)

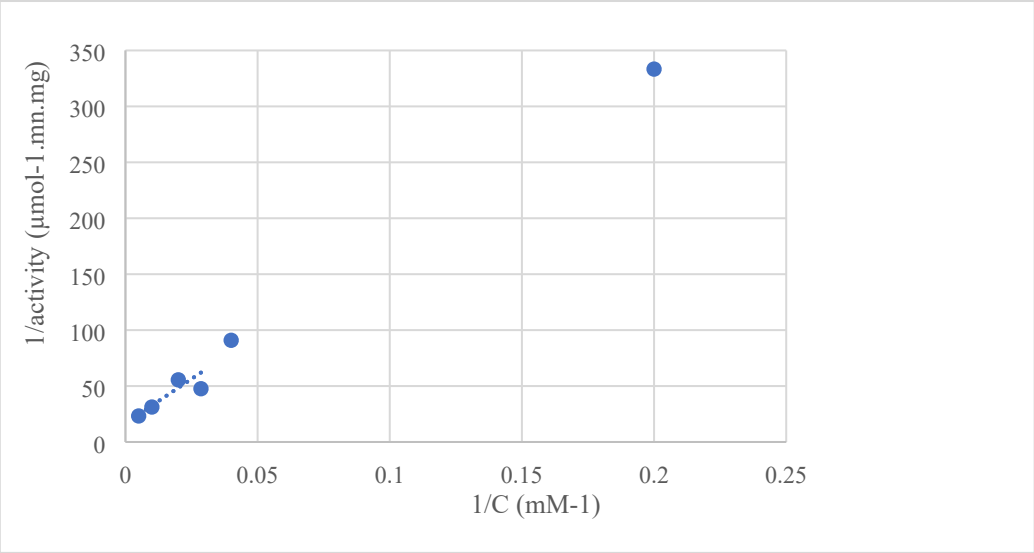


Figure S13 : Lineweaver-Burk plot for $FSA_{E.coli}$ with D-arabinose in glycylglycine buffer (pH 7.5, 50 mM)

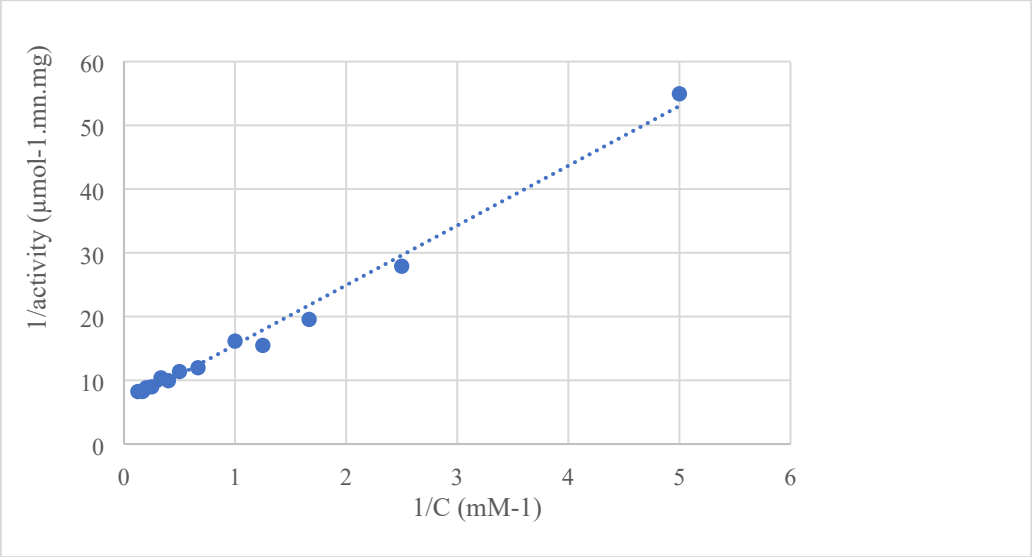


Figure S14 : Lineweaver-Burk plot for FSA_{AciBac} with D-arabinose-5-phosphate in glycylglycine buffer (pH 7.5, 50 mM)

2.6. Isotopic exchange measurements on nucleophiles by ^{1}H NMR

To a solution of 2 mg/mL of aldolase (FSA_{Ecoli}, $\underline{FSA}_{AciBac}$, FSA_{AcAeo}) in D_2O (500 µL) adjusted to pH 7.5, were added 25 µmol (50 mM) of the studied nucleophile. ^{1}H NMR spectra were recorded after 1 h and 24 h. Activity and selectivity were determined by integration of the signal corresponding to the methylene group of each nucleophile.

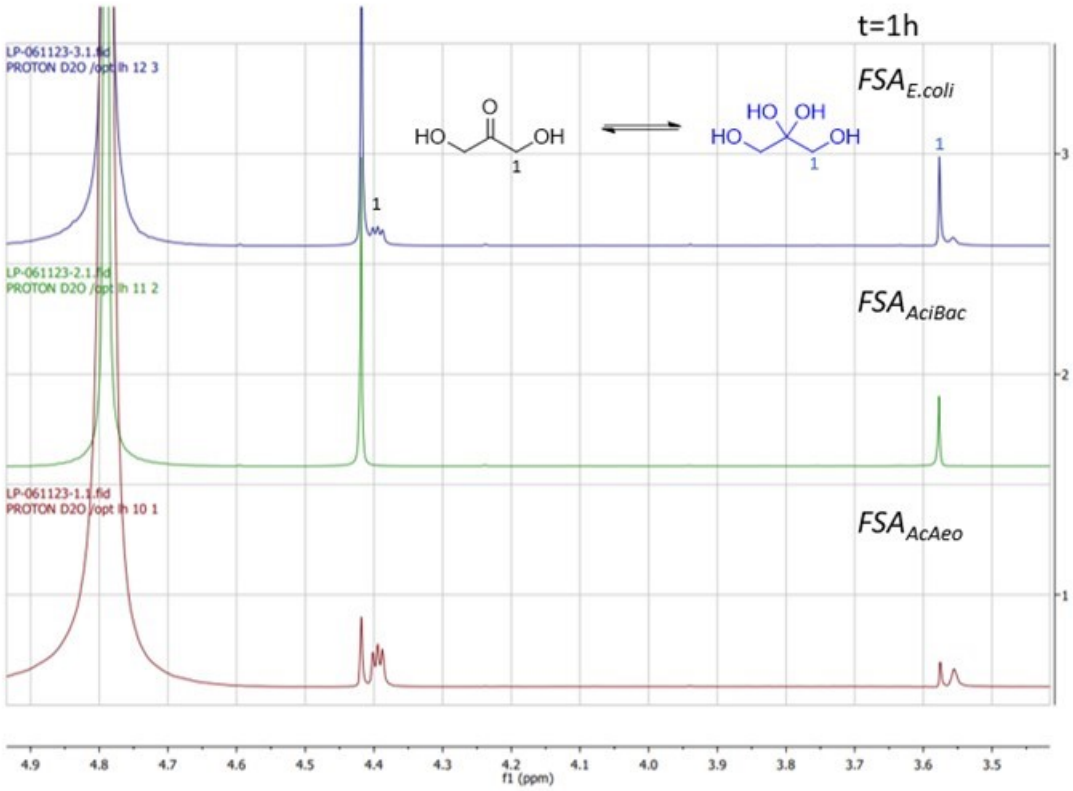


Figure S15 : Isotopic exchange after 1 h for DHA in the presence of FSA_{Ecoli}, FSA_{AciBac} and FSA_{AcAeo}

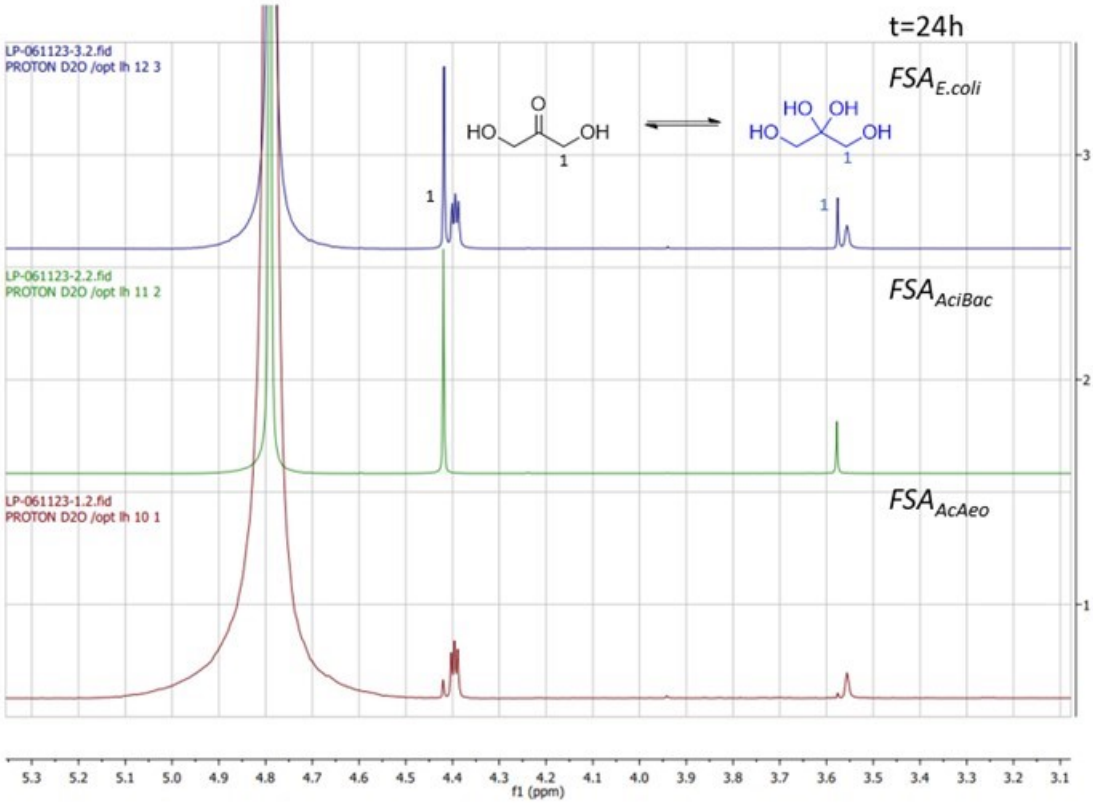


Figure S16 : Isotopic exchange after 24 h for DHA in the presence of FSA_{Ecoli}, FSA_{AciBac} and FSA_{AcAeo}

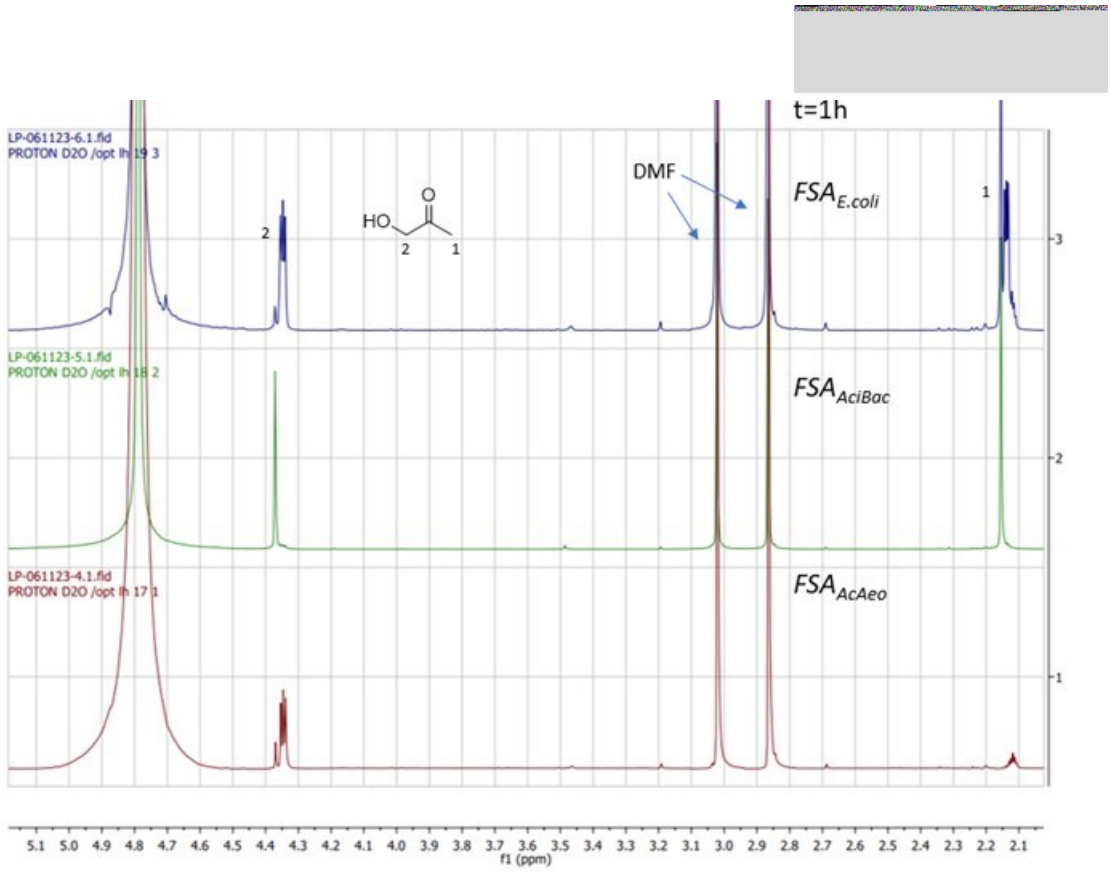


Figure S17 : Isotopic exchange after 1 h for HA in the presence of FSA_{Ecoli}, FSA_{AciBac} and FSA_{AcAeo}

