

## Electronic Supplementary Information (ESI)

### Cross-acyloin condensation of aldehydes catalyzed by Transketolase variants for the synthesis of aliphatic $\alpha$ -hydroxyketones

Giuseppe Arbia,<sup>a</sup> Camille Gadona,<sup>a</sup> Hubert Casajus,<sup>a</sup> Lionel Nauton,<sup>a</sup> Franck Charmantray,<sup>a\*</sup> Laurence Hecquet<sup>a\*</sup>

<sup>a</sup>Université Clermont Auvergne, CNRS, Clermont Auvergne INP, Institut de Chimie de Clermont-Ferrand (ICCF) F-63000 Clermont-Ferrand (France). Email : laurence.hecquet@uca.fr, franck.charmantray@uca.fr

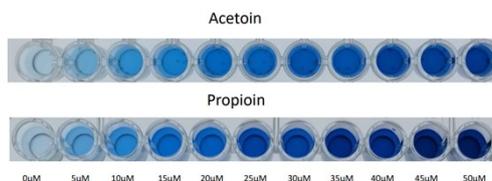
#### Table of Contents

<b>ESI 1 : determination of acetoin and propion standard curves with WST-1 assay</b> .....	2
<b>ESI 2 : screening of 4 TK<sub>gst</sub> variants for acyloin reaction</b> .....	3
<b>ESI 3 : monitoring of preparative scale syntheses of products 7-13 with the best TK<sub>gst</sub> variants</b> .....	7
3-hydroxybutan-2-one <b>7</b> synthesis .....	7
4-hydroxyhexan-3-one <b>8</b> synthesis .....	8
1,2-dihydropentan-3-one <b>9</b> synthesis .....	9
1,2-dideoxy-D-threo-hex-3-ulose <b>10</b> synthesis .....	10
1,2-dideoxy-D-arabino-hept-3-ulose <b>11</b> synthesis .....	11
1,2-dihydroxy-4-methylpentan-3-one <b>12</b> synthesis (protocol A).....	12
1,2-dihydroxy-4-methylpentan-3-one <b>12</b> synthesis (protocol B).....	12
4,5,6-trihydroxy-2-methylhexan-3-one <b>13</b> synthesis (protocol B).....	13
<b>ESI 4 : Enantiomeric excess determination</b> .....	14
<b>ESI 5 : NMR spectra of compounds 7-13</b> .....	15
<sup>1</sup> H NMR spectrum of 3-hydroxybutan-2-one <b>7</b> .....	15
<sup>13</sup> C NMR spectrum of 3-hydroxybutan-2-one <b>7</b> .....	15
<sup>1</sup> H NMR spectrum of 4-hydroxyhexan-3-one <b>8</b> .....	16
<sup>13</sup> C NMR spectrum of 4-hydroxyhexan-3-one <b>8</b> .....	16
<sup>1</sup> H NMR spectrum of 1,2-dihydropentan-3-one <b>9</b> .....	17
<sup>13</sup> C NMR spectrum of 1,2-dihydropentan-3-one <b>9</b> .....	17
<sup>1</sup> H NMR spectrum of 1,2-dideoxy-D-threo-hex-3-ulose <b>10</b> .....	18
<sup>13</sup> C NMR spectrum of 1,2-dideoxy-D-threo-hex-3-ulose <b>10</b> .....	18
<sup>1</sup> H NMR spectrum of 1,2-dideoxy-D-arabino-hept-3-ulose <b>11</b> .....	19
<sup>13</sup> C NMR spectrum of 1,2-dideoxy-D-arabino-hept-3-ulose <b>11</b> .....	19
<sup>1</sup> H NMR spectrum of 1,2-dihydroxy-4-methylpentan-3-one <b>12</b> .....	20
<sup>13</sup> C NMR spectrum of 1,2-dihydroxy-4-methylpentan-3-one <b>12</b> .....	20
<sup>1</sup> H NMR spectrum of 4,5,6-trihydroxy-2-methylhexan-3-one <b>13</b> .....	21
<sup>13</sup> C NMR spectrum of 4,5,6-trihydroxy-2-methylhexan-3-one <b>13</b> .....	21
<b>ESI 6 Molecular dynamics studies</b> .....	22
<b>ESI 7 Determination of E factors</b> .....	27

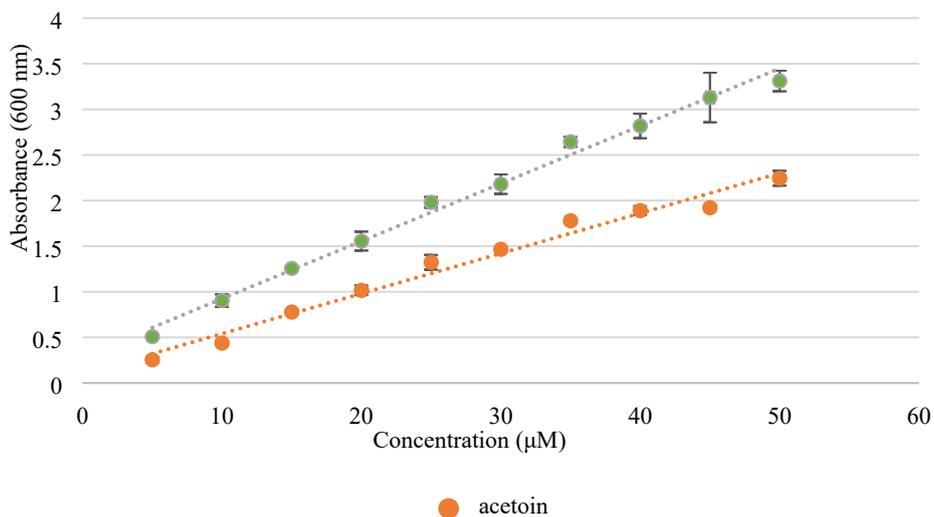
## ESI 1: Determination of acetoin 7 and propionin 8 standard curves with WST-1 assay

### Acetoin 7 and propionin 8 (5–50 $\mu\text{M}$ ) standard curves:

Standard curves were determined in the 5–50  $\mu\text{M}$  concentration range in 96-well plates using WST-1 assay. The experiments were performed in triplicate and standard curves were generated by linear regression analysis. The limits of detection (LOD) and of quantification (LOQ) were calculated according to the formula  $LOD=3.3Sb/a$  and  $LOQ=10Sb/a$ , where  $Sb$  is the standard deviation of the control and  $a$  is the slope of the standard curve.



**Figure S1** : colour range using WST-1 assay with 5–50  $\mu\text{M}$  concentration of acetoin 7 and propionin 8



**Figure S2** : standard curves determined in the 5–50  $\mu\text{M}$  concentration range of acetoin 7 and propionin 8 in 96-well plates using WST-1

Acyloins	LOD ( $\mu\text{M}$ )	LOQ ( $\mu\text{M}$ )	$\epsilon$ ( $\text{M}^{-1}.\text{cm}^{-1}$ )	$SD^a$ ( $\text{M}^{-1}.\text{cm}^{-1}$ )	Slope	y-intercept
Acetoin 7	3.6	10.7	81660	2600	0.044	0.1009
Propionin 8	6.7	11.5	112900	2480	0.063	0.2924

<sup>a</sup>SD: standard deviation

**Table S1** : limits of detection (LOD) and of quantification (LOQ) using WST-1 with acetoin 7 and propionin 8

## ESI 2 : screening of TK<sub>gst</sub> variants for acyloin reaction

### Screening of TK<sub>gst</sub> variants for self-acyloin condensation of 1-3 by <sup>1</sup>H *in situ* NMR

Reaction conditions : aldehydes 1-3 at 50 mM, 2 mg TK<sub>gst</sub> variant . mL<sup>-1</sup>, phosphate buffer 50 mM pH 7.0, 37°C. The reaction progress was determined by *in situ* <sup>1</sup>H NMR after 48 h of reaction time at 37 °C. Substrates: 1, ethanal; 2, propanal; 3 *iso*-butanal. A, blank reaction without enzyme; B, wt TK<sub>gst</sub>; C, TK<sub>gst</sub> variants H102L/L118I/H474S; D H102L/L118I/H474G, E H102L/L118I, F H102L/474S

Reaction monitoring : Aliquots of reaction mixtures (450 μL) were taken out from reaction mixture at 48 h and mixed with 50 μL of TSP-d<sub>4</sub> (50 mM, 8.5 mg . mL<sup>-1</sup> of D<sub>2</sub>O) before analysis by using quantitative *in situ* <sup>1</sup>H NMR relative to 3-trimethylsilyl-2,2,3,3-tetrauteropropionate (TSP-d<sub>4</sub>) as internal standard.