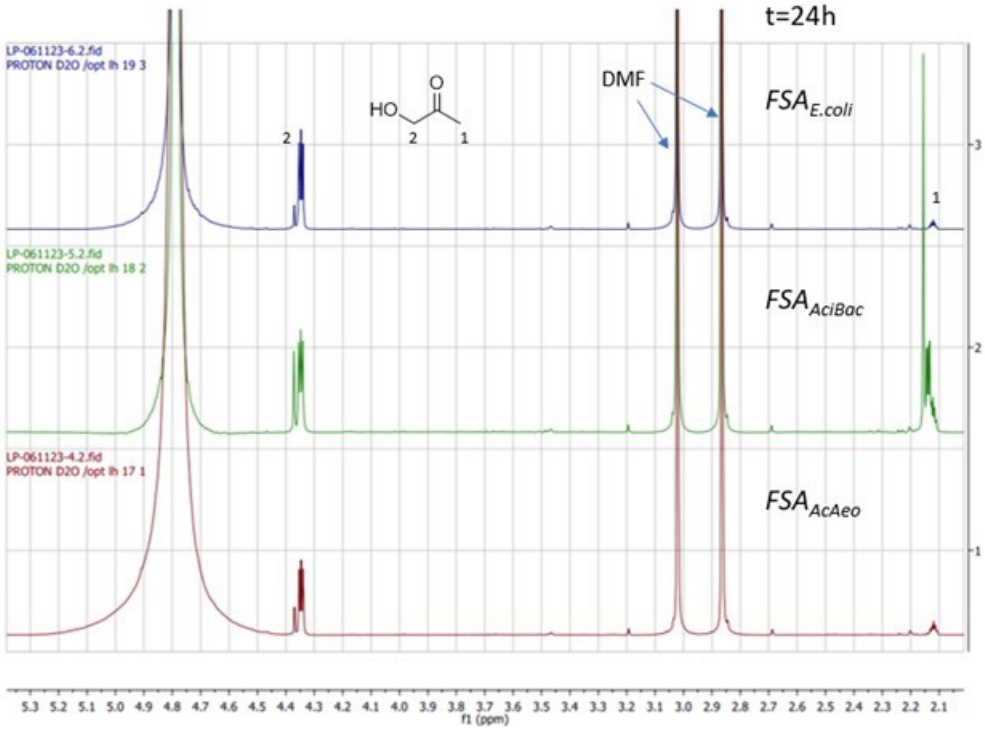


Figure S18 : Isotopic exchange after 24 h for HA in the presence of FSA_{Ecoli}, FSA_{AciBac} and FSA_{AcAeo}

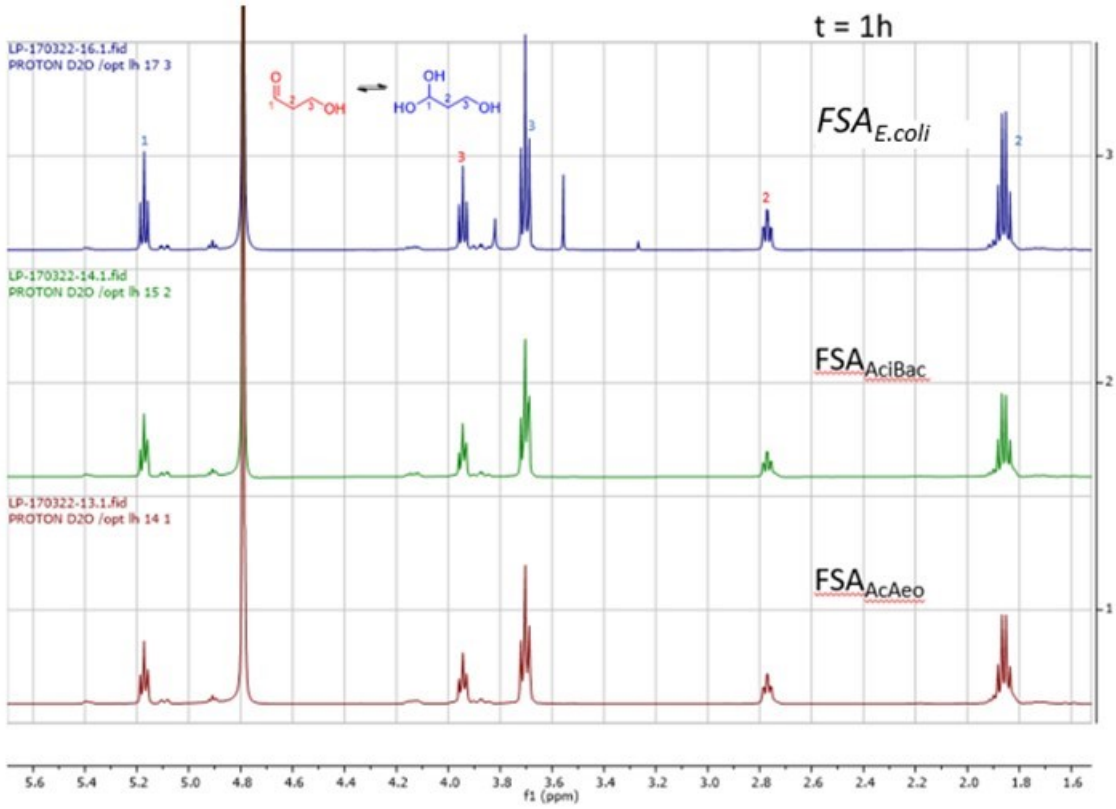


Figure S19 : Isotopic exchange after 1 h for 3-hydroxypropanal in the presence of FSA_{Ecoli}, FSA_{AciBac} and FSA_{AcAeo}

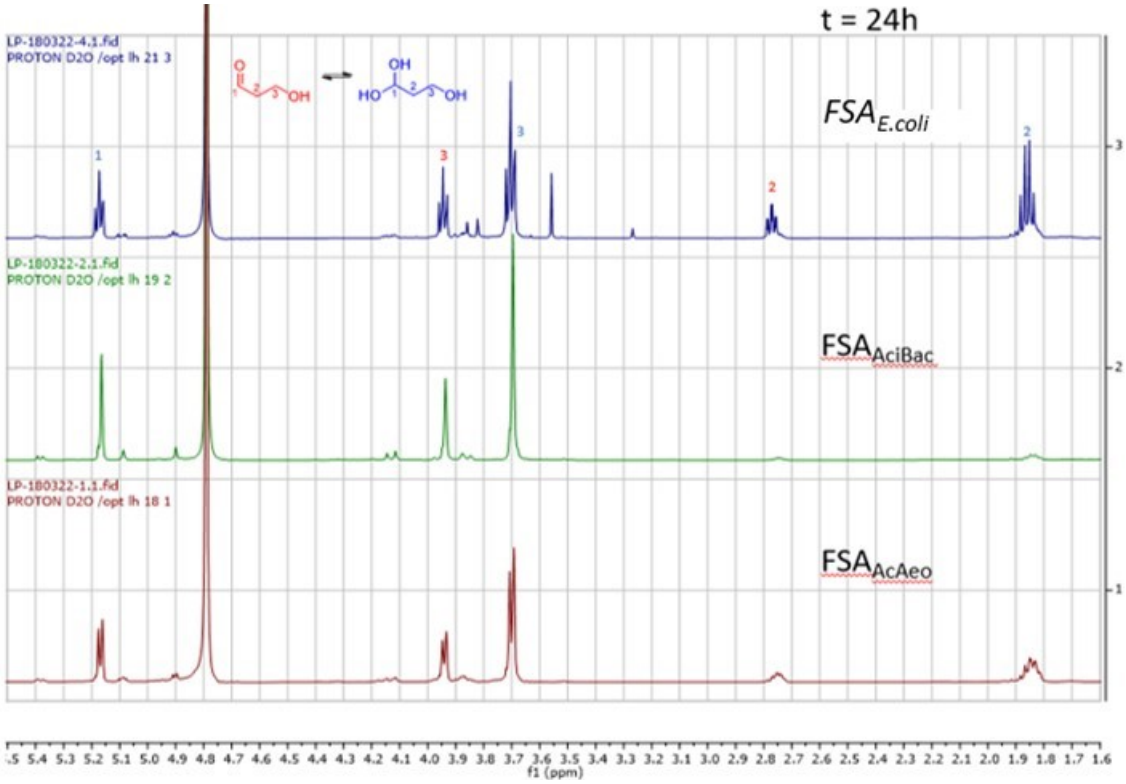


Figure S20 : Isotopic exchange after 24 h for 3-hydroxypropanal in the presence of FSA_{Ecoli}, FSA_{AciBac} and FSA_{AcAeo}

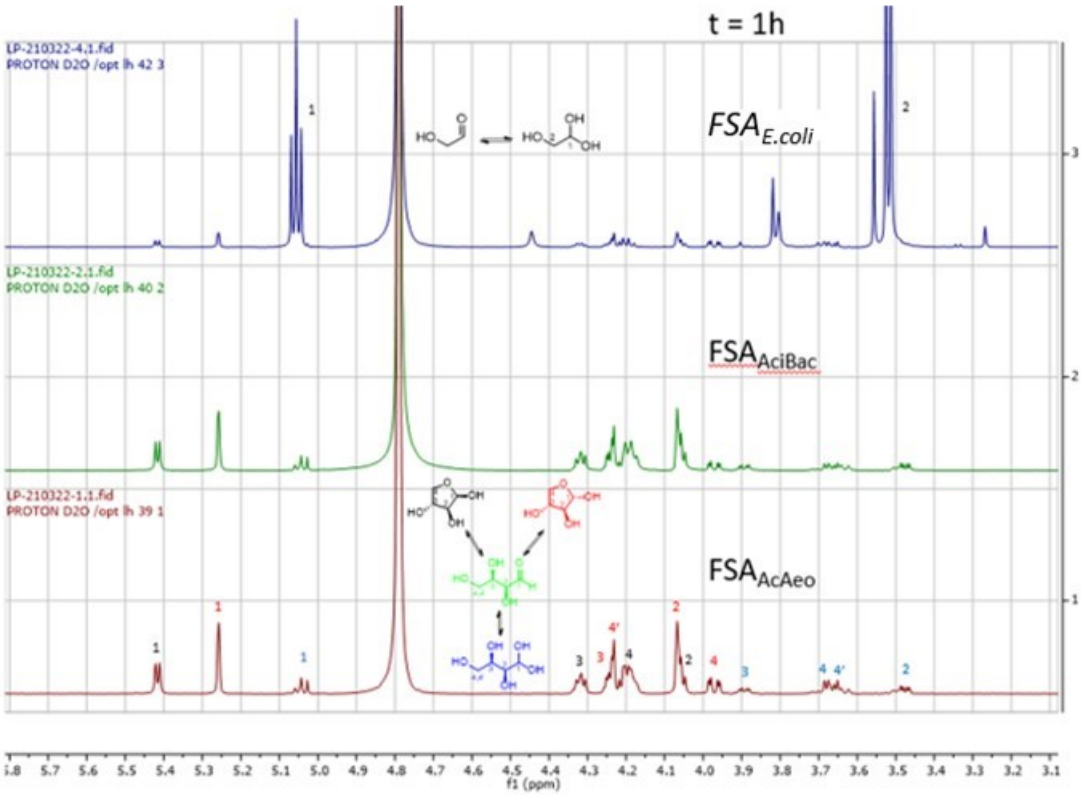


Figure S21 : Isotopic exchange after 1 h for glycolaldehyde in the presence of FSA_{Ecoli}, FSA_{AciBac} and FSA_{AcAeo}

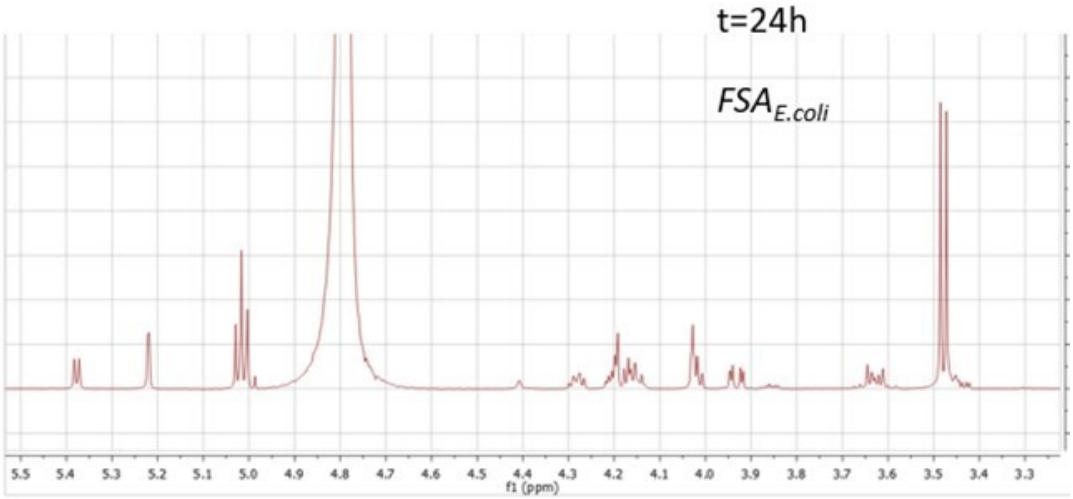


Figure S22 : Isotopic exchange after 24 h for glycolaldehyde in the presence of FSA_{Ecoli}

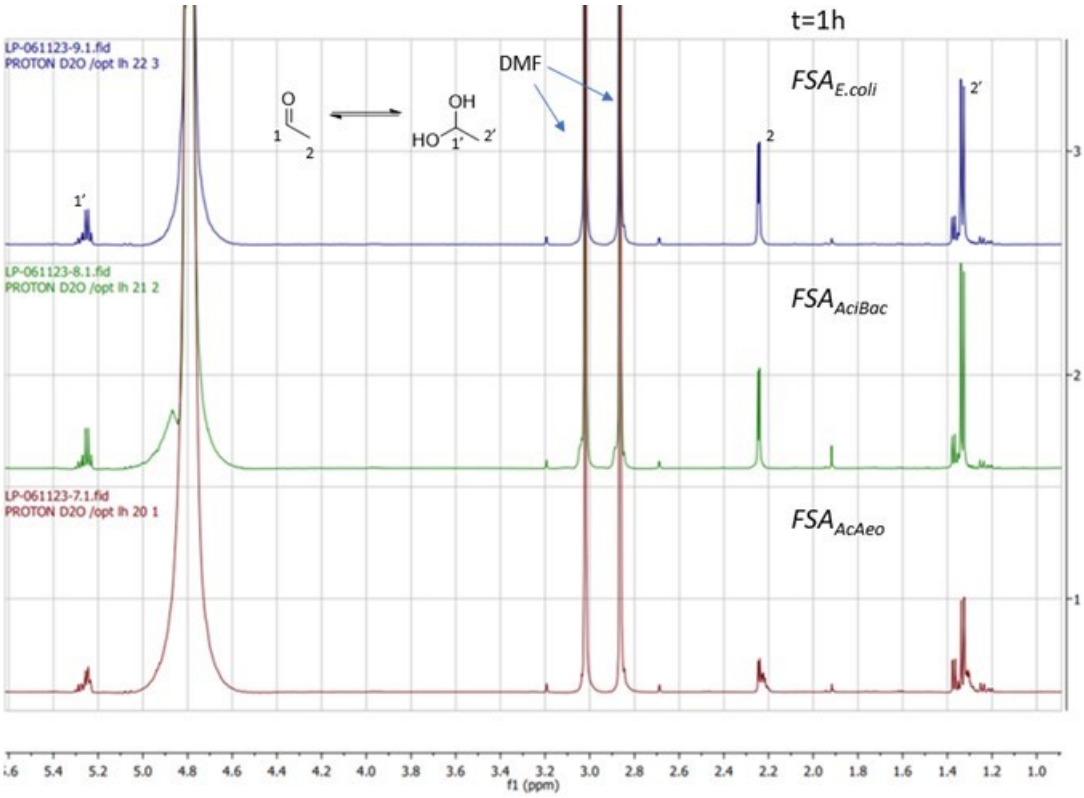


Figure S23 : Isotopic exchange after 1 h for acetaldehyde in the presence of FSA_{Ecoli}, FSA_{AciBac} and FSA_{AcAeo}

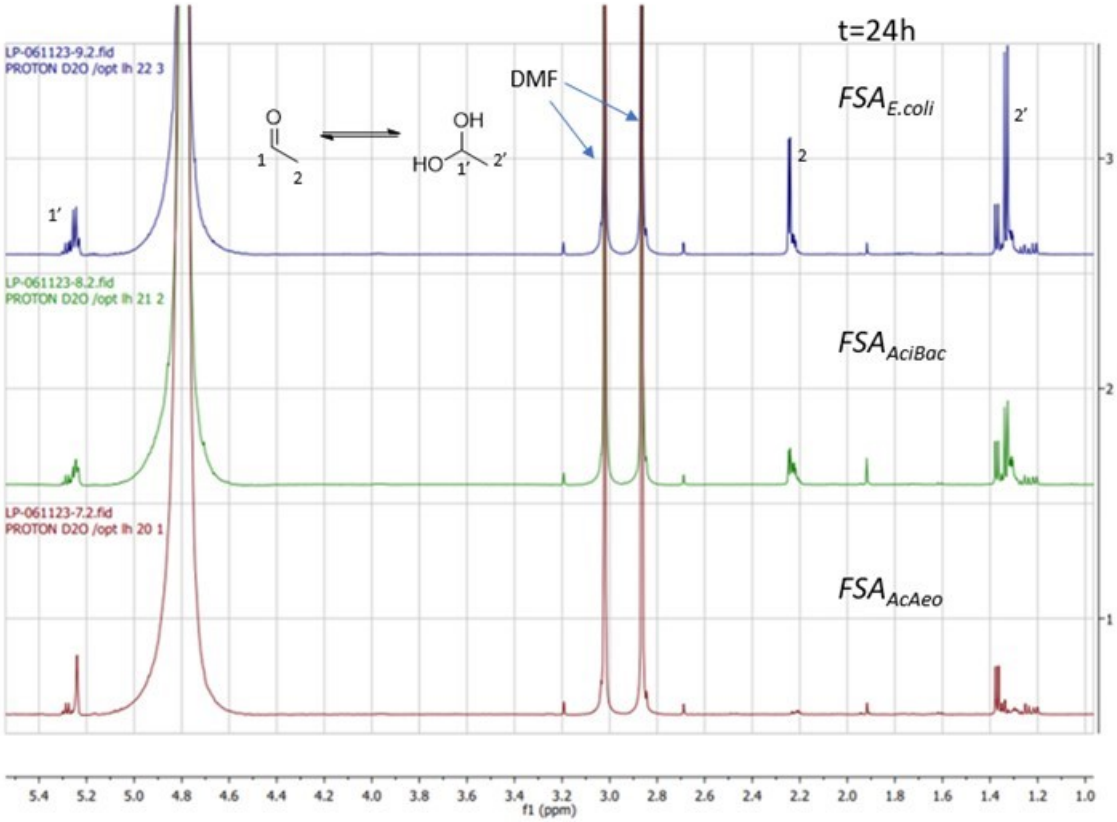


Figure S24 : Isotopic exchange after 24 h for acetaldehyde in the presence of FSA_{Ecoli}, FSA_{AciBac} and FSA_{AcAeo}

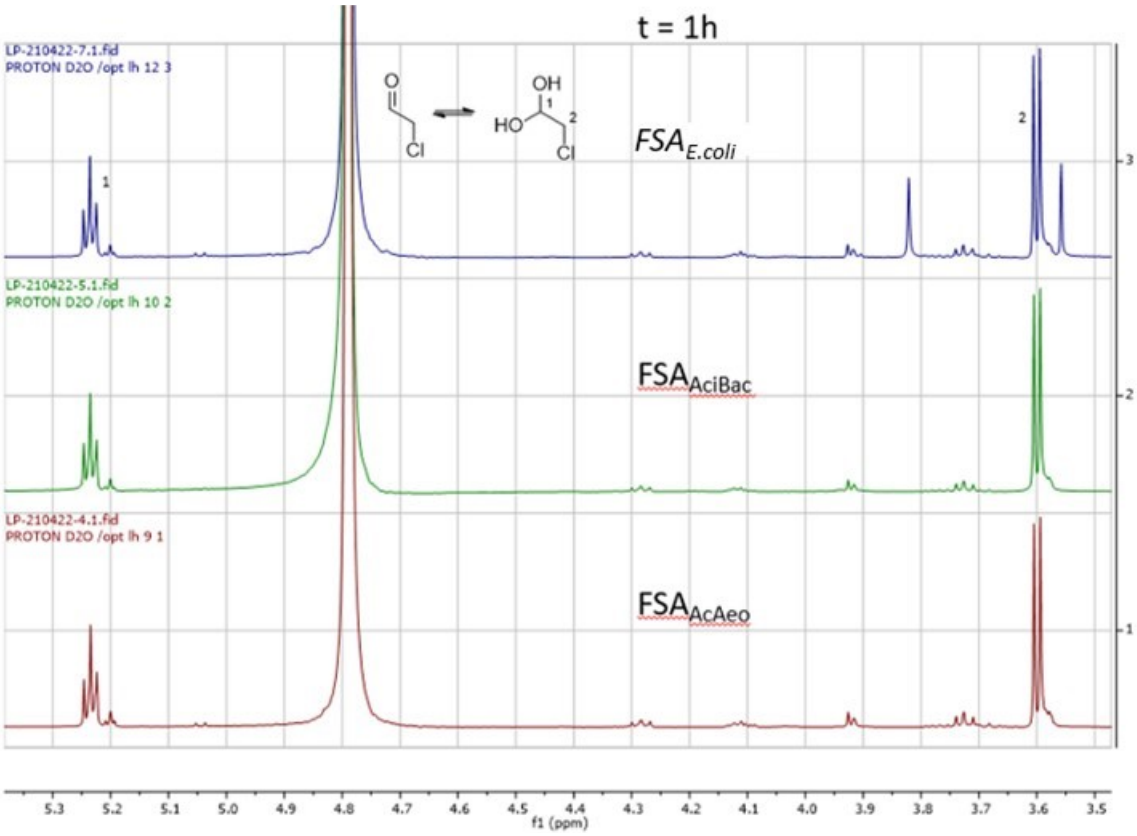


Figure S25 : Isotopic exchange after 1 h for chloroacetaldehyde in the presence of FSA_{Ecoli}, FSA_{AciBac} and FSA_{AcAeo}

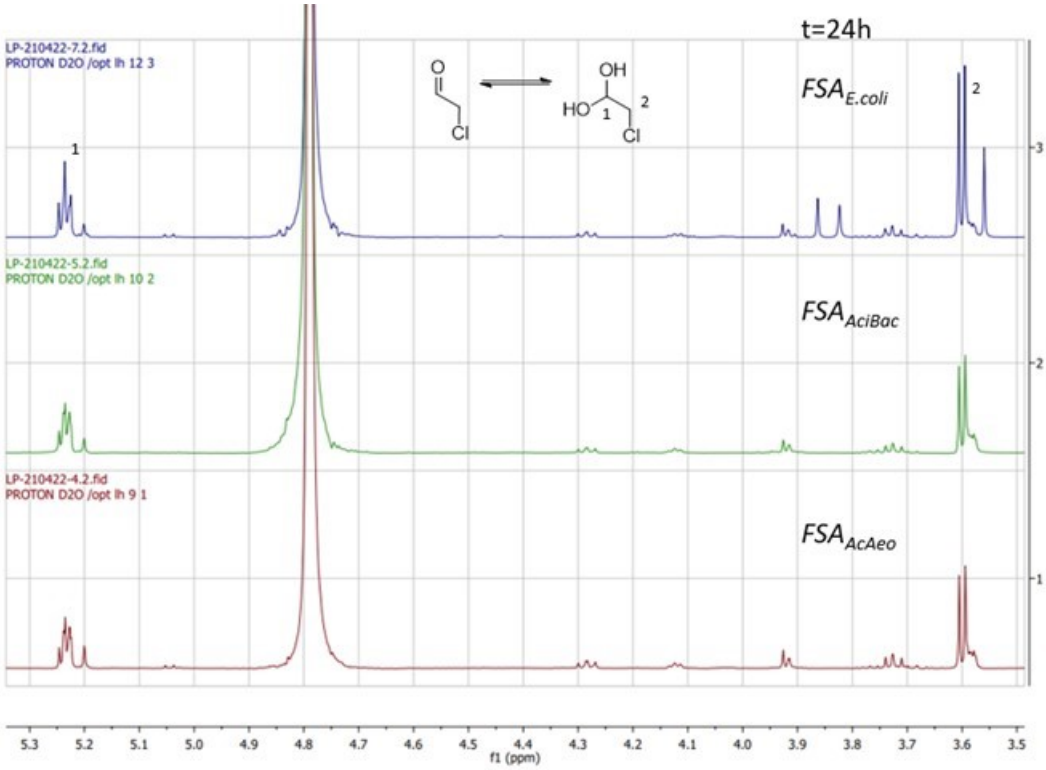


Figure S26 : Isotopic exchange after 24 h for chloroacetaldehyde in the presence of FSA_{Ecoli}, FSA_{AciBac} and FSA_{AcAeo}

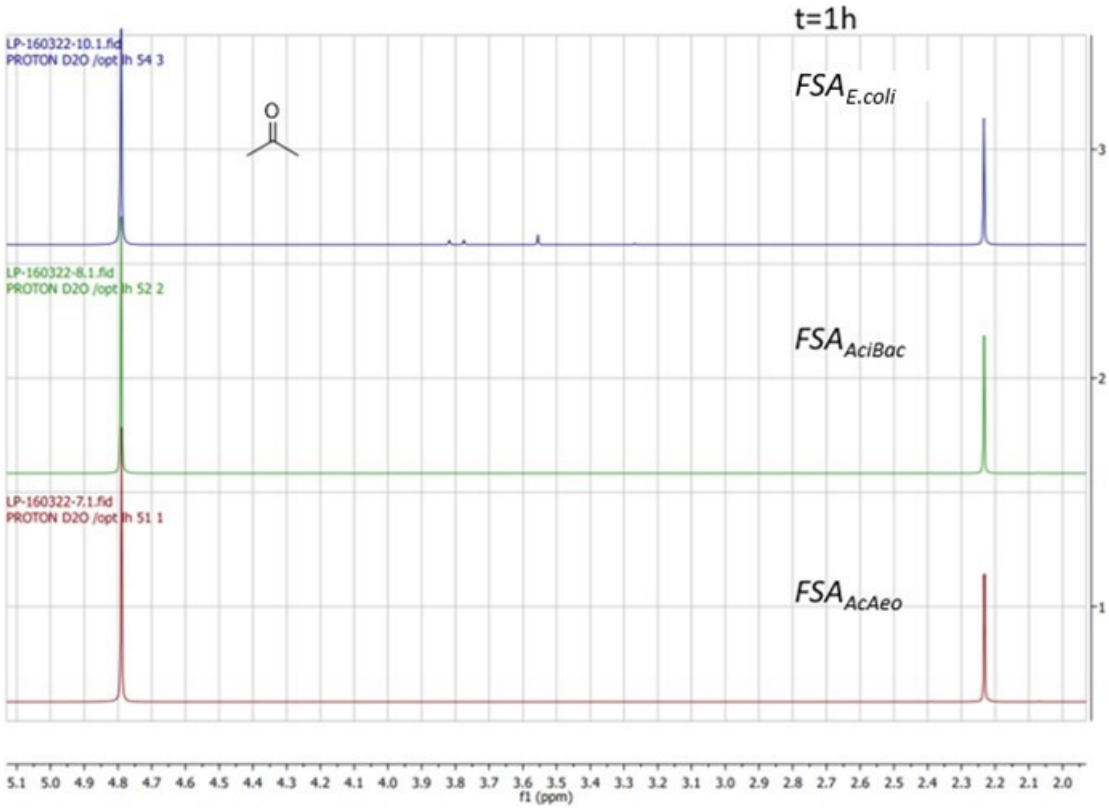


Figure S27 : Isotopic exchange after 1 h for acetone in the presence of FSA_{Ecoli}, FSA_{AciBac} and FSA_{AcAeo}