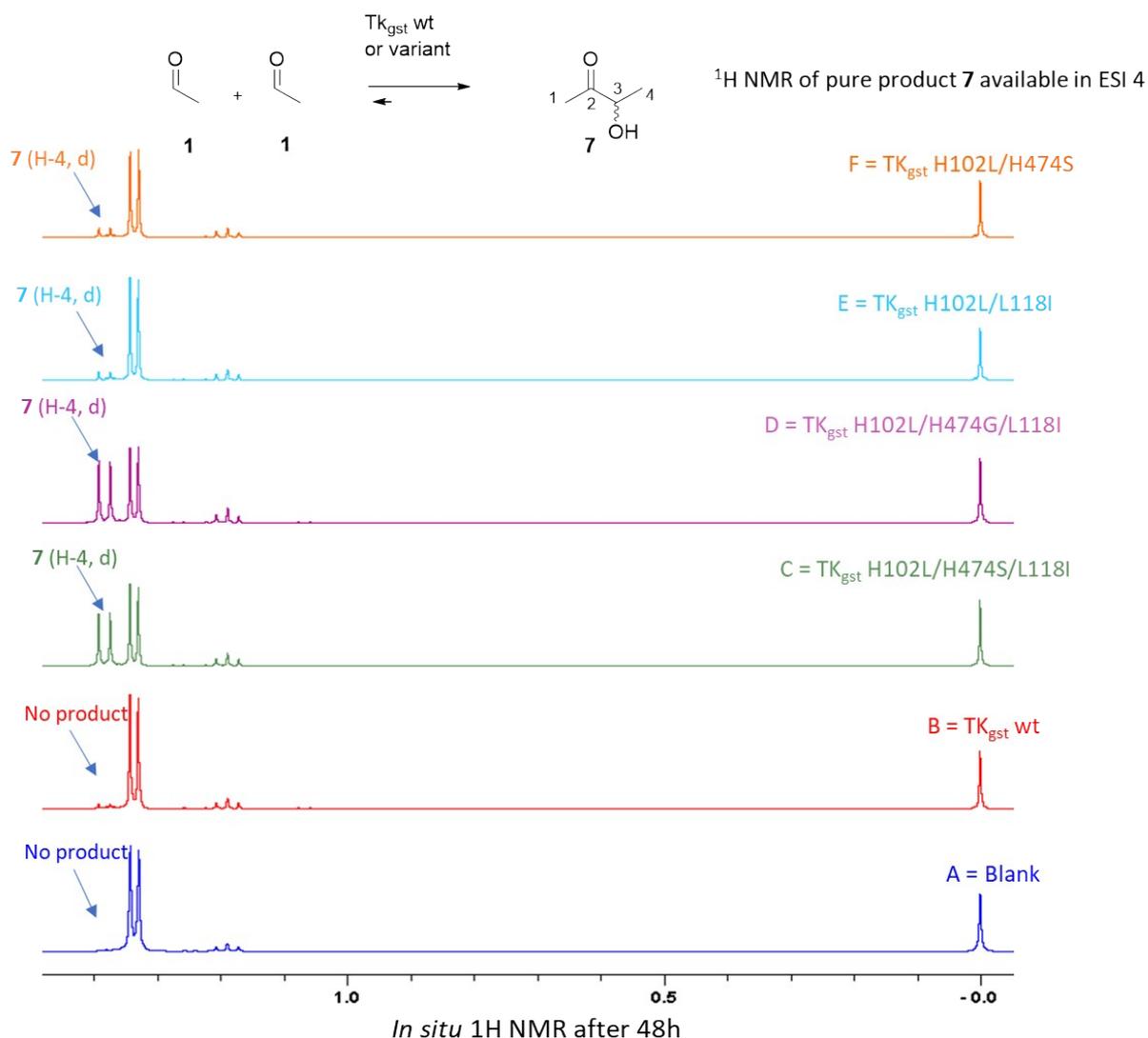
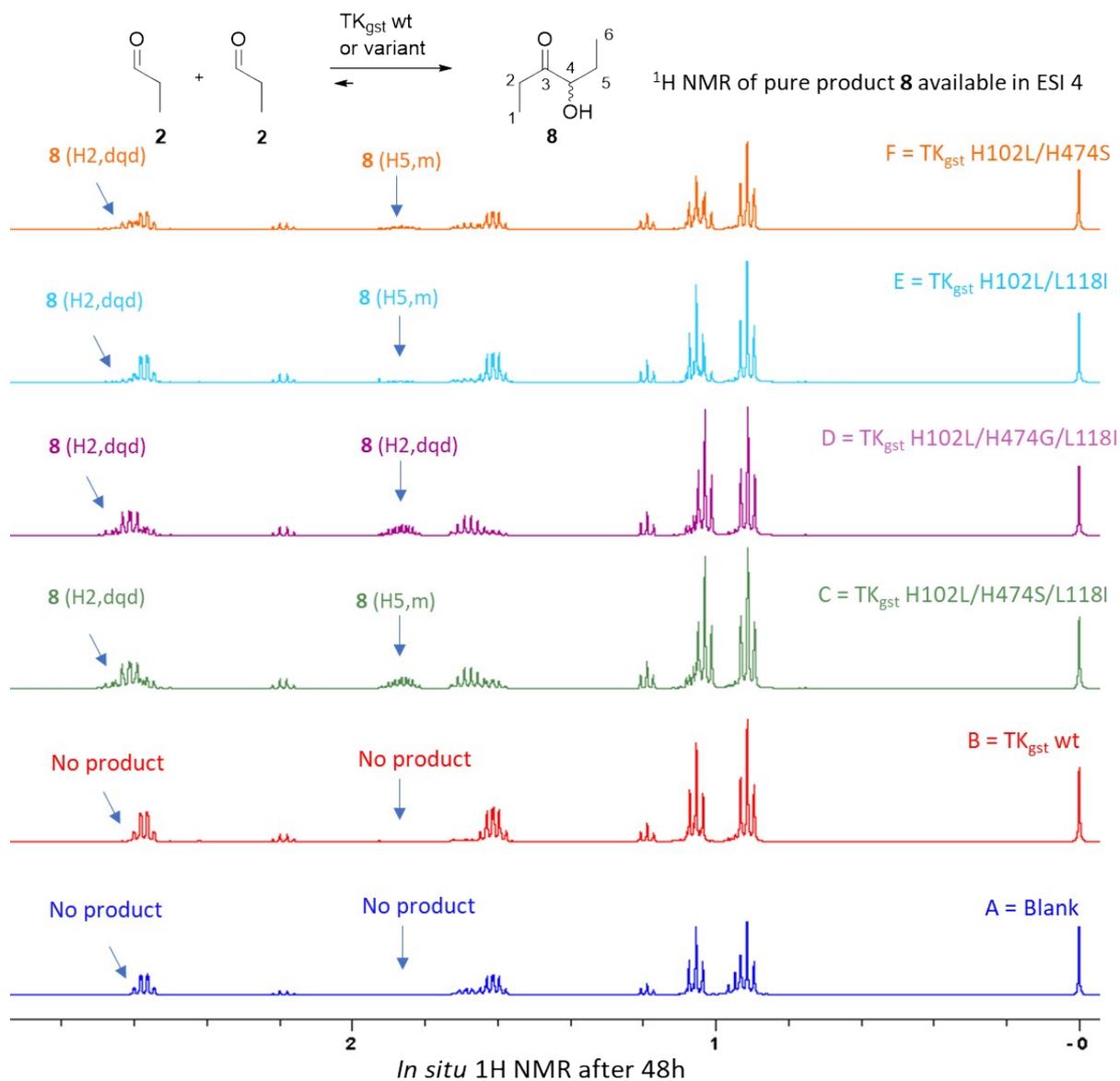
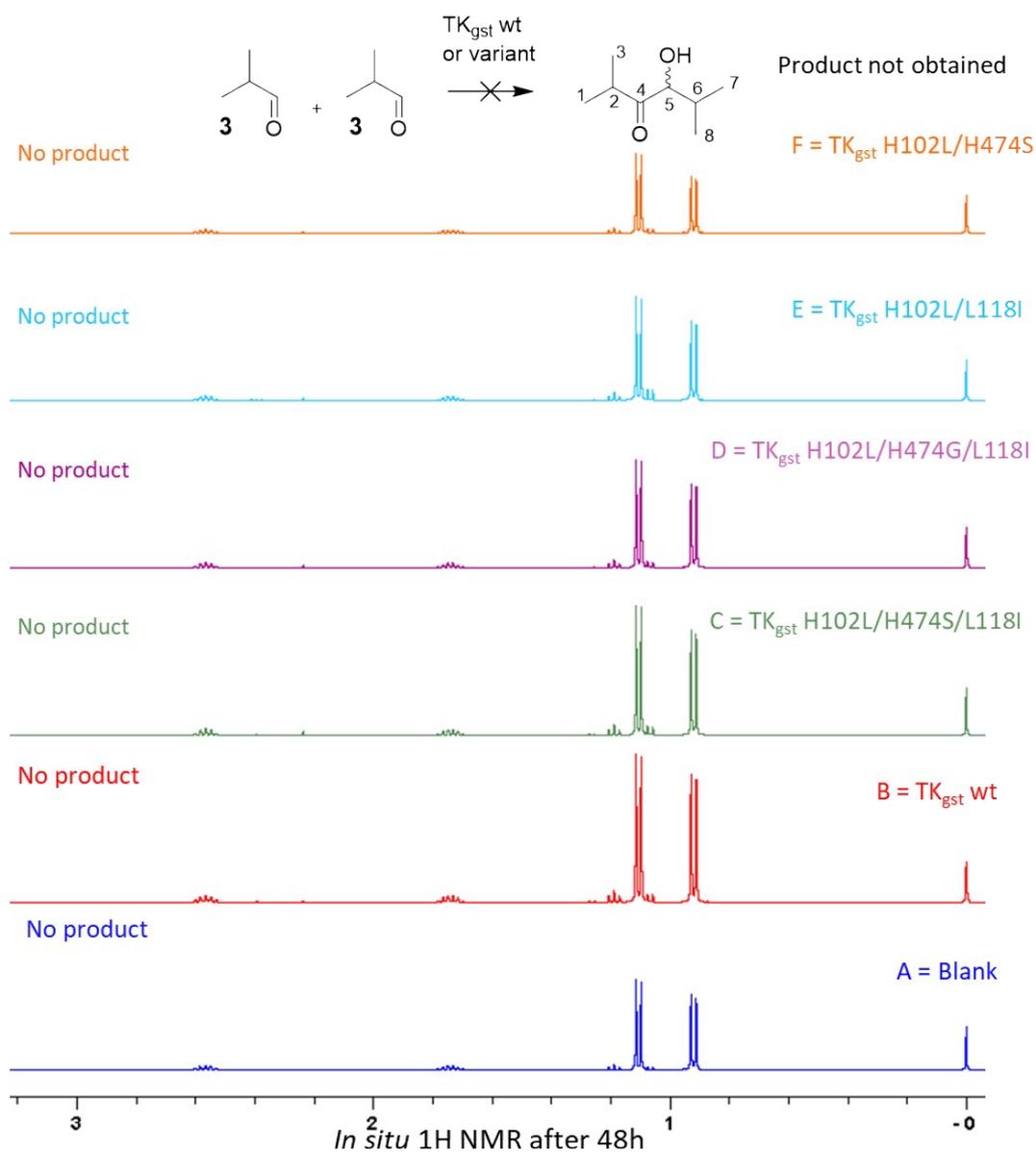


Figure S3 : monitoring by *in situ* NMR of product 7 synthesis

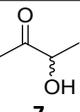
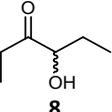


**Figure S4** : monitoring by *in situ* NMR of product **8** synthesis



**Figure S5** : monitoring by *in situ* NMR of *iso*-butanal **3** (self-condensation product not obtained)

**Table S2.** Synthesis of products **7** and **8** at 37°C using  $\text{TK}_{\text{gst}}$  variants in the presence of ethanal **1** and propanal **2** as nucleophiles (50 mM) and acceptors (50mM) for self-condensation reactions. Reactions were performed at 37 °C in phosphate buffer pH 7.0, 50 mM

Nucleophile	Electrophile	Product	Time (h)	$\text{TK}_{\text{gst}}$ variant	$\text{TK}_{\text{gst}}$ mg · mL <sup>-1</sup> <sub>1</sub>	<i>In situ</i> yield (%) <sup>a</sup>	Isolated yield (%)
 <b>1</b>	 <b>1</b>	 <b>7</b>	96	H102L/L118I/ H474G	3	86	67
 <b>2</b>	 <b>2</b>	 <b>8</b>	48	H102L/L118I/ H474G	1.5	90	85

### Screening of the TK<sub>gst</sub> variants for self and cross-acyloin condensation by using quantitative *in situ* (%) <sup>1</sup>H NMR

Reaction monitoring : aliquots of reaction mixtures (450 μL) were taken out from reaction mixture at 48 h and mixed with 50 μL of TSP-d<sub>4</sub> (50 mM, 8.5 mg . mL<sup>-1</sup> of D<sub>2</sub>O) before analysis by using quantitative *in situ* <sup>1</sup>H NMR relative to 3-trimethylsilyl-2,2,3,3-tetrauteropropionate (TSP-d<sub>4</sub>) as internal standard.