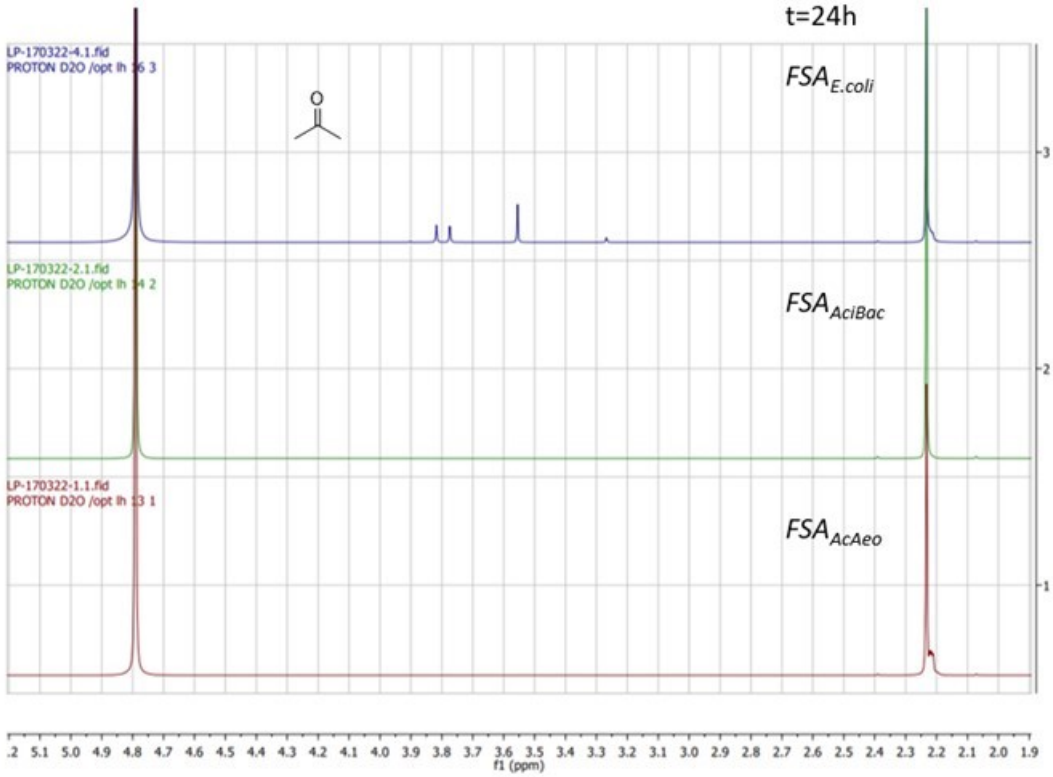


Figure S28 : Isotopic exchange after 24 h for acetone in the presence of FSA_{Ecoli}, FSA_{AciBac} and FSA_{AcAeo}

3. Molecular modelling

3.1. Construction of the structures

Structures (Resolution, date)	A0A399XY01: FSA_{AciBac} percent identity	A0A0Q0RVA3: FSA_{AcAeo} percent identity	Organism	Publications (DOI)
1L6W (1.93, 2002)[4]	26.9	26.8	*Escherichia coli*	10.1016/S0022-2836(02)00258-9
1VPX (2.40, 2004)	29.2	36.0	*Thermotoga maritima*	*To be published.*
1WX0 (2.27, 2005)	31.0	31.8	*Thermoplasma acidophilum DSM 1728*	
3S0C (1.78, 2011)[5]	28.2	42.2	*Thermoplasma acidophilum DSM 1728*	10.1038/nchembio.633
3S1U (1.90, 2011)	28.2	42.2	*Thermoplasma acidophilum DSM 1728*	
3S1V (1.80, 2011)	28.2	42.2	*Thermoplasma acidophilum DSM 1728*	
3S1X (1.65, 2011)	28.2	42.2	*Thermoplasma acidophilum DSM 1728*	
3S1W (1.80, 2011, K86A)	27.8	41.7	*Thermoplasma acidophilum DSM 1728*	
3R8R (1.90, 2011)[6]	31.1	34.4	*Bacillus subtilis*	10.1111/j.1742-4658.2011.08467.x
4RXF (2.40, 2014, Y131F)[7]	25.9	25.3	*Escherichia coli K-12*	10.1002/cctc.201500478
4RZ4 (1.75, 2014, Q59E, Y131F)	25.5	24.9	*Escherichia coli K-12*	
4RXG (2.15, 2014)	26.4	26.4	*Escherichia coli K-12*	
4S1F (2.24, 2015)	26.4	25.8	*Escherichia coli K-12*	
4XZ9 (1.80, 2015)[8]	29.2	43.1	*Thermoplasma acidophilum DSM 1728*	10.1021/acs.biochem.5b00283
5ZOL (2.17, 2018, I31T, Q59T, I195Q)[9]	26.4	25.3	*Escherichia coli K-12*	10.1039/d0cc02437f
6YR3 (1.48, 2020)	28.2	42.2	*Thermoplasma acidophilum DSM 1728*	*To be published*
6YS0 (1.70, 2020, D211A)	28.7	41.7	*Thermoplasma acidophilum DSM 1728*	
6YRT (1.65, 2020, T30D)	28.2	42.2	*Thermoplasma acidophilum DSM 1728*	
6YRM (1,70, 2020, T30A)	28.2	42.2	*Thermoplasma acidophilum DSM 1728*	
6YRE (1.96, 2020, T30C, D211C)	28.2	42.2	*Thermoplasma acidophilum DSM 1728*	
6YRH (1.80, 2020, T30C, D211C)	28.2	42.2	*Thermoplasma acidophilum DSM 1728*	

*Table 3. 21 p*rotein sequences found from the PDB after Basic Local Alignment with FSA_{AciBac} and FSA_{AcAeo} using Chimera software.

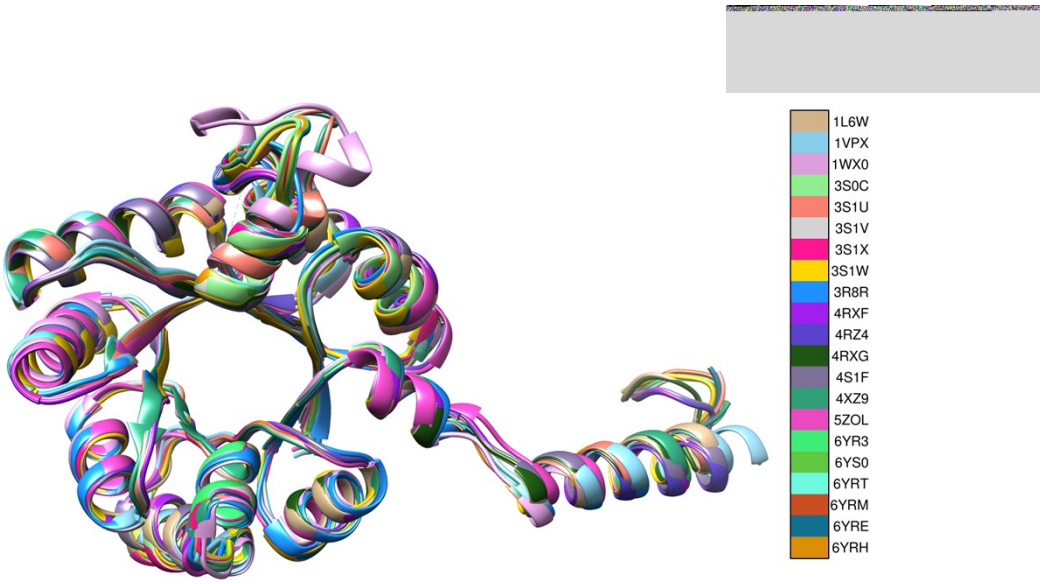


Figure S29. Superimposition of the 21 proteins showing an identical (β/α) 8-barrel fold secondary structure near the reactive lysine and a Cter alpha helix varying in length.

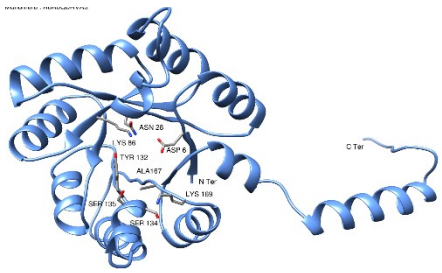


Figure S30. Proposed structure for FSA_{AcAeo}

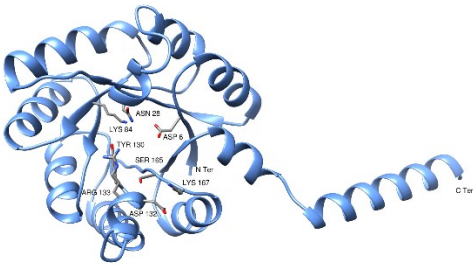


Figure S31. Proposed structure for FSA_{AciBac}

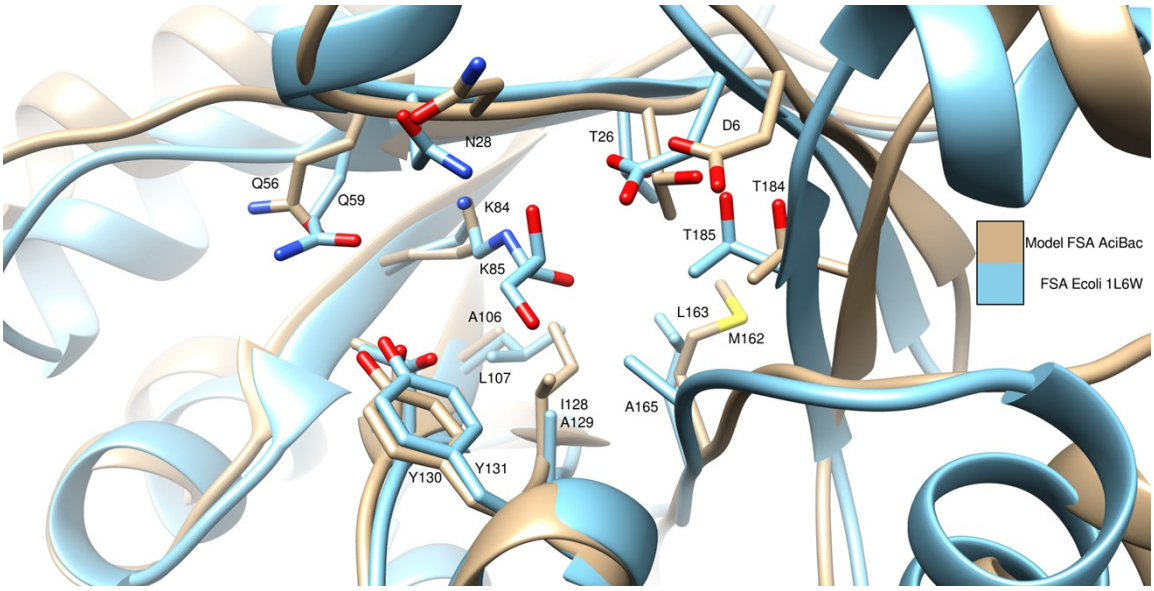


Figure S32. Superimposition of 1L6W crystallographic structure of FSA_{Ecoli} in blue with our model structure of FSA_{AciBac} in brown determined using Modeller 10.4,[10] Chimera 1.17.1[11] and AlphaFold[12].

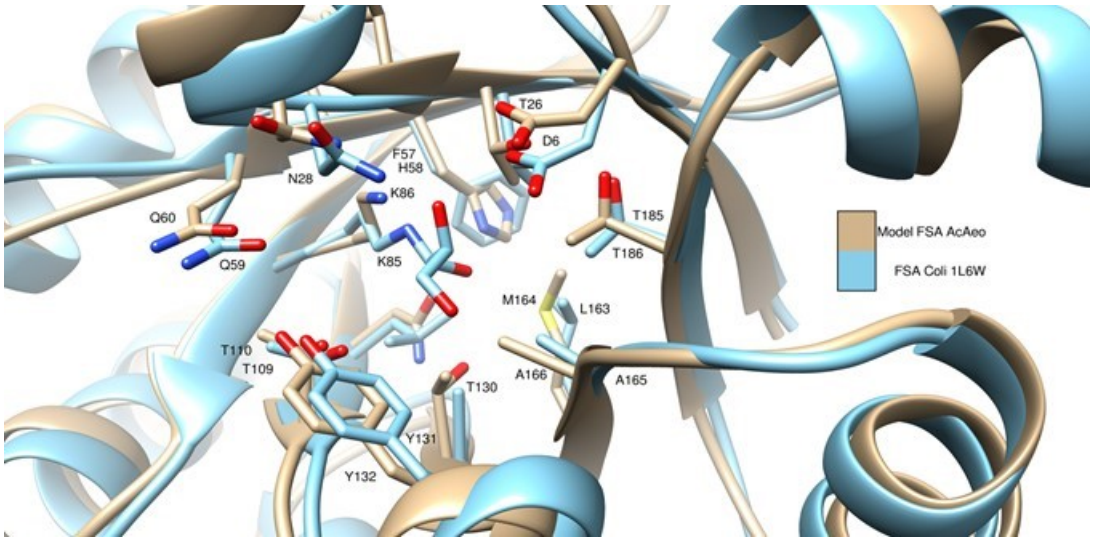


Figure S33. Superimposition of 1L6W crystallographic structure of FSA_{Ecoli} in blue with our model structure of FSA_{AcAeo} in brown.

3.2. Docking experiments

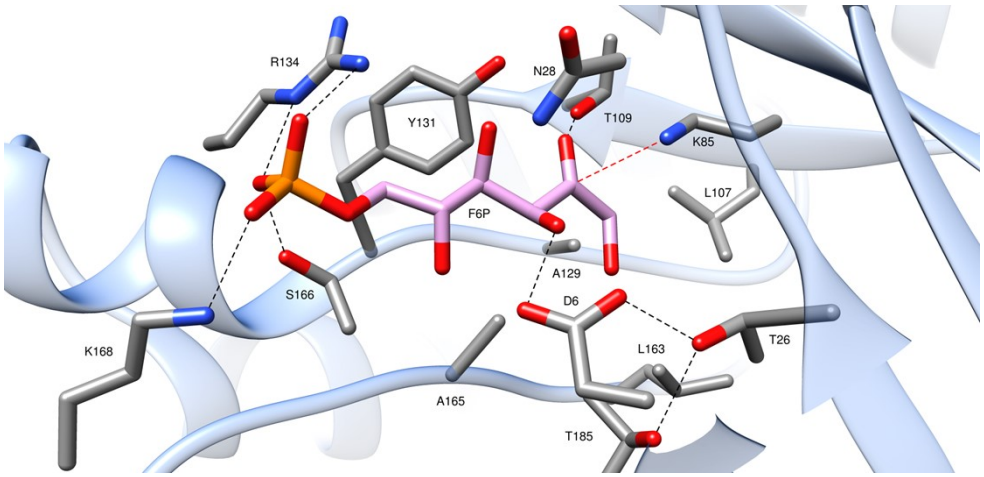


Figure S34. Docking of fructose-6-phosphate (pink) in FSA_{Ecoli} active site with Autodock 4.2™,[13] leading to a binding energy of -8.5 kcal/mol for the best solution over 500 belonging to a same cluster, and an estimated Ki of 572 nM.

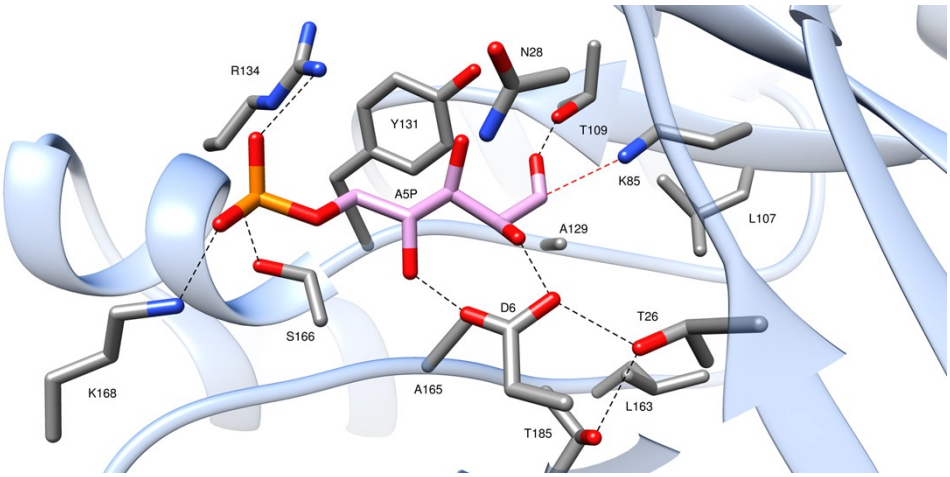


Figure S35. Docking of arabinose-5-phosphate (pink) in FSA_{Ecoli} active site with Autodock 4.2™, leading to a binding energy of -8.9 kcal/mol for the best solution over 500 belonging to a same cluster, and an estimated Ki of 322 nM.

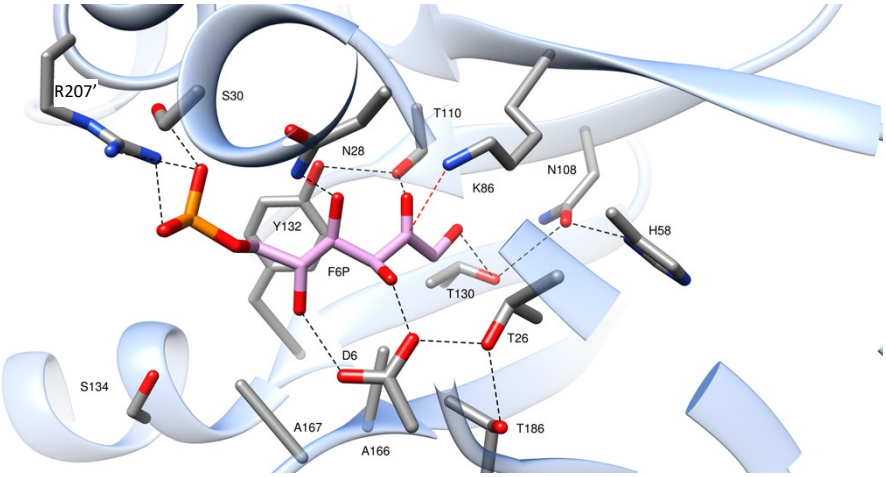


Figure S36. Docking of fructose-6-phosphate (pink) in FSA_{AcAeo} active site with Autodock 4.2™, leading to a binding energy of -5.7 kcal/mol for the best solution over 497 belonging to a same cluster, and an estimated Ki of 65 µM.

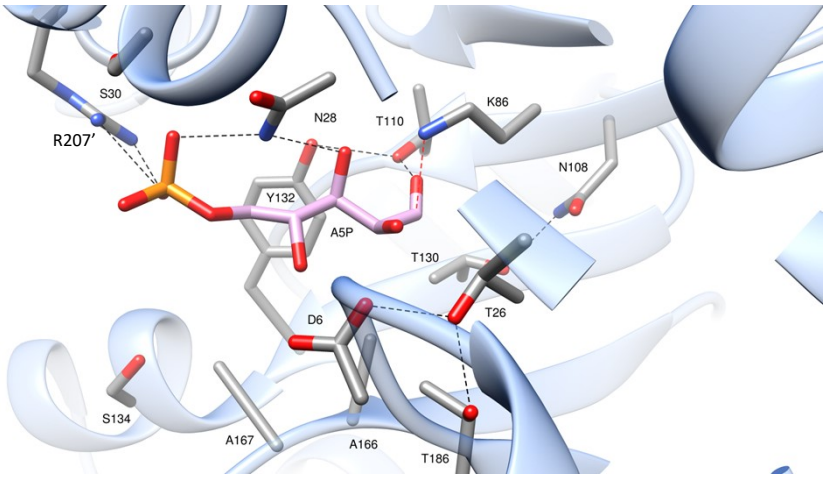


Figure S37. Docking of arabinose-5-phosphate (pink) in FSA_{AcAeo} active site with Autodock 4.2™, leading to a binding energy of -5.6 kcal/mol for the best solution over 500 belonging to a same cluster, and an estimated Ki of 75 µM.

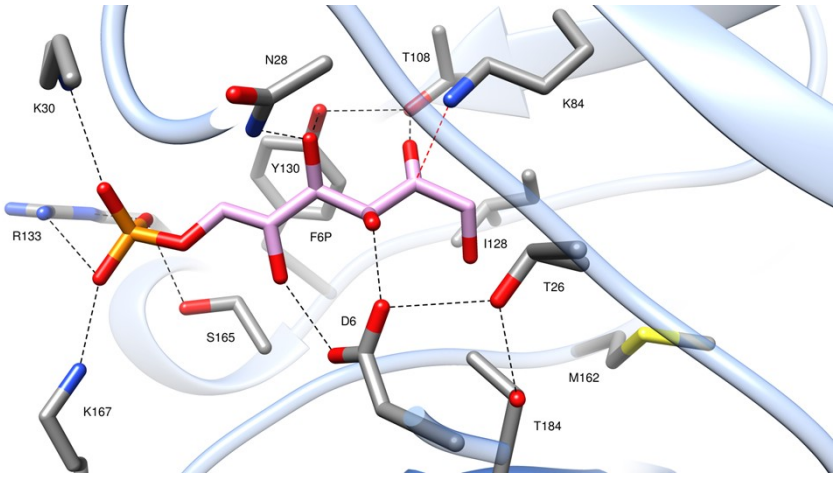


Figure S38. Docking of fructose-6-phosphate (pink) in FSA_{AciBac} active site with Autodock 4.2™, leading to a binding energy of -7.7 kcal/mol for the best solution over 494 belonging to a same cluster, and an estimated Ki of 2.3 µM.

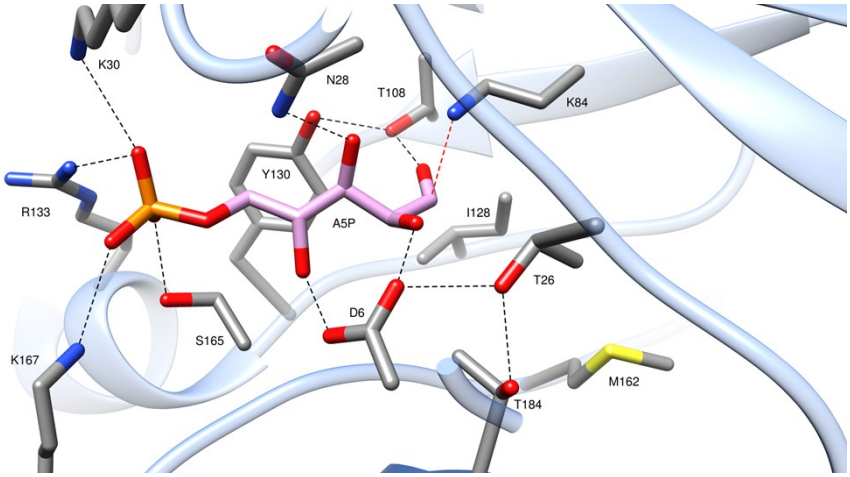


Figure S39. Docking of arabinose-5-phosphate (pink) in FSA_{AciBac} active site with Autodock 4.2™, leading to a binding energy of -8.3 kcal/mol for the best solution over 500 belonging to a same cluster, and an estimated Ki of 781 nM.

4. Electrophiles study

40 mM of hydroxyacetone (HA) and 60 mM of the studied electrophile in solution in water at pH 7.5 were placed in a vial, a blank was collected and 2 mg of enzyme (FSA_{AcAeo}, FSA_{AciBac} or FSA_{Ecoli}) were added. Samples were taken at regular intervals and HA content was determined using glycerol dehydrogenase, allowing spectroscopic monitoring of the reaction at 340 nm.

5. Syntheses

5.1. D-threose from glycolaldehyde

150 mg (500 mM) of GoA in solution in 5 mL of water were placed in a vial. The pH of the medium was adjusted to 8.3 and 10 mg of FSA_{AciBac} were added. In parallel, for comparison,

60 mg of GoA (500 mM) in solution in 2 mL of water were placed in a vial and the pH of the medium was adjusted to 8.4 and 4.2 mg of FSA_{Ecoli} were then added. The reactions were stirred at room temperature. The reaction was complete after 4 h for FSA_{AciBac} whereas for FSA_{Ecoli} the reaction was almost complete after 24 h (see NMR spectrum). After completion, FSA_{AciBac} was discarded via IMAC technique. The pH was then adjusted to 5 and the mixture was freeze-dried. 151.8 mg of a crystalline solid was obtained. The reaction was quantitative, the product containing 2 mg of salts due to the addition of sodium hydroxide and HCl to balance the pH.