**Table S3.** TK<sub>gst</sub> variant substrate scope with nucleopiles **1**, **2** and **3** (50 mM) and electrophiles **1-6**. (50 mM) electrophile, 2 mg . mL<sup>-1</sup> TK<sub>gst</sub> variant of reaction, phosphate buffer 50 mM, pH 7.0, 37 °C. The reaction progress was determined by *in situ* <sup>1</sup>H NMR after 48 h of reaction time at 37 °C.

Nucleophile 1 ethanal			<i>In situ</i> yield of products determined by <sup>1</sup> H NMR			
Electrophiles	No TK	wt TK	H102L/L118I/H474S	H102L/L118I/H474G	H102L/L118I	H102L/H474S.
<b>1</b> ethanal	0	0	50	61	11	11
<b>4</b> glycolaldehyde	0	0	7	3	0	0
<b>5</b> D-glyceraldehyde	0	0	6	4	0	0
<b>6</b> D-erythrose	0	0	5	4	0	0
Nucleophile 2 propanal			<i>In situ</i> yield of products determined by <sup>1</sup> H NMR			
Electrophiles	No TK	wt TK	H102L/L118I/H474S	H102L/L118I/H474G	H102L/L118I	H102L/H474S
<b>2</b> propanal	0	0	92	98	22	60
<b>4</b> glycolaldehyde	0	0	86	90	9	7
<b>5</b> D-glyceraldehyde	0	0	35	24	4	5
<b>6</b> D-erythrose	0	0	70	70	21	22
Nucleophile 3 iso-butanal			<i>In situ</i> yield of products determined by <sup>1</sup> H NMR			
Electrophiles	No TK	wt TK	H102L/L118I/H474S	H102L/L118I/H474G	H102L/L118I	H102L/H474S
<b>3</b> iso-butanal;	0	0	0	0	0	0
<b>4</b> glycolaldehyde	0	0	31	18	9	4
<b>5</b> D-glyceraldehyde	0	0	15	10	0	0
<b>6</b> D-erythrose.	0	0	10	12	0	0

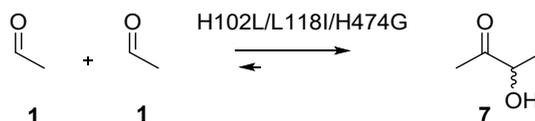
### ESI 3: monitoring of preparative scale syntheses of compounds 7-13 with the best TK<sub>gst</sub> variants

Reaction conditions: electrophile **1-6** at 50 mM and, nucleophile **1-3** at 50 mM, TK<sub>gst</sub> variant 1.5-3 mg · mL<sup>-1</sup>, phosphate buffer 50 mM pH 7.0, 37°C.

Reaction monitoring: aliquots of reaction mixtures (450 μL) were taken out from reaction mixture at several times (0-96h) and mixed with 50 μL of TSP-d<sub>4</sub> (50 mM, 8.5 mg · mL<sup>-1</sup> of D<sub>2</sub>O) before analysis by using quantitative in situ <sup>1</sup>H NMR relative to 3-trimethylsilyl-2,2,3,3-tetra-deuterio-propionate (TSP-d<sub>4</sub>) as internal standard.

#### Synthesis of 3-hydroxybutan-2-one **7**

Reaction conditions: ethanal **1** (100 mM), H102L/L118I/H474G 3 mg · mL<sup>-1</sup>, 37°C, phosphate buffer, pH 7.0, 96 h.



compound	<b>7</b>
<i>in situ</i> conc. (mM)	42
<i>in situ</i> yield (%)	84
Isolated yield (%)	67

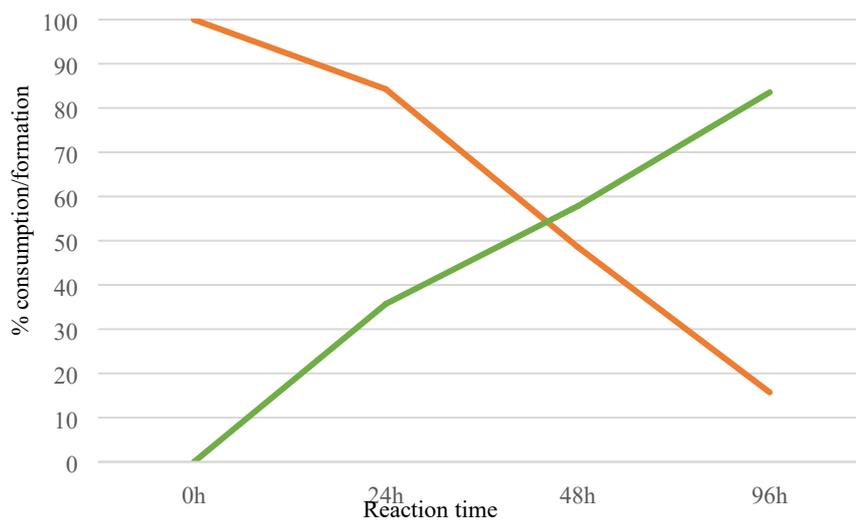
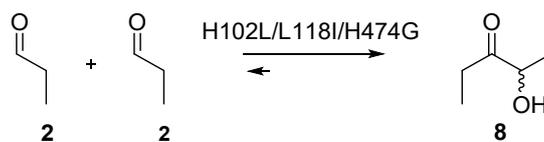


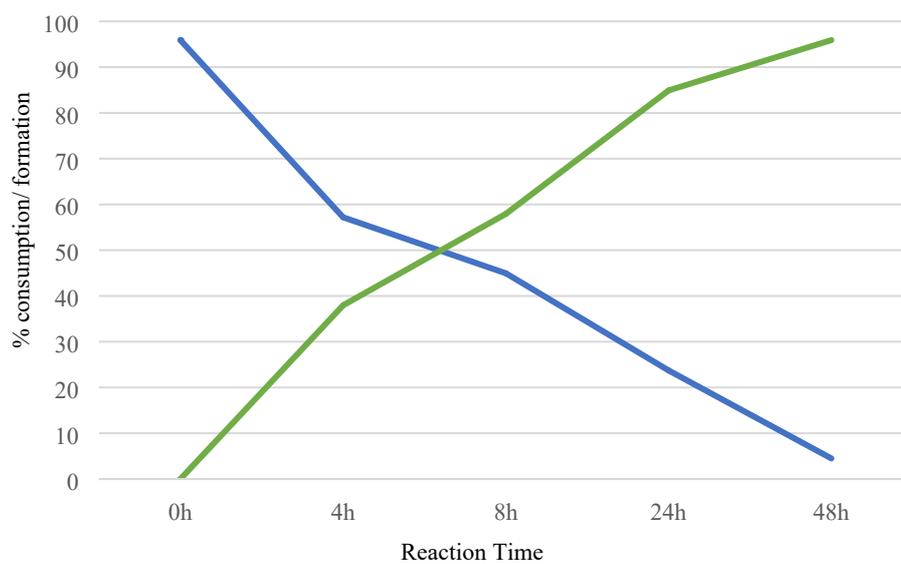
Figure S6 : Self-acyloin condensation progress of product **7** (■) from ethanal **1** (■)

### Synthesis of 4-hydroxyhexan-3-one **8**

Reaction conditions: propanal **2** (100 mM), H102L/L118I/H474G 1.5 mg . mL<sup>-1</sup>, phosphate buffer, pH 7.0, 48 h.



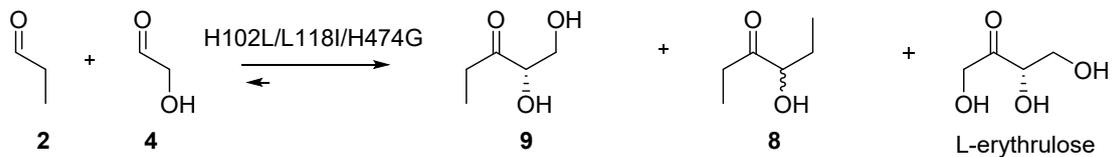
compounds	<b>8</b>
<i>in situ</i> conc. (mM)	48
<i>in situ</i> yield (%)	96
Isolated yield (%)	60



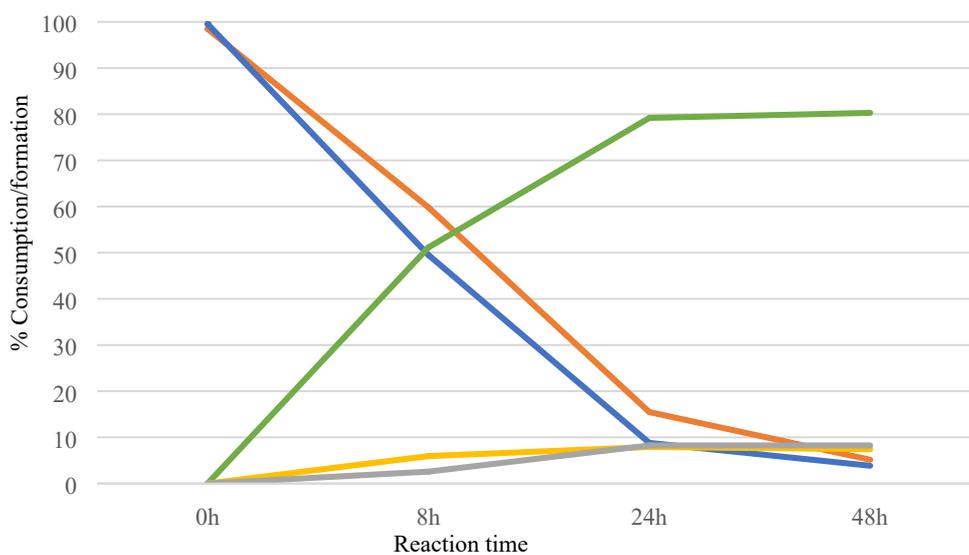
**Figure S7** : Self-acyloin condensation progress of product **8** (■) from propanal **2** (■)

### Synthesis of 1,2-dihydroxypentan-3-one **9**

Reaction conditions: propanal **2** (50 mM), GoA **4** (50 mM), H102L/L118I/H474G 2.5 mg · mL<sup>-1</sup>, phosphate buffer, pH 7.0, 48 h.