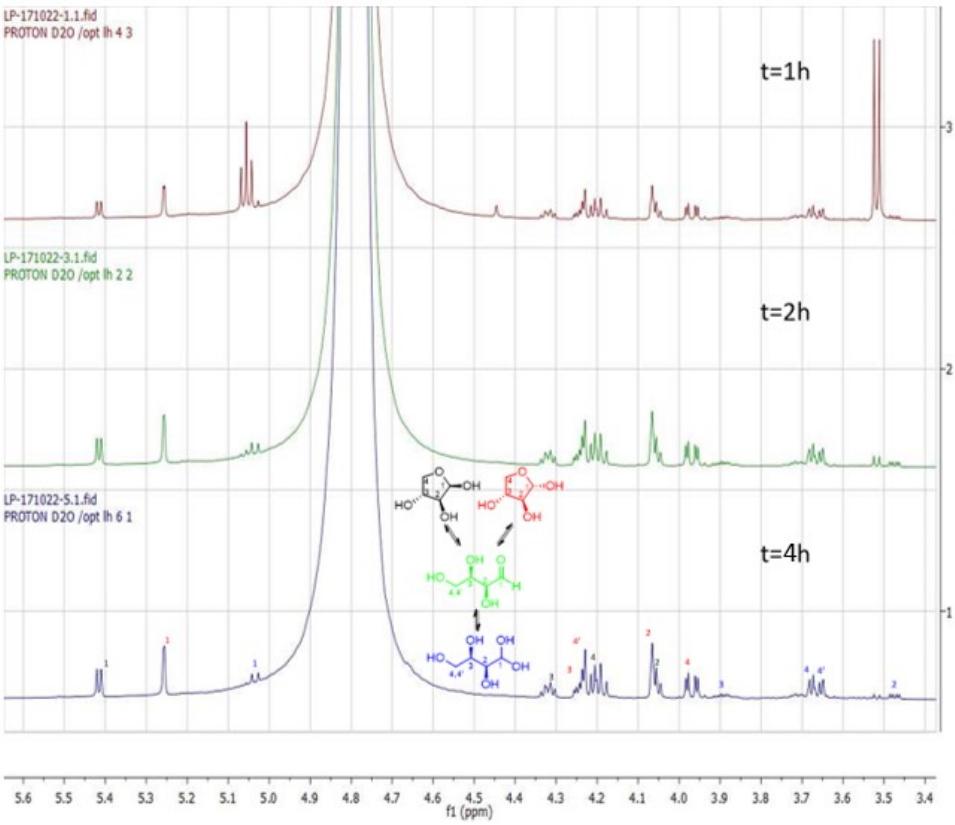


Figure S40 : ^{1}H NMR monitoring of D-threose formation with FSA_{AciBac}

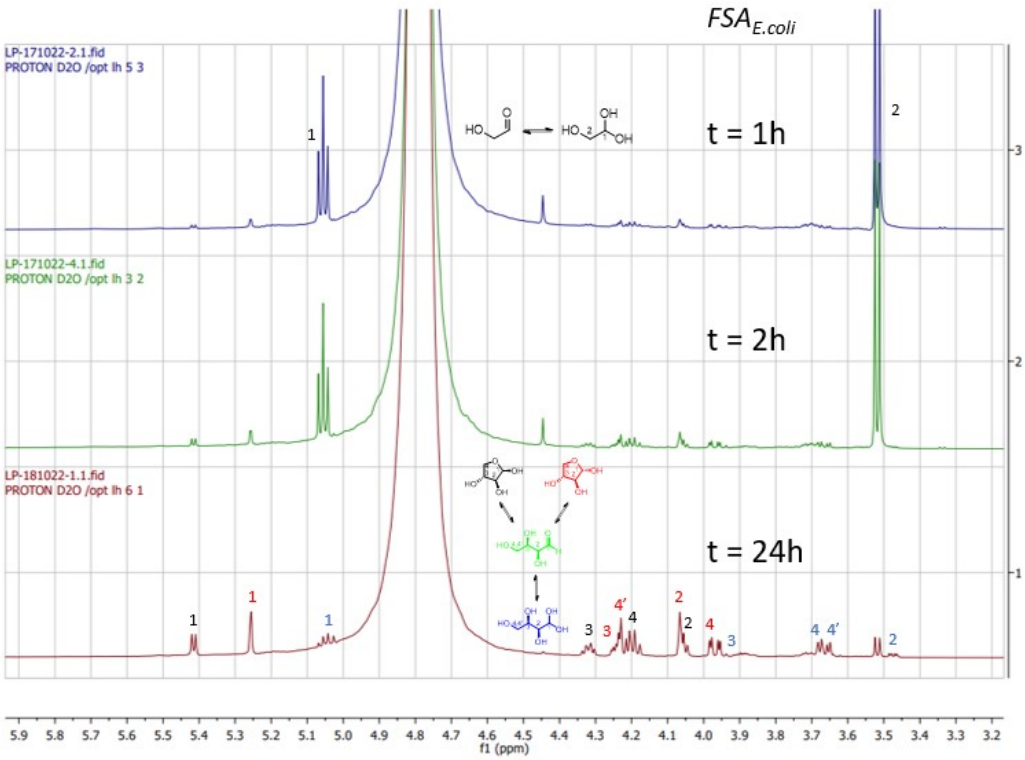


Figure S41 : ^{1}H NMR monitoring of D-threose formation with FSA$_{Ecoli}$

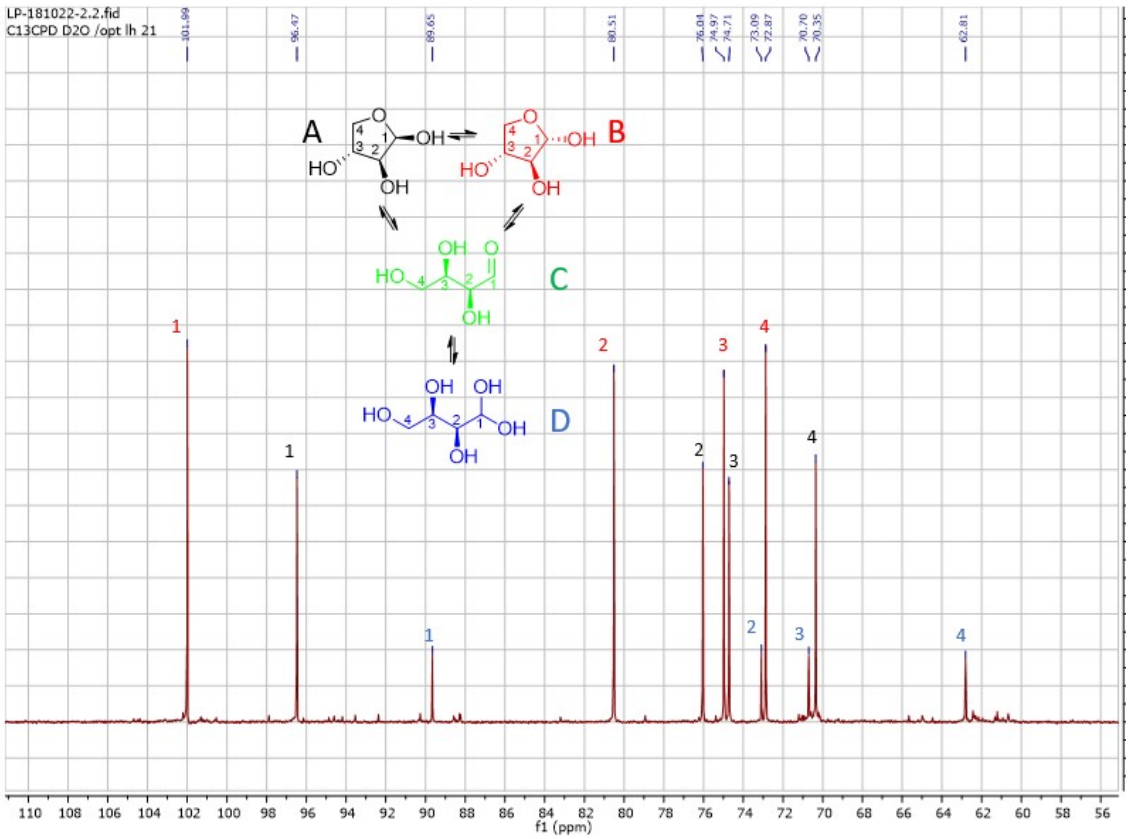


Figure S42 : ^{13}C NMR spectrum of D-threose obtained with FSA$_{AciBac}$

RMN ^{13}C (100 MHz, D_2O) δ 102.0 (C^{B1}), 96.5 (C^{A1}), 89.7 (C^{D1}), 80.5 (C^{B2}), 76.0 (C^{A2}), 75.0 (C^{B3}), 74.7 (C^{A3}), 73.1 (C^{D2}), 72.9 (C^{B4}), 70.7 (C^{D3}), 70.4 (C^{A4}), 62.8 (C^{D4}).

5.2. 2-deoxy-D-ribose-5-phosphate from acetaldehyde + D-glyceraldehyde-3-phosphate

In a flask were placed 600 µL of an acetaldehyde solution (1 M, final 100 mM), 465 µL of a DHAP solution (final 25 mM) and 5.1 mL of water. The pH of the solution was then adjusted to 7.5 and 100 µL were taken to provide a blank. 25 mg of FSA_{AciBac} and 30 µL of TPI (545 U) were then added. After 2 h of reaction at room temperature, 12 mg of FSA_{AciBac} and 200 µL of acetaldehyde solution (1 M) were added. After additional 4 h of reaction at room temperature, 10 mg of FSA_{AciBac} enzyme and 200 µL of acetaldehyde solution (1 M) were added. After 6h and 23 h of reaction, 200 µL of acetaldehyde solution (1 M) was added to prevent acetaldehyde evaporation. An 84% disappearance of DHAP was achieved in 26 h. Then, the enzyme was discarded by IMAC technique and the solution lyophilized at pH 7.5. The solid was then washed with acetone, recovered by centrifugation and washed with ethanol. The solid was then taken up in water, passed through an ultra-centrifugal filter (Amicon™) and lyophilized again. 85 mg of product was obtained. To accurately determine the amount of deoxyribose-5-phosphate, the solid was taken up in 1 mL of water and the deoxyribose-5-phosphate content determined using deoxyribose-5-phosphate aldolase (DERA) by UV-visible spectroscopy. 77 µmol of product were found, corresponding to a 50% yield.

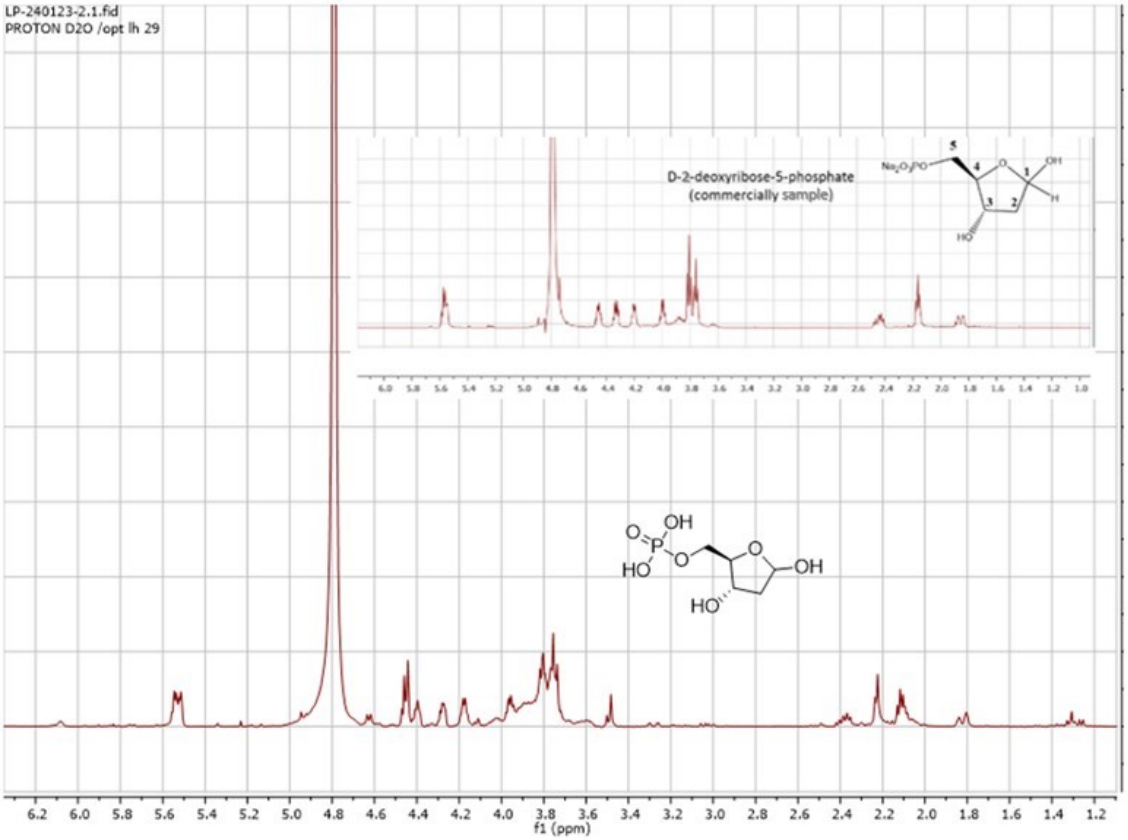


Figure S43 : ^{1}H NMR of deoxyribose-5-phosphate obtained

5.3. (3*S*,4*S*) 5-chloro-3,4-dihydroxypentan-2-one from HA + chloroacetaldehyde

30 µL (0.26 mmol) of a 55% chloroacetaldehyde solution in water, 26 µL of HA (0.31 mmol) and 2.5 mL of water were placed in a flask. The pH of the solution was adjusted to 8.2 and 100 µL was taken as a blank. 361 µL of a FSA_{AcAeo} solution (in ammonium sulphate 29 mg/mL) was centrifuged, the ammonium sulphate solid removed, and the supernatant containing the enzyme was added (final enzyme concentration: 4 mg/mL). After 12 h of reaction, the protein was removed on an ultra-centrifugal filter (Amicon™) and the mixture evaporated at reduced pressure. 52.9 mg of product were obtained as a white/translucent solid. Since the compound tended to sublimate easily, to determine the salts quantity, the compound was totally sublimated by lyophilization leading to 19.4 mg of salts thus corresponding to an 85% yield (33.5 mg).

RMN ^{1}H (400 MHz, D2O) δ 4,54 (d, J = 2.0 Hz, 1H, H^3), 4,38 (m, 1H, H^4), 3.71 (m, 2H, H^5, $H^{5'}$), 2,31 (s, 3H, H^1).

RMN ^{13}C (100 MHz, D2O) δ 212.5 (C^2), 77.0 (C^3), 71.4 (C^4), 44.5 (C^5), 25.7 (C^1).

HRMS ESI-: m/z calcd. for $[C_5H_9ClO_3, HCO_2]^-$ = 197.0222; found 197.0210.

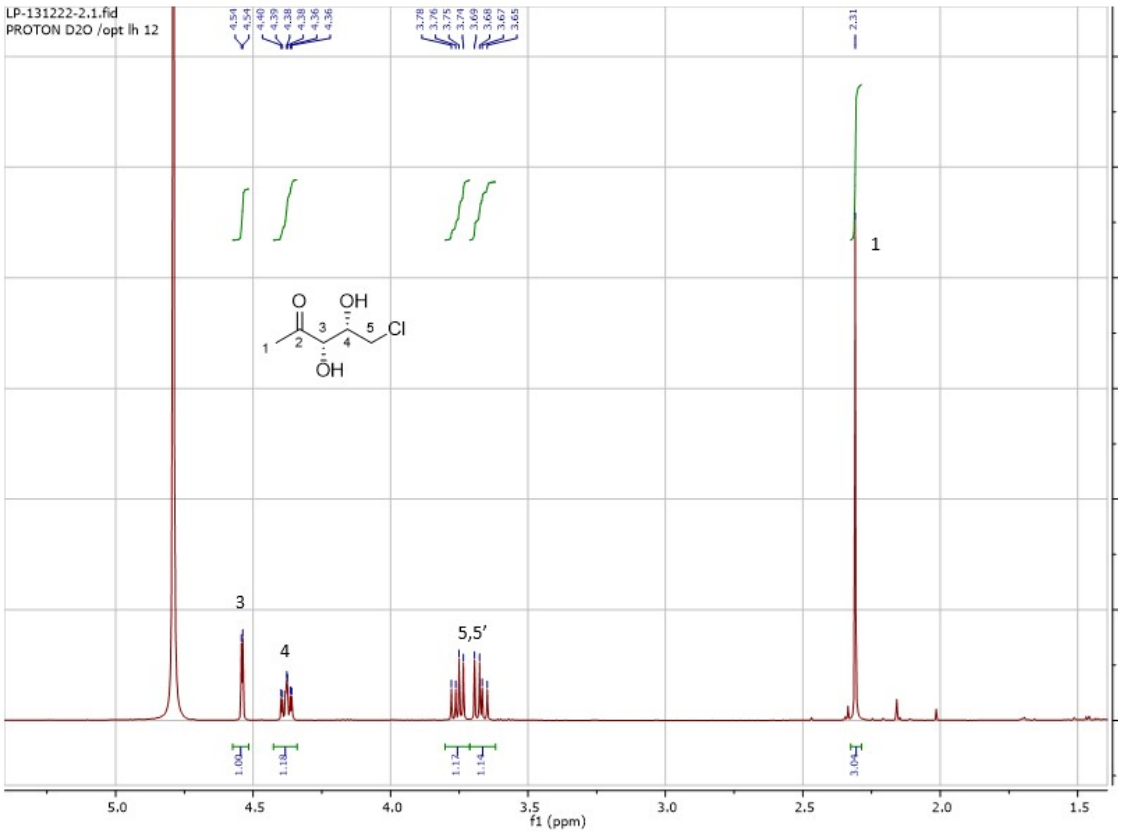


Figure S44 : ^{1}H NMR of the product obtained by coupling HA+chloroacetaldehyde with FSA$_{AcAeo}$

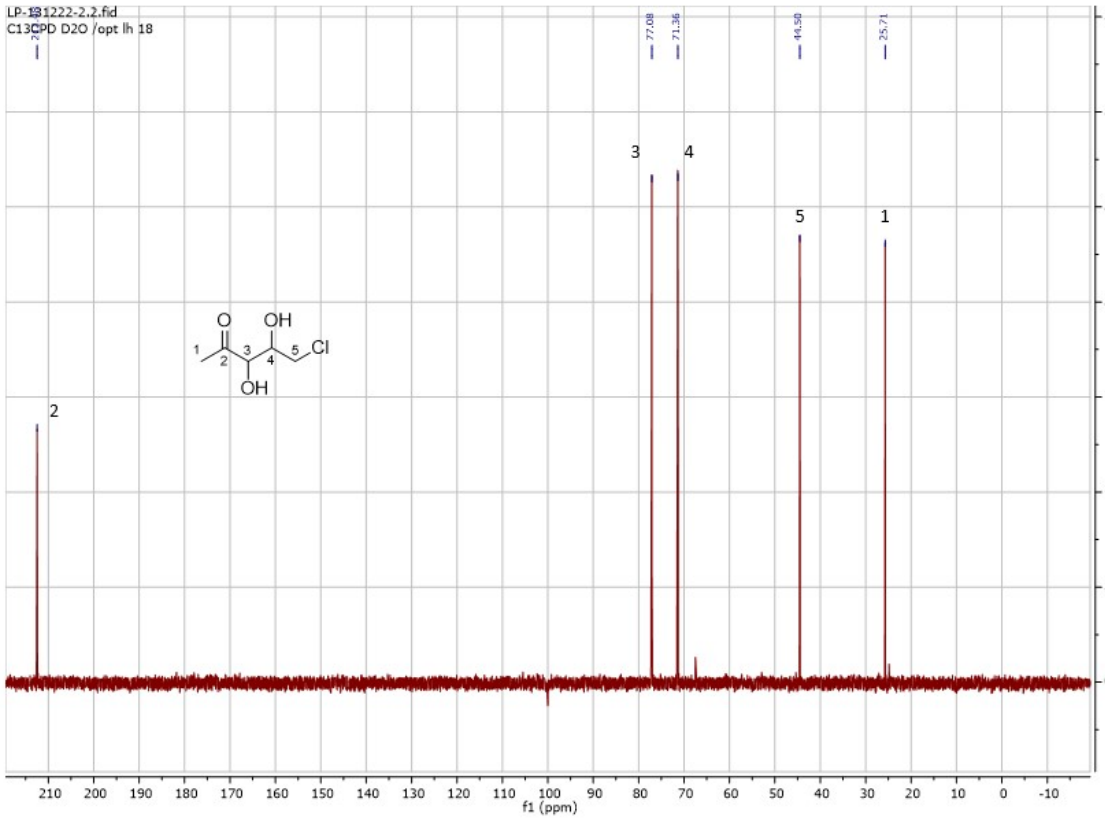


Figure S45 : ^{13}C NMR of the product obtained by coupling HA+chloroacetaldehyde with FSA$_{AcAeo}$

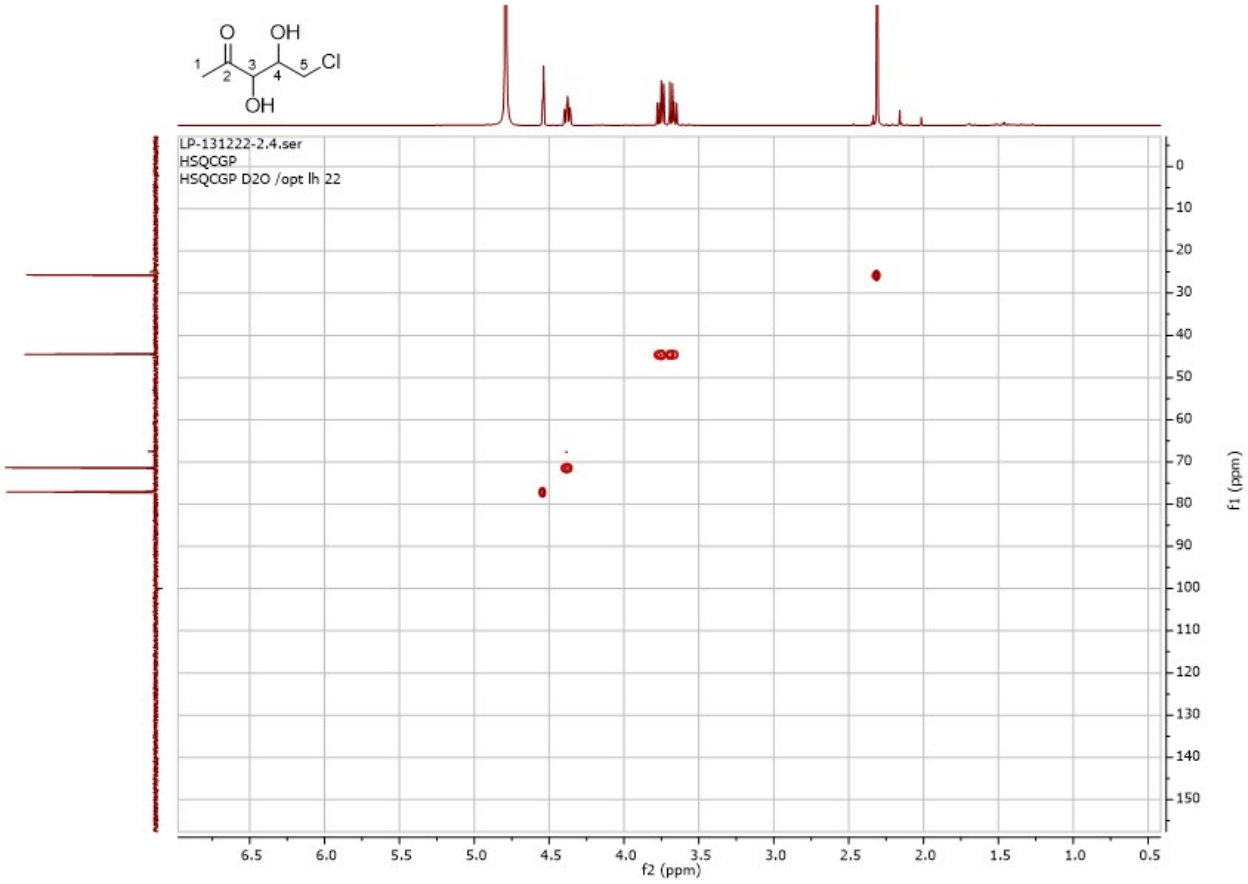


Figure S46 : HSQC of the product obtained by coupling HA+chloroacetaldehyde with FSA$_{AcAeo}$

5.4. (3*S*,4*S*) 3,4-dihydroxypentanal from acetaldehyde + L-lactaldehyde

A flask was filled with 4.5 mL of a solution of L-lactaldehyde (93 mM, final 50 mM), 485 µL of an acetaldehyde solution (1.75 M, final 100 mM) and 3.4 mL of water. The pH was then adjusted to 7.9 and 100 µL were collected for a blank. 1.4 g of cells were added after being washed 3 times with water to remove the glycerol added for conservation. After 48 h, the mixture was centrifuged at 8000 rpm for 20 min and the supernatant collected. The cells were washed with 2x15 mL of water and the mixture centrifuged. The supernatants were combined and evaporated *in vacuo*. Purification was performed on a silica gel column (eluent: AcOEt). 30.9 mg of product were obtained (62% yield).

RMN ^{1}H (400 MHz, D2O) δ 5.57 (t, *J* = 4.7 Hz, 1H, H^{1}), 4,25 (m, 1H, H^{3}), 4.21-4.14 (m, 1H, H^{4}), 2.15 (m, *2*H, H^{2}, H$^{2'}$), 1.16 (d, 3H, H^{5}).

RMN ^{13}C (major form) (100 MHz, D2O) δ 96.8 (C^{1}), 77.5 (C^{4}), 72.5 (C^{3}), 42.4 (C^{2}), 12.9 (C^{5}).

RMN ^{13}C (minor form) (100 MHz, D2O) δ 97.2 (C^{1}), 79.3 (C^{4}), 71.0 (C^{3}), 41.2 (C^{2}), 14.1 (C^{5}).

HRMS ESI-: m/z calcd. for $[C_5H_{10}O_3, HCO_2]^-$ = 163.0612; found 163.0600.

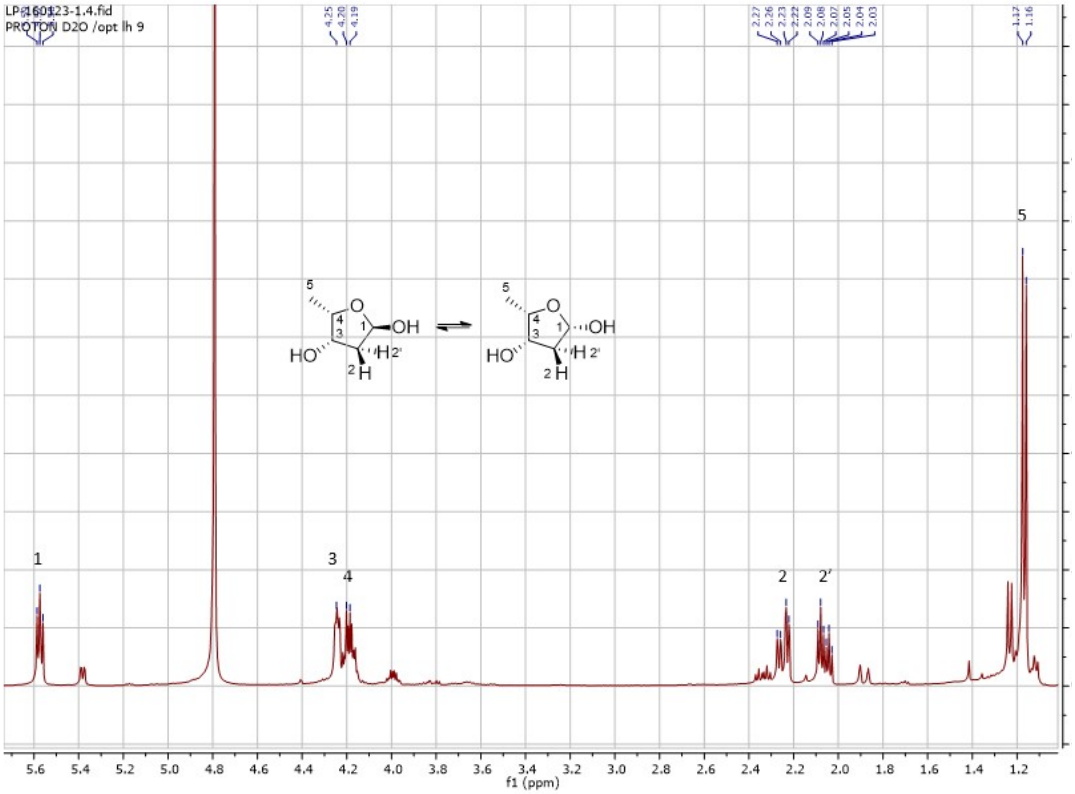


Figure S47 : ^{1}H NMR of the product obtained by coupling acetaldehyde+L-lactaldehyde with FSA_{AcAeo}