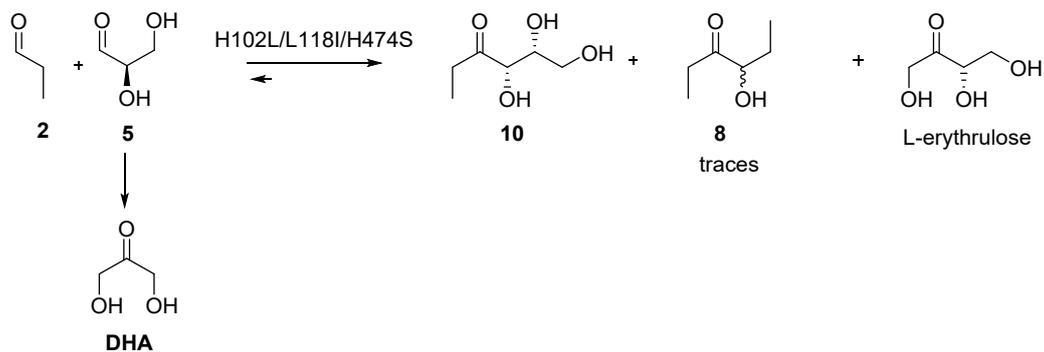
compounds	<b>9</b>	<b>8</b>	L-erythrulose
<i>in situ</i> conc. (mM)	40	4	4
<i>in situ</i> yield (%)	80	8	8
Isolated yield (%)	50	n.d (volatile)	n.d.



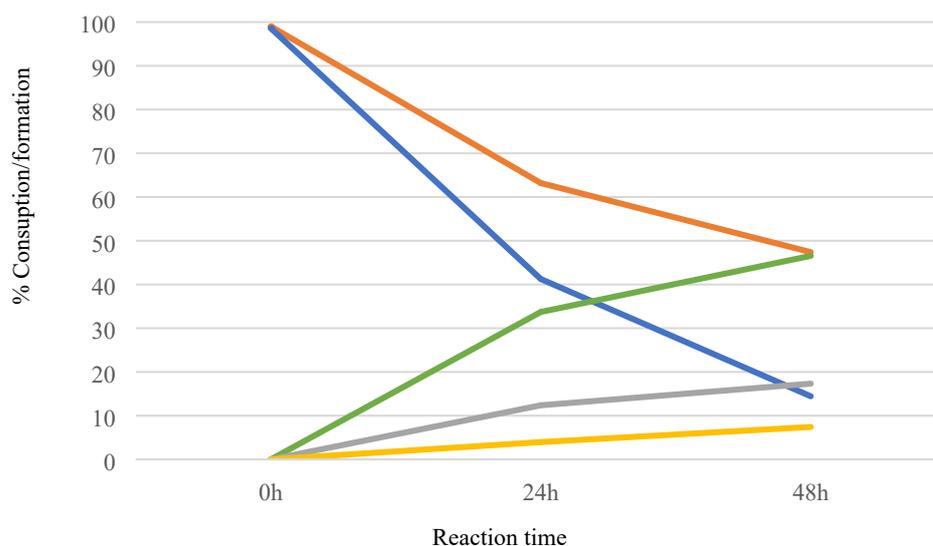
**Figure S8** : Cross-acyloin condensation progress of product **9** (■) from propanal **2** (■) and GoA **4** (■) giving by-products L-erythrulose (■) and propionol **8** (■)

### Synthesis of 1,2-dideoxy-D-threo-hex-3-ulose **10**

Reaction conditions: propanal **2** (50 mM), D-glyceraldehyde **5** (50 mM), H102L/L118I/H474S, 3 mg · mL<sup>-1</sup>, phosphate buffer, pH 7.0, 48 h.



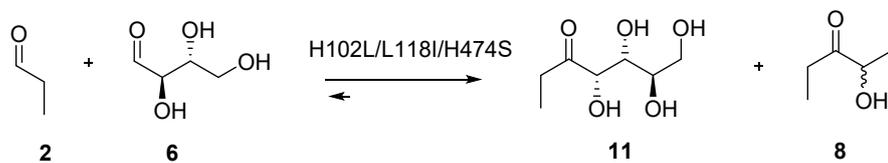
compounds	<b>10</b>	<b>DHA</b>	<b>L-erythrulose</b>
<i>in situ</i> conc. (mM)	23	8	4
<i>in situ</i> yield (%)	48	17	8
Isolated yield (%)	31	n.d.	n.d.



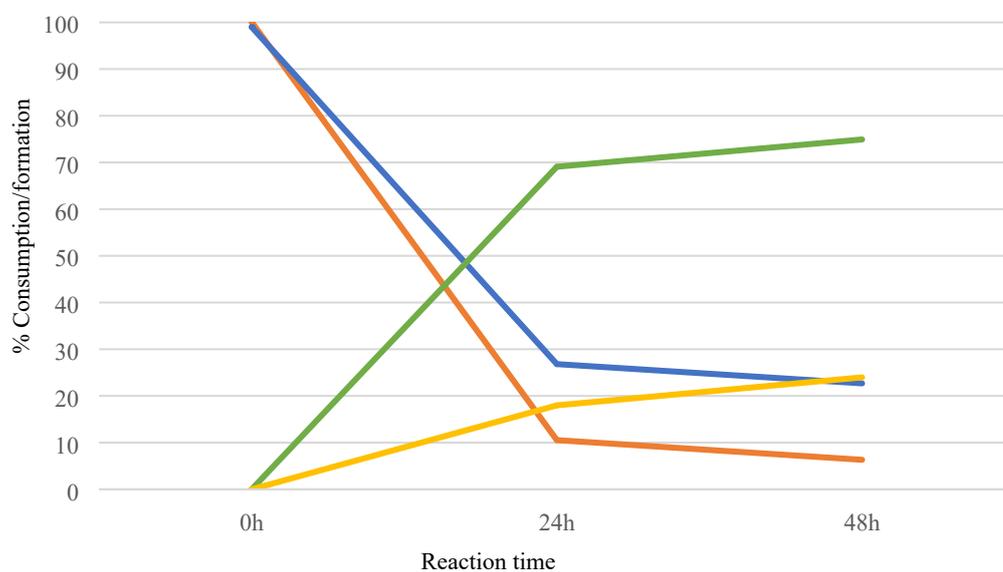
**Figure S9** : Cross-acyloin condensation progress of product **10** (■) from propanal **2** (■) and D-glyceraldehyde **5** (■) giving by-products dihydroxyketone **DHA** (■) and **L-erythrulose** (■)

### Synthesis of 1,2-dideoxy-D-arabino-hept-3-ulose **11**

Reaction conditions: propanal **2** (50 mM), D-erythrose **6** (50 mM), H102L/L118I/H474S 2.5 mg · mL<sup>-1</sup>, phosphate buffer, pH 7.0, 48 h.



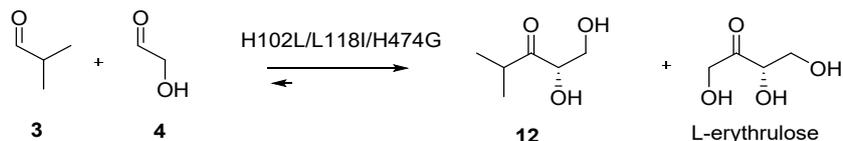
compounds	<b>11</b>	<b>8</b>
<i>in situ</i> conc. (mM)	37	12
<i>in situ</i> yield (%)	75	24
Isolated yield (%)	50	n.d. (volatile)



**Figure S10** : Cross-acyloin condensation progress of product **11** (■) from propanal **2** (■) and D-erythrose **6** (■) giving by-products propion **8** (■)

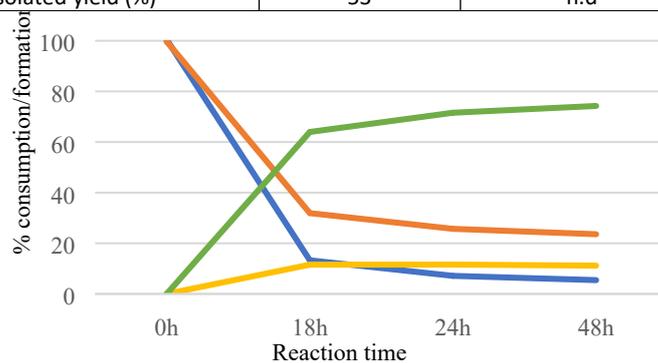
### Synthesis of 1,2-dihydroxy-4-methylpentan-3-one **12**

Reaction conditions: *iso*-butanal **3** (50 mM), GoA **4** (50 mM), H102L/L118I/H474G 2 mg · mL<sup>-1</sup>, phosphate buffer, pH 7.0, 48 h or 72 h.



a) **Protocol A (48h)** : *iso*-butanal **3** (50 mM), GoA **4** (50 mM) added simultaneously

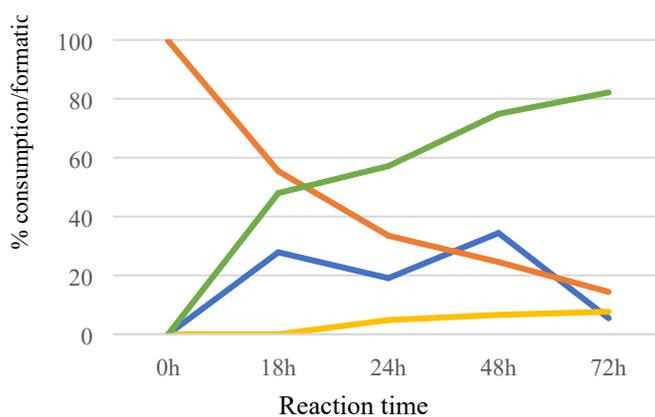
compounds	<b>12</b>	<b>L-erythrulose</b>
<i>in situ</i> conc. (mM)	37	6
<i>in situ</i> yield (%)	74	12
Isolated yield (%)	53	n.d.



**Figure S11** : Cross-acyloin condensation progress of product **12** (■) from *iso*-butanal **3** (■) and GoA **4** (■) giving by-products L-erythrulose (■)

b) **Protocol B (72 h)**: *iso*-butanal **3** (50 mM) at start, GoA **4** added by syringe perfusion (2.5 mM · h<sup>-1</sup>, 20 h)

compounds	<b>12</b>	<b>L-erythrulose</b>
<i>in situ</i> conc. (mM)	41	4
<i>in situ</i> yield (%)	82	8
Isolated yield (%)	60	n.d.

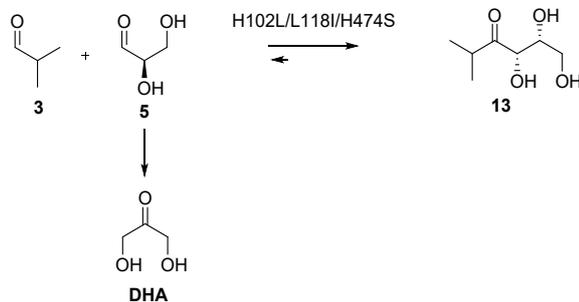


**Figure S12** : Cross-acyloin condensation progress of product **12** (■) from *iso*-butanal **3** (■) and GoA **4** (■) giving by-products L-erythrulose (■)

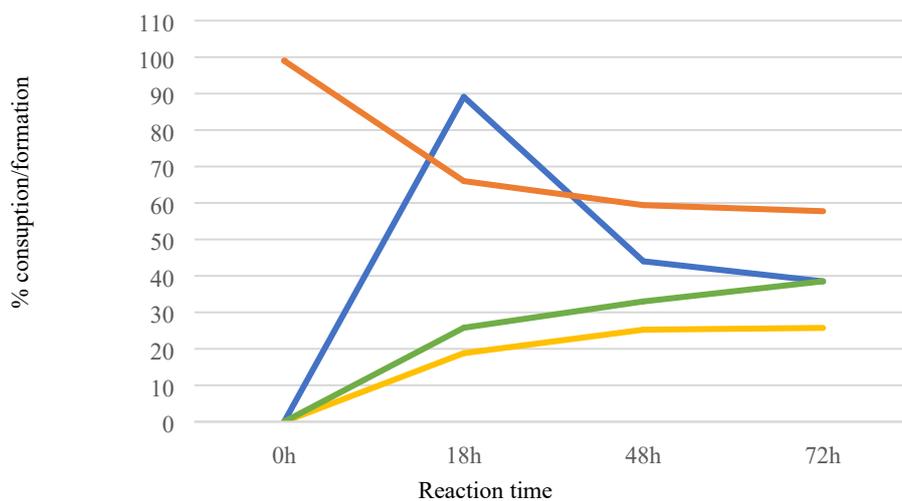
### Synthesis of 4,5,6-trihydroxy-2-methylhexan-3-one 13

Reaction conditions: *iso*-butanal **3** (50 mM), D-glyceraldehyde **5** (50 mM), H102L/L118I/H474S 3 mg . mL<sup>-1</sup>, phosphate buffer, pH 7.0, 72 h.

**Protocol B (72 h)** : *iso*-butanal **3** (50 mM) added at start, D-glyceraldehyde **5** added by syringe perfusion 2.5 mM . h<sup>-1</sup>, 20 h.



compounds	<b>13</b>	<b>DHA</b>
<i>in situ</i> conc. (mM)	19	13
<i>in situ</i> yield (%)	39	26
Isolated yield (%)	31	n.d.



**Figure S13** : Cross-acyloin condensation progress of product **13** (■) from *iso*-butanal **3** (■) and D-glyceraldehyde(■) giving by-products dihydroxyacetone (■)

## ESI 4: Product characterization ; NMR spectra of compounds 7-13

## Enantiomeric excess (e.e.) determination for products 7 and 8 obtained by self-acyloin condensation of aldehydes 1 or 2

**Table S4:** enantiomeric excess of  $\alpha$ -hydroxyketones 7 and 8 obtained at 37 °C using TK<sub>gst</sub> variants in the presence of ethanal 1 or propanal 2 as both nucleophiles and electrophiles (100 mM) in phosphate buffer 50 mM, pH 7.0. Enantiomeric excess (ee) were determined by chiral GC-MS as described earlier.<sup>17d</sup>

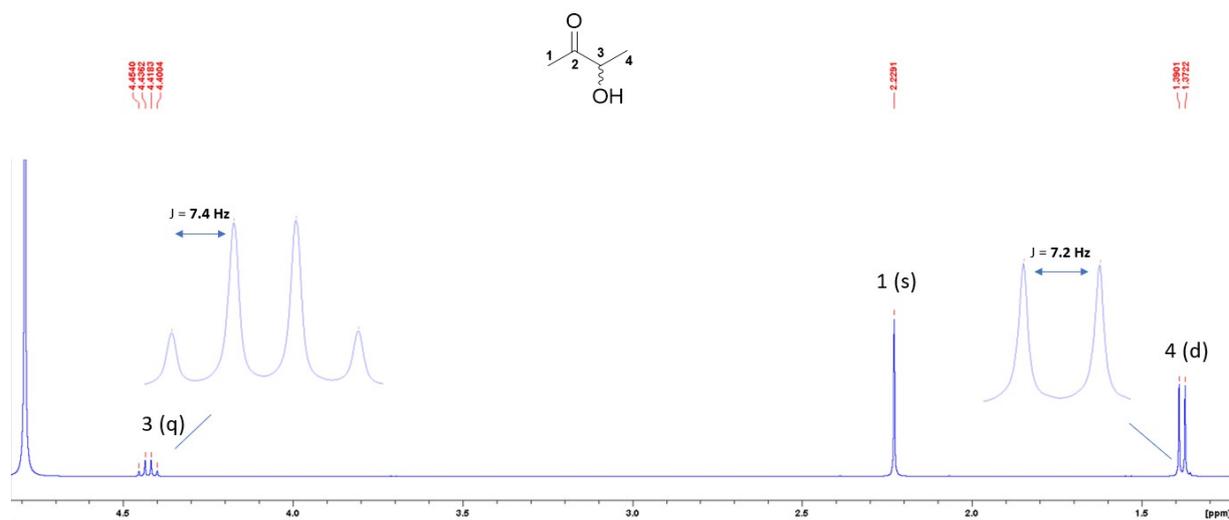
Nucleophile	Electrophile	Product	Time (h)	TK <sub>gst</sub> variant	TK <sub>gst</sub> mg . mL <sup>-1</sup>	ee (%)
 <b>1</b>	 <b>1</b>	 <b>7</b>	96	H102L/L118I/H474G	3	16 (S)
 pyruvate <b>1'</b>	 <b>1</b>	 <b>7</b>	48h	H102L/L118I/H474G	2	4
 <b>2</b>	 <b>2</b>	 <b>8</b>	48	H102L/L118I/ H474G	1.5	16 (S)

Enantiomeric excess (e.e.) determination for products 9 and 12 obtained from cross acyloin condensation of aldehydes 2 with 4 and 3 with 4 or using  $\alpha$ -ketoacids 2' with aldehyde 4 or  $\alpha$ -ketoacids 3' with aldehyde 4 .**Table S5:** enantiomeric excess of  $\alpha$ -hydroxyketones 9 and 12 obtained at 37 °C using TK<sub>gst</sub> variants in the presence of nucleophiles and electrophiles at 100 mM in phosphate buffer 50 mM pH 7.0. Enantiomeric excess (ee) were determined by chiral GC as described earlier.<sup>17d</sup>

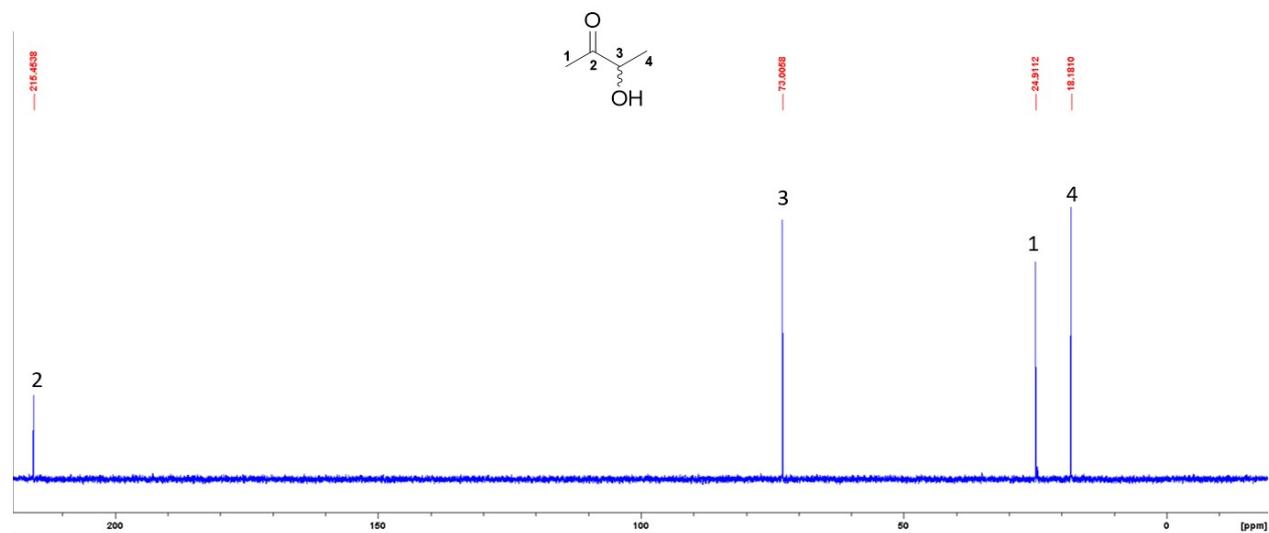
Nucleophile	Electrophile	Product	Time (h)	TK <sub>gst</sub> variant	TK <sub>gst</sub> mg . mL <sup>-1</sup>	ee (%)
 propanal <b>2</b>	 glycolaldehyde <b>4</b>	 <b>9</b>	48	H102L/L118I/H474G	3	65 (S)
 oxobutanoate <b>2'</b>	 glycolaldehyde <b>4</b>	 <b>9</b>	48	H102L/H474S	1	69 (S)
 isobutyraldehyde <b>3</b>	 glycolaldehyde <b>4</b>	 <b>12</b>	48h	H102L/H474G/L118I	2	34 (S)
 methyl oxobutanoate <b>3'</b>	 glycolaldehyde <b>4</b>	 <b>12</b>	48h	H102L/H474G/L118I	1	26 (S)