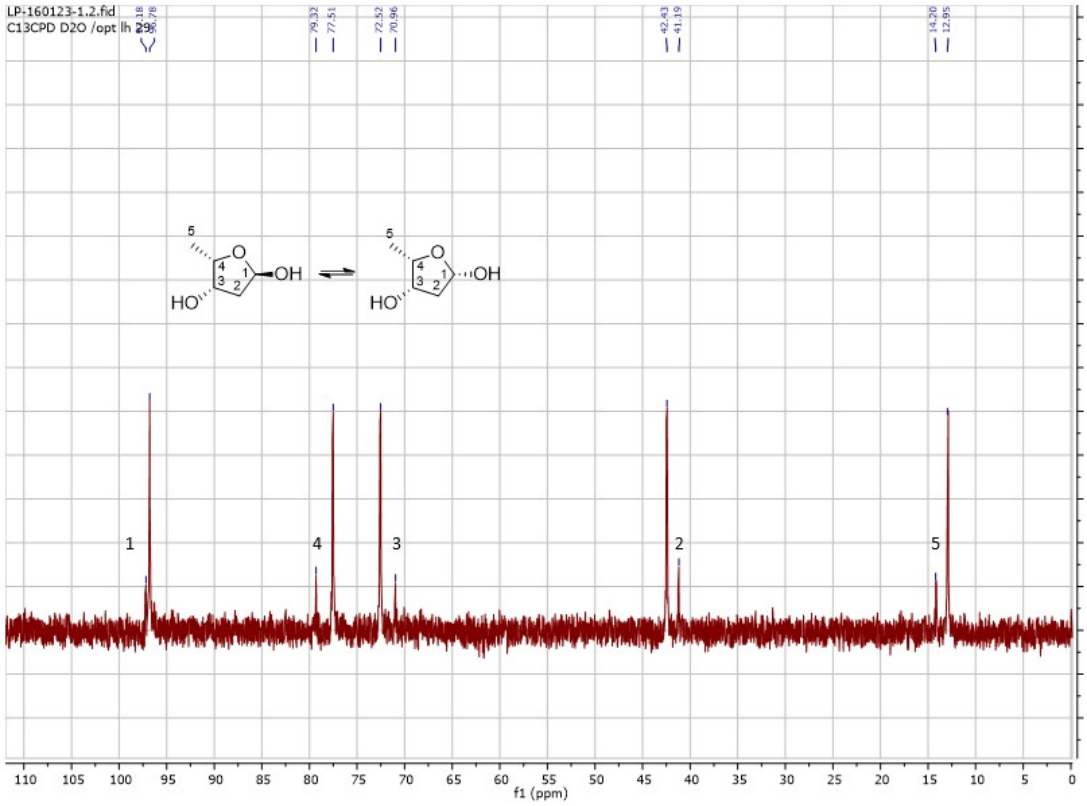


Figure S48 : ^{13}C NMR of the product obtained by coupling acetaldehyde+L-lactaldehyde with FSA$_{AcAeo}$

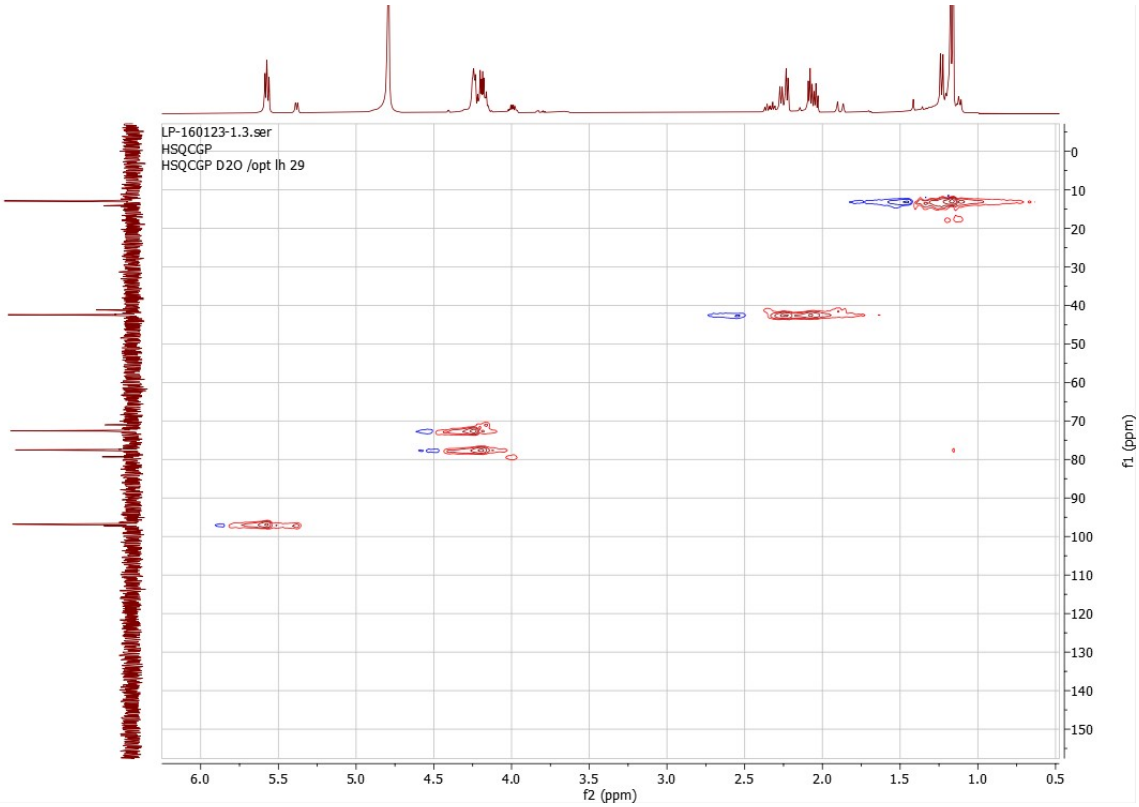


Figure S49 : HSQC of the product obtained by coupling acetaldehyde+L-lactaldehyde with FSA$_{AcAeo}$

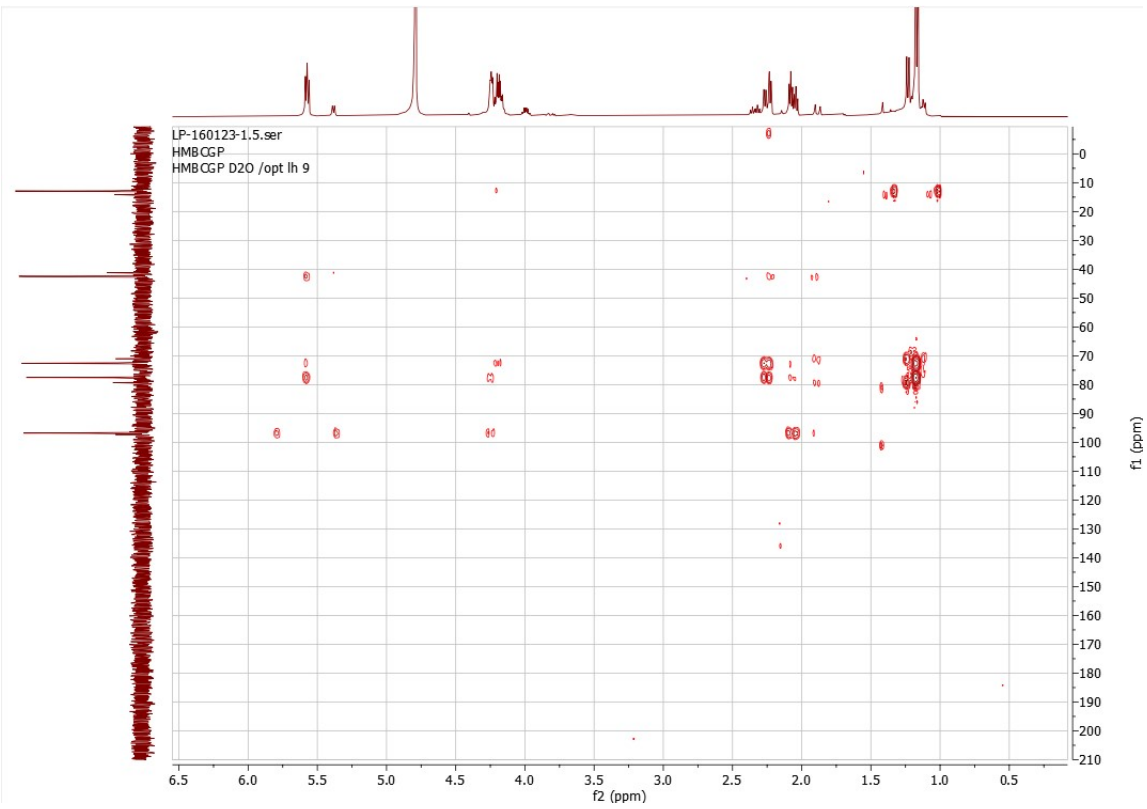


Figure S50 : HMBC of the product obtained by coupling acetaldehyde+L-lactaldehyde with FSA_{AcAeo}

References

(1) Schneider, S.; Gutiérrez, M.; Sandalova, T.; Schneider, G.; Clapés, P.; Sprenger, G. A.; Samland, A. K. Redesigning the Active Site of Transaldolase TalB from Escherichia Coli: New Variants with Improved Affinity towards Nonphosphorylated Substrates. *ChemBioChem* **2010**, *11* (5), 681–690. https://doi.org/10.1002/cbic.200900720.

(2) Vergne-Vaxelaire, C.; Bordier, F.; Fossey, A.; Besnard-Gonnet, M.; Debard, A.; Mariage, A.; Pellouin, V.; Perret, A.; Petit, J.-L.; Stam, M.; Salanoubat, M.; Weissenbach, J.; De Berardinis, V.; Zaparucha, A. Nitrilase Activity Screening on Structurally Diverse Substrates: Providing Biocatalytic Tools for Organic Synthesis. *Adv. Synth. Catal.* **2013**, *355* (9), 1763–1779. https://doi.org/10.1002/adsc.201201098.

(3) Paulat, L.; Laurent, V.; Hélaine, V.; Théveniot, M.; Petit, J.; Lemaire, M.; Delmas, V.; Bouzon, M.; De Berardinis, V.; Guérard-Hélaine, C. Insights on DHAP Aldolases Ability to Convert Dioxygen or a Ketone as Electrophile: Use of a Strain Depleted in Triose Phosphate Isomerase. *ChemCatChem* **2024**, e202400202. https://doi.org/10.1002/cctc.202400202.

(4) Thorell, S.; Schürmann, M.; Sprenger, G. A.; Schneider, G. Crystal Structure of Decameric Fructose-6-Phosphate Aldolase from Escherichia Coli Reveals Inter-Subunit Helix Swapping as a Structural Basis for Assembly Differences in the Transaldolase Family. *J. Mol. Biol.* **2002**, *319* (1), 161–171. https://doi.org/10.1016/S0022-2836(02)00258-9.

(5) Lehwess-Litzmann, A.; Neumann, P.; Parthier, C.; Lüdtke, S.; Golbik, R.; Ficner, R.; Tittmann, K. Twisted Schiff Base Intermediates and Substrate Locale Revise Transaldolase Mechanism. *Nat. Chem. Biol.* **2011**, *7* (10), 678–684. https://doi.org/10.1038/nchembio.633.

(6) Samland, A. K.; Baier, S.; Schürmann, M.; Inoue, T.; Huf, S.; Schneider, G.; Sprenger, G. A.; Sandalova, T. Conservation of Structure and Mechanism within the Transaldolase Enzyme Family. *FEBS J.* **2012**, *279* (5), 766–778. https://doi.org/10.1111/j.1742-4658.2011.08467.x.
(7) Stellmacher, L.; Sandalova, T.; Leptihn, S.; Schneider, G.; Sprenger, G. A.; Samland, A. K. Acid-Base Catalyst Discriminates between a Fructose 6-Phosphate Aldolase and a Transaldolase. *ChemCatChem* **2015**, *7* (19), 3140–3151. https://doi.org/10.1002/cctc.201500478.
(8) Sautner, V.; Friedrich, M. M.; Lehwess-Litzmann, A.; Tittmann, K. Converting Transaldolase into Aldolase through Swapping of the Multifunctional Acid–Base Catalyst: Common and Divergent Catalytic Principles in F6P Aldolase and Transaldolase. *Biochemistry* **2015**, *54* (29), 4475–4486. https://doi.org/10.1021/acs.biochem.5b00283.
(9) Yang, X.; Wu, L.; Li, A.; Ye, L.; Zhou, J.; Yu, H. The Engineering of Decameric D -Fructose-6-Phosphate Aldolase A by Combinatorial Modulation of Inter- and Intra-Subunit Interactions. *Chem. Commun.* **2020**, *56* (55), 7561–7564. https://doi.org/10.1039/D0CC02437F.
(10) Webb, B.; Sali, A. Comparative Protein Structure Modeling Using MODELLER. *Curr. Protoc. Bioinforma.* **2016**, *54* (1). https://doi.org/10.1002/cpbi.3.
(11) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera?A Visualization System for Exploratory Research and Analysis. *J. Comput. Chem.* **2004**, *25* (13), 1605–1612. https://doi.org/10.1002/jcc.20084.
(12) Varadi, M.; Anyango, S.; Deshpande, M.; Nair, S.; Natassia, C.; Yordanova, G.; Yuan, D.; Stroe, O.; Wood, G.; Laydon, A.; Žídek, A.; Green, T.; Tunyasuvunakool, K.; Petersen, S.; Jumper, J.; Clancy, E.; Green, R.; Vora, A.; Lutfi, M.; Figurnov, M.; Cowie, A.; Hobbs, N.; Kohli, P.; Kleywegt, G.; Birney, E.; Hassabis, D.; Velankar, S. AlphaFold Protein Structure Database: Massively Expanding the Structural Coverage of Protein-Sequence Space with High-Accuracy Models. *Nucleic Acids Res.* **2022**, *50* (D1), D439–D444. https://doi.org/10.1093/nar/gkab1061.
(13) Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *J. Comput. Chem.* **2009**, *30* (16), 2785–2791. https://doi.org/10.1002/jcc.21256.