## ESI 5 : NMR spectra of compounds 7-13

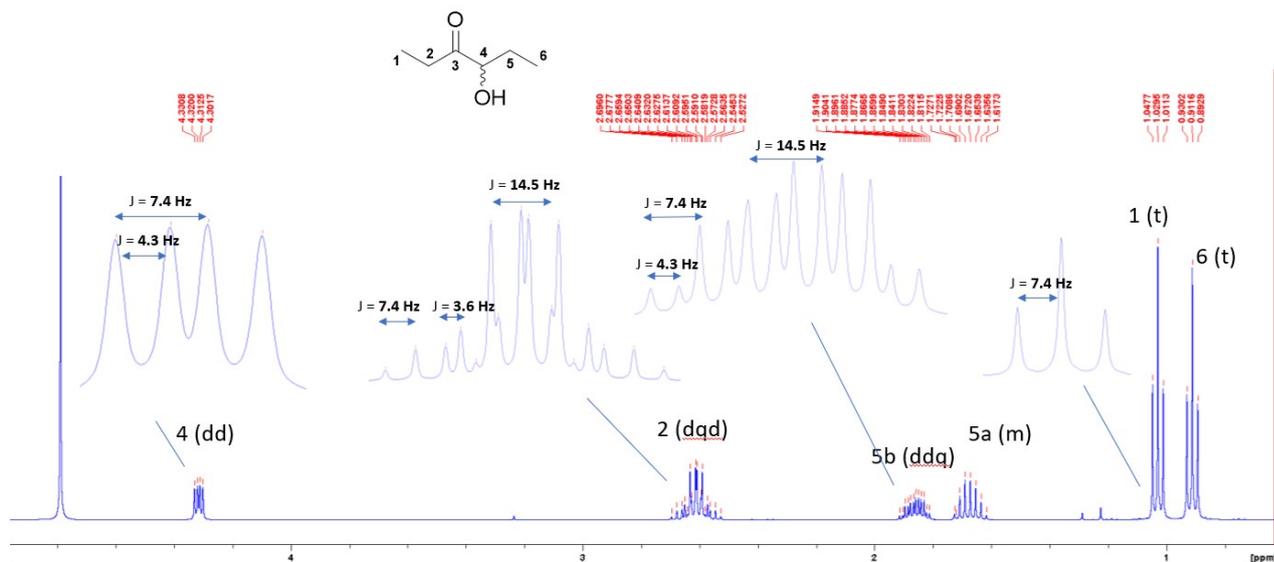
<sup>1</sup>H NMR spectrum of 3-hydroxybutan-2-one



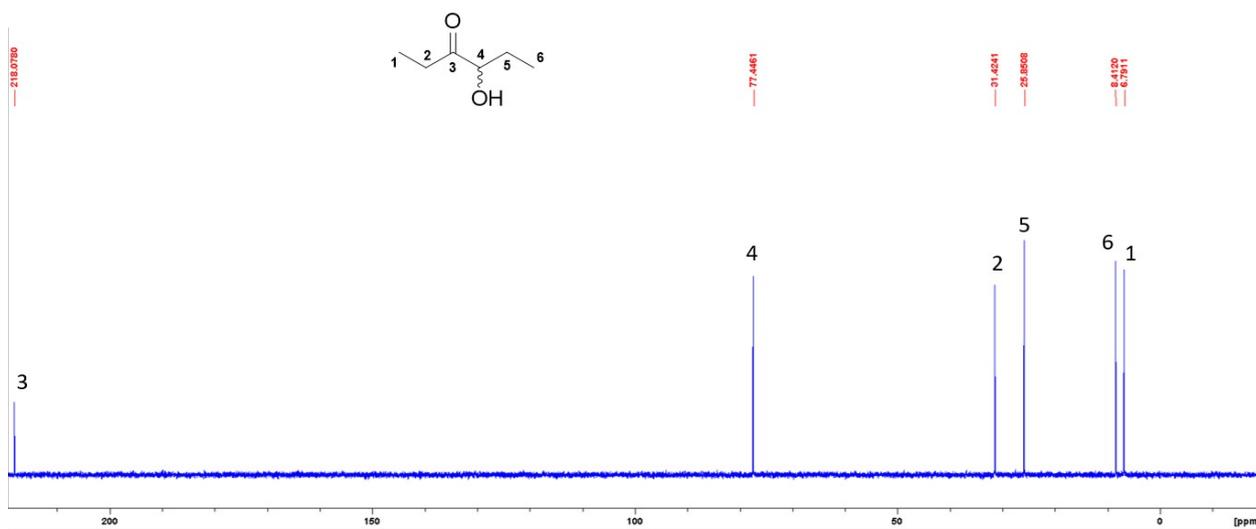
<sup>13</sup>C NMR spectrum of 3-hydroxybutan-2-one **7**



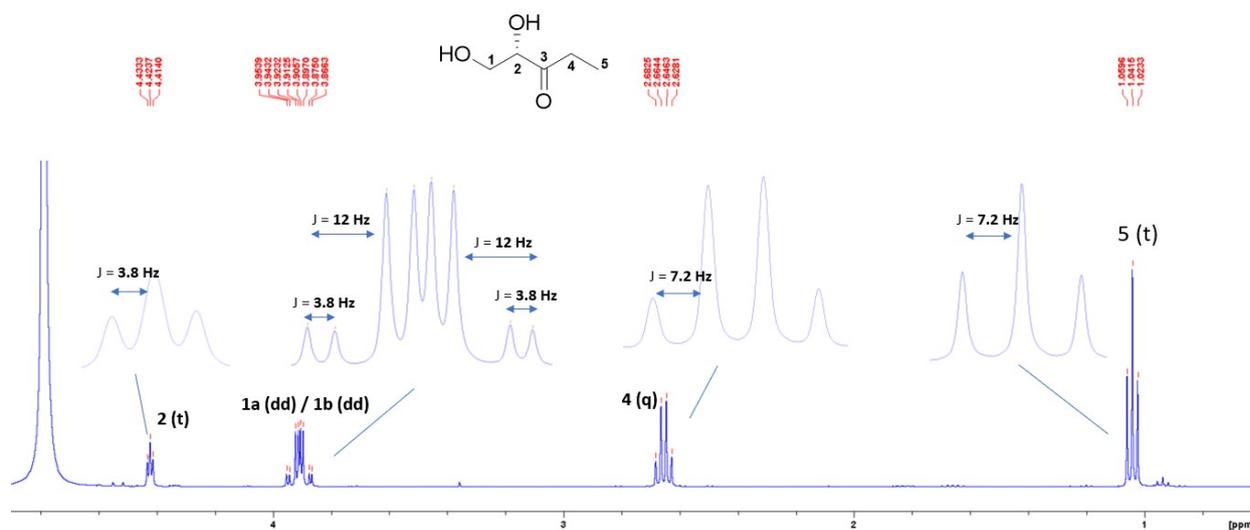
<sup>1</sup>H NMR spectrum of 4-hydroxyhexan-3-one **8**



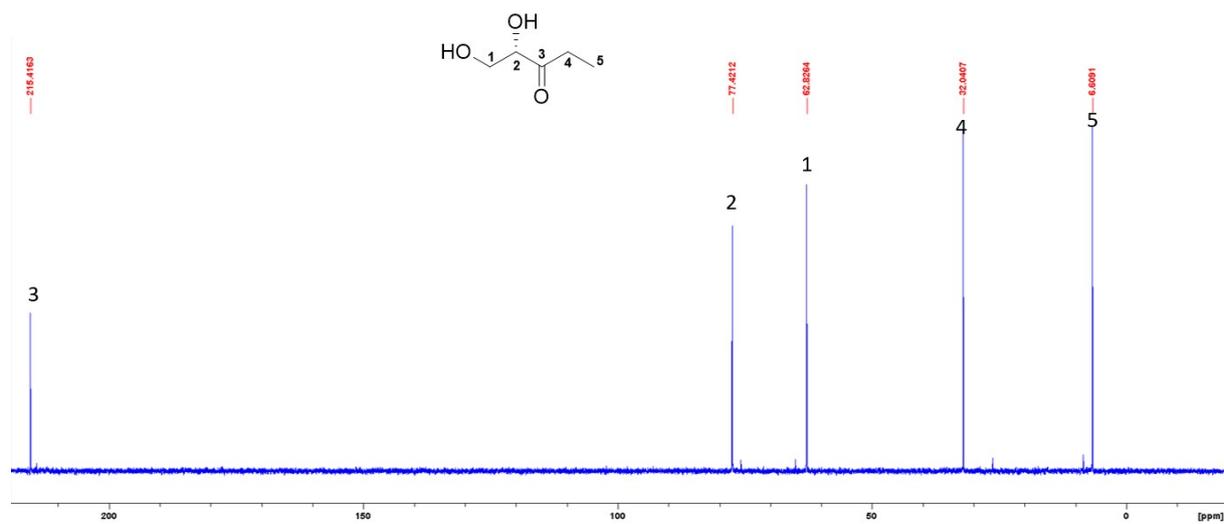
<sup>13</sup>C NMR spectrum of 4-hydroxyhexan-3-one **8**



<sup>1</sup>H NMR spectrum of 1,2-dihydroxypentan-3-one **9**

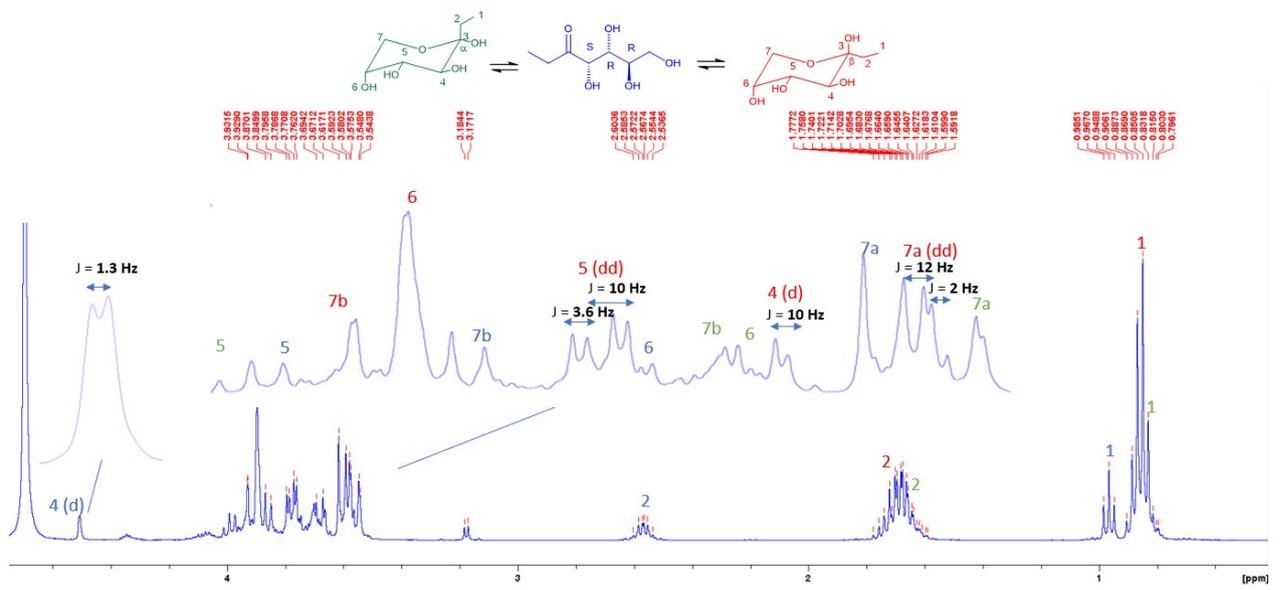


<sup>13</sup>C NMR spectrum of 1,2-dihydroxypentan-3-one **9**

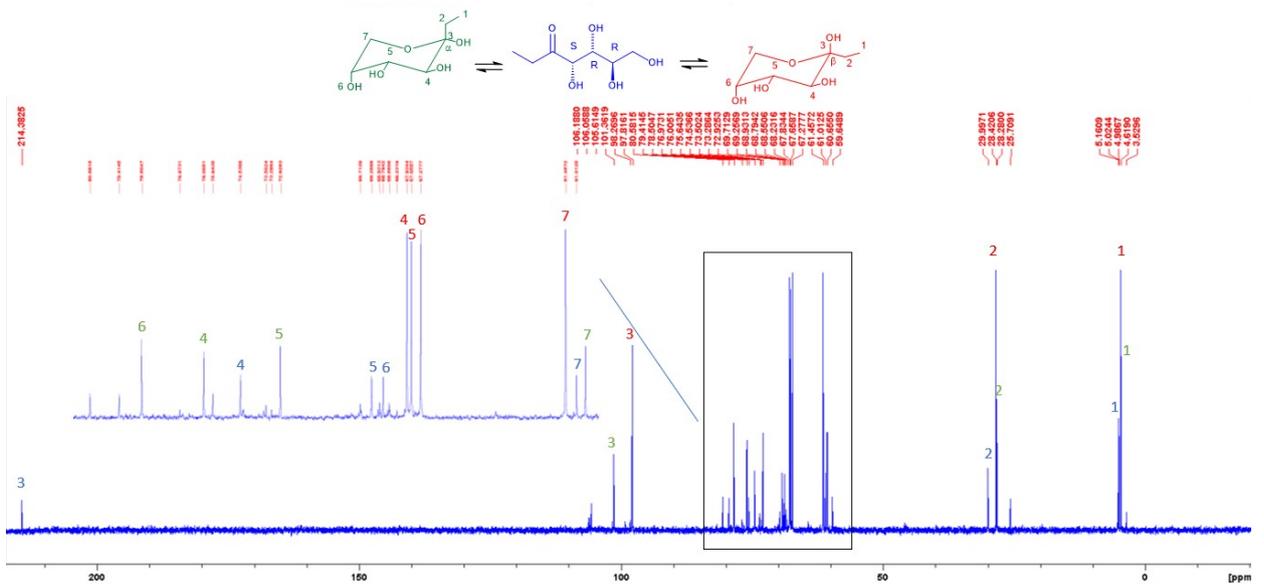




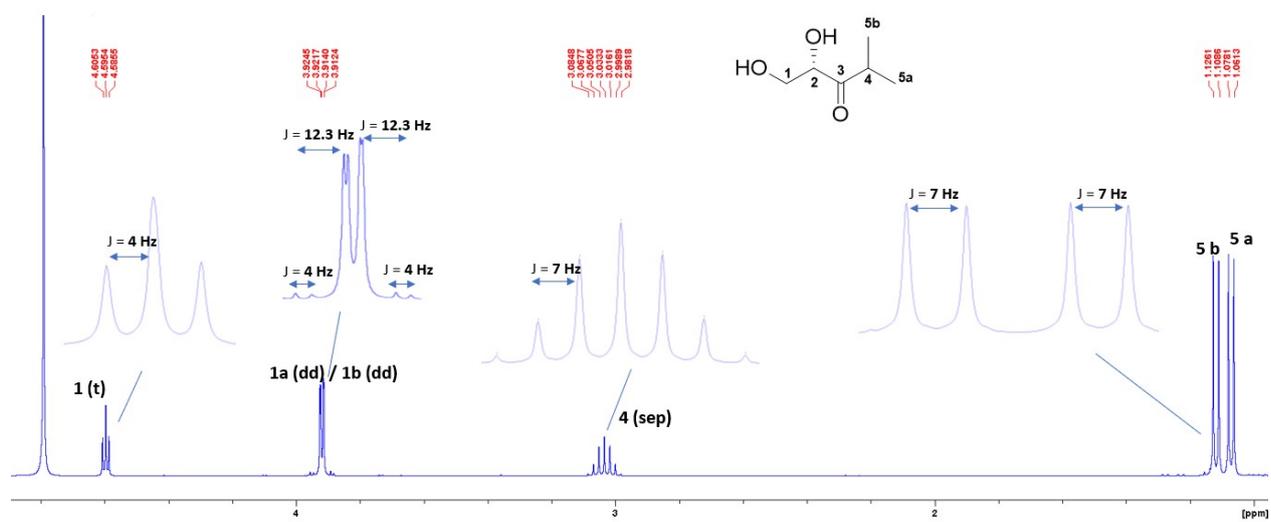
<sup>1</sup>H NMR spectrum of 1,2-dideoxy-D-arabino-hept-3-ulose **11**



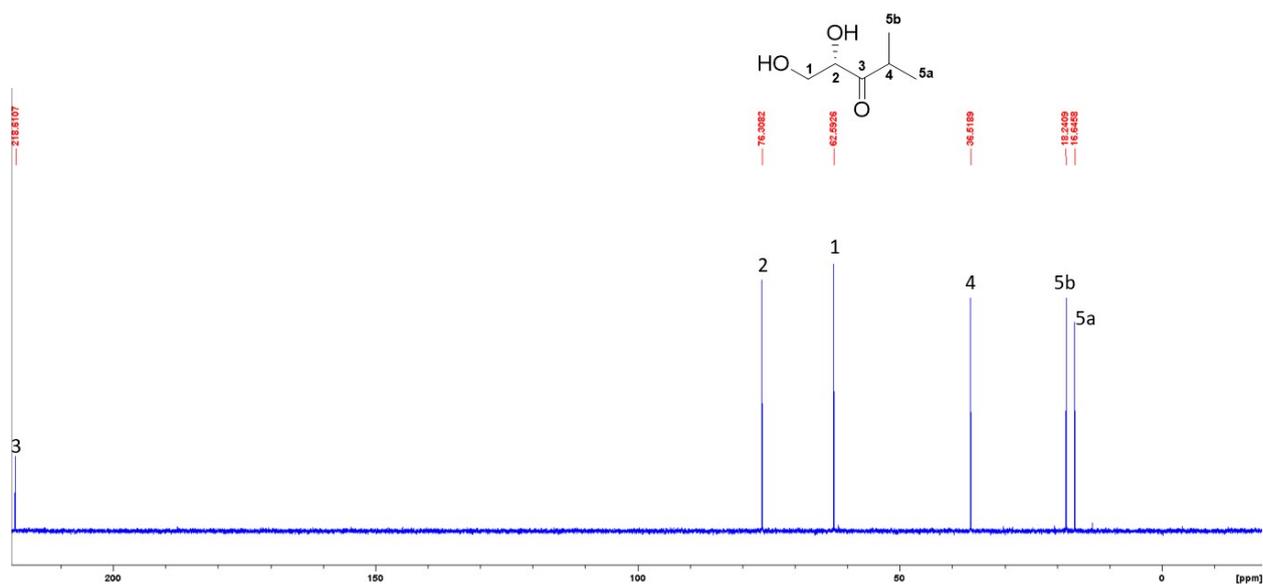
<sup>13</sup>C NMR spectrum of 1,2-dideoxy-D-arabino-hept-3-ulose **11**



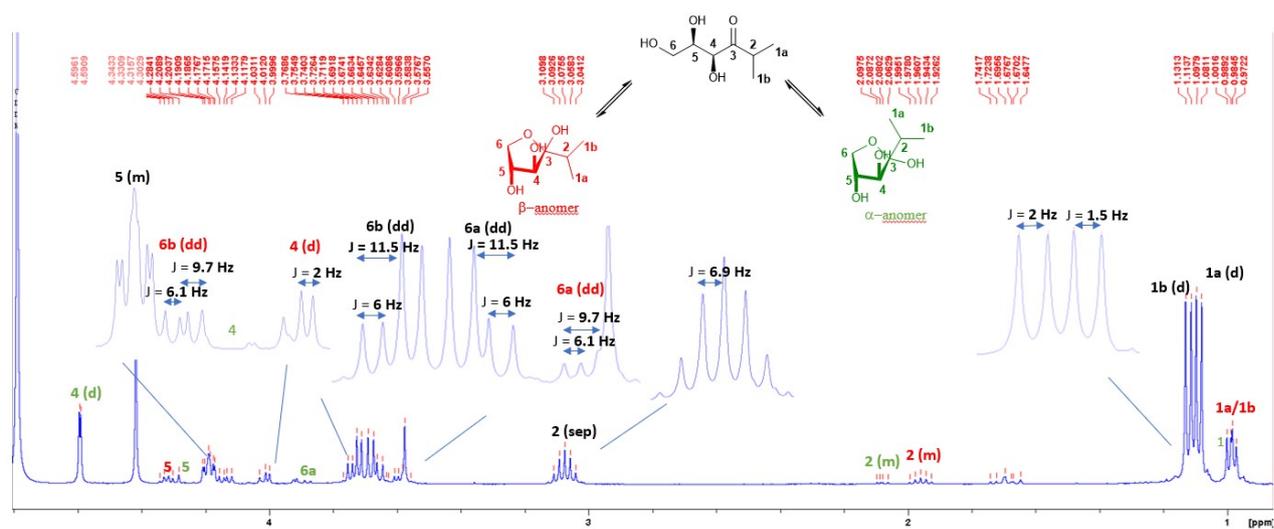
$^1\text{H}$  NMR spectrum of 1,2-dihydroxy-4-methylpentan-3-one **12**



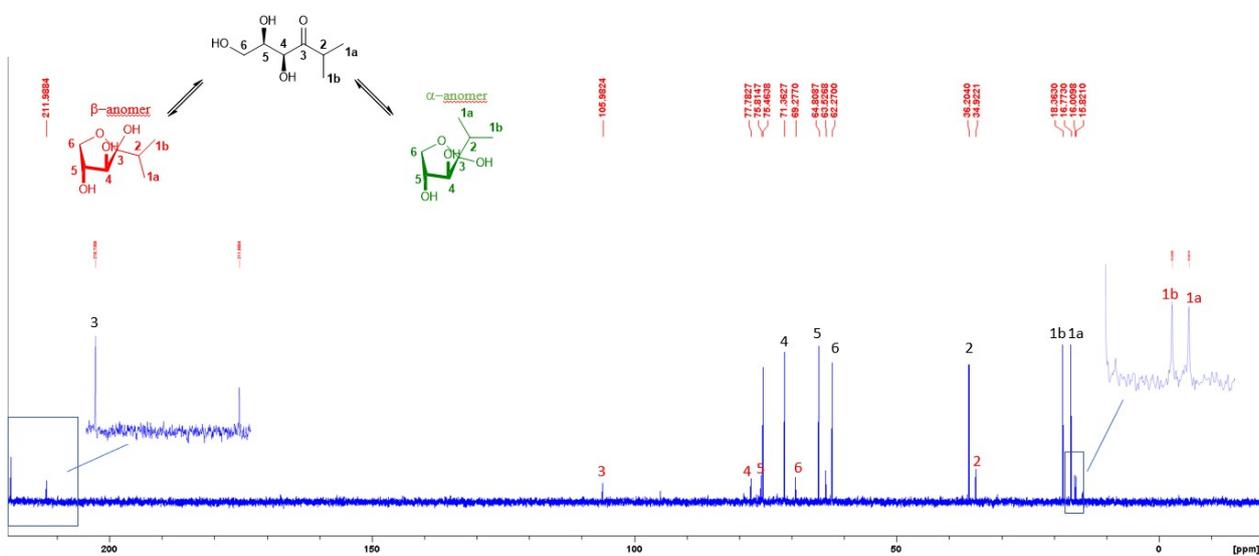
$^{13}\text{C}$  NMR spectrum of 1,2-dihydroxy-4-methylpentan-3-one **12**



<sup>1</sup>H NMR spectrum of 4,5,6-trihydroxy-2-methylhexan-3-one **13**



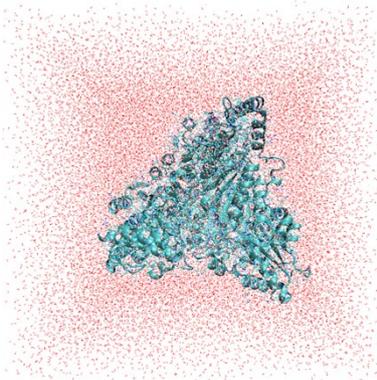
<sup>13</sup>C NMR spectrum of 4,5,6-trihydroxy-2-methylhexan-3-one **13**



## ESI 6 Molecular dynamics studies

### *In silico* study of H102L/L118I/H474G variant with two different substrates as nucleophilic donor, propanal or iso-butanal.

**Method :** The 3D structures of microbial TKs such as TK<sub>sce</sub>, TK<sub>eco</sub>, and TK from *Bacillus anthracis* (TK<sub>ban</sub>) have a strong protein sequence homology.<sup>17a</sup> All TKs are dimeric proteins with an identical active site in each monomer and the key residues stabilizing the TK substrates are identical and have similar orientations to ThDP. The TK<sub>gst</sub> 3D structure being unknown, a model of the TK<sub>gst</sub> active-site pocket containing its phosphorylated acceptor aldose, D-erythrose-4-phosphate (E4P), was constructed using the X-ray crystal structure of TK<sub>ban</sub> as a template, which belongs to the same microbial species and has a high percentage of identity (74%).<sup>17</sup> The residues H102/L118I/H474 of wild type TK<sub>gst</sub> were replaced by L/I/G respectively using the Rotamers module of Chimera software. Each starting position of 102, 118 and 474 residues were defined according to the most favorable “dunbrack” proposition of Chimera software giving any steric clash with the rest of the structure. Molecular dynamics of this new model was carried out taking into account the changes induced by the mutations. This construction was therefore solvated in a box of water (120 Å x 120 Å x 120 Å) with a sufficient quantity of ions to electrically neutralize the system to be able to use the periodic conditions during the dynamics. The system was minimized and a temperature ramp applied to reach the set temperature at 310 K in NVT. Then, the volume was equilibrated to reach the set pressure of 1 atm in NPT. The structure of this mutant (Figure S14) was studied with propanal or with iso-butanal placed randomly in the two available active sites of TK<sub>gst</sub>. All input files were generated with VMD and molecular dynamics made with NAMD.



**Figure S14:** TK<sub>gst</sub> H102L/L118I/H474G variant in solvation box

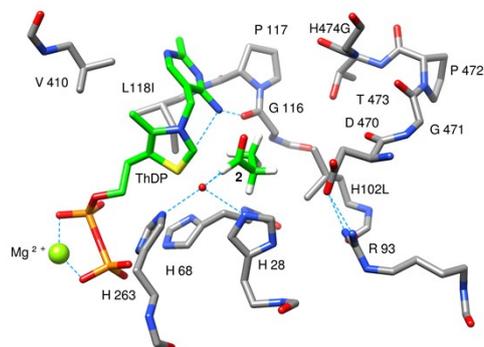
**Analysis :** Structurally, the modifications induced by the mutations led to slight modifications : we note that the thiazole ring of ThDP was slightly reoriented due to the L118I mutation, which can have a slight impact on the stability of activated ThDP because the C2(thiazolium) – H4'' (aminogroup of pyrimidine cycle) distance is greater compared to that observed in wild type TK<sub>gst</sub>. The other interactions of ThDP with the variant are all conserved. As we already published on the mechanism of TK<sub>sce</sub> and TK<sub>eco</sub>,<sup>29</sup> we have chosen to start from a system in which ThDP was already activated and the carbanion C2 was closed to the ThDP H4'' proton. As observed in TK<sub>coli</sub>,<sup>29b</sup> H263 and H28 need to be protonated to initiate the C-C bond cleavage by forming a pinch able to fix an hydroxyl group. Due to H263 direct interaction with the pyrophosphate group, only this residue could be able to catch a proton. ThDP could therefore be protonated on Nδ1 and H28 on Nε2. We also demonstrated that the distance between the C2 carbanion and the electrophilic carbon of the donor substrate was in a range of 2.8 or 3.2 Å, and that the oxygen of the aldehyde was oriented towards H4'' of the ThDP. Using all of these informations, we carried out two molecular dynamics of propanal and iso-butanal while maintaining the distance between C2 ThDP and the electrophilic carbon at 3.2 Å by adding an extrabond in our calculation.

As described before, the system consists of a box of water of 120 Å x 120 Å x 120 Å, neutralized in charge by the addition of sodium ions. The system was minimized with a temperature reaching 300 K. After a short equilibration (4ns), because the system was already balanced beforehand, we start a dynamic over 20 ns. This duration may seem short, but we are on a balanced system, and we only wish to look at the placement of propanal or iso-butanal in the TK<sub>gst</sub> active sites in the presence of constraints linked to the expected reactivity, in these conditions it is more than sufficient time. The aim of these dynamics was only to look at how the hydrophobic parts of these aldehydes were positioned in the TK<sub>gst</sub> variants.

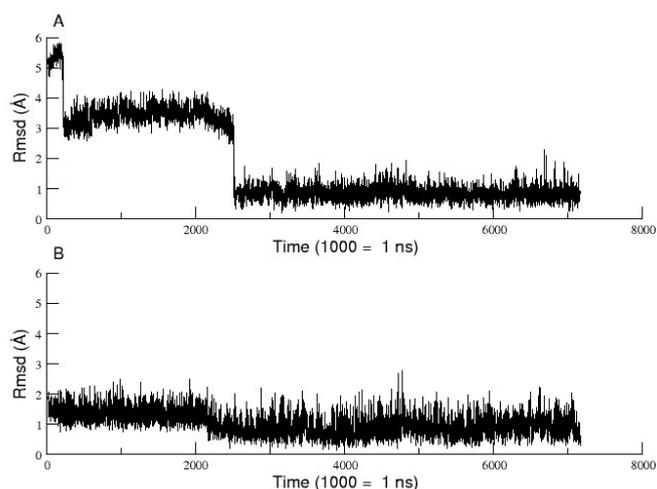
#### **Molecular dynamics of propanal 2 in H102L/L118I/H474G TK<sub>gst</sub> variant**

In the case of propanal, the dynamics gave identical positions in the two active sites (Figure S15). The methyl group was fixed under the ThDP by making hydrophobic interactions with the proton of the Cβ of H68, the protons of the Ca of G67 and

especially with the side chains of two of the three mutated residues H102L and L118I. This position propanal was compatible with a nucleophilic attack of C2 ThDP on the carbonyl of the propanal. We also notice that the propanal did not disturb the position of the water molecule trapped between H28 and H263, found at this position in all TK structures when the resolution is sufficient. This water molecule could play the role of H bond acceptor for the hydrogen atom coming from the propanal. Rmsd of propanal in active site (Figure S15) indicates one favorite position for this compound after a very short time (2.5 ns) similar whatever the active sites and the starting point of the dynamics : far (A) or close (B) of the ThDP.



**Figure S15** : position and interactions of propanal obtained by molecular dynamics in H102L/L118I/H474G TK<sub>gst</sub> variant.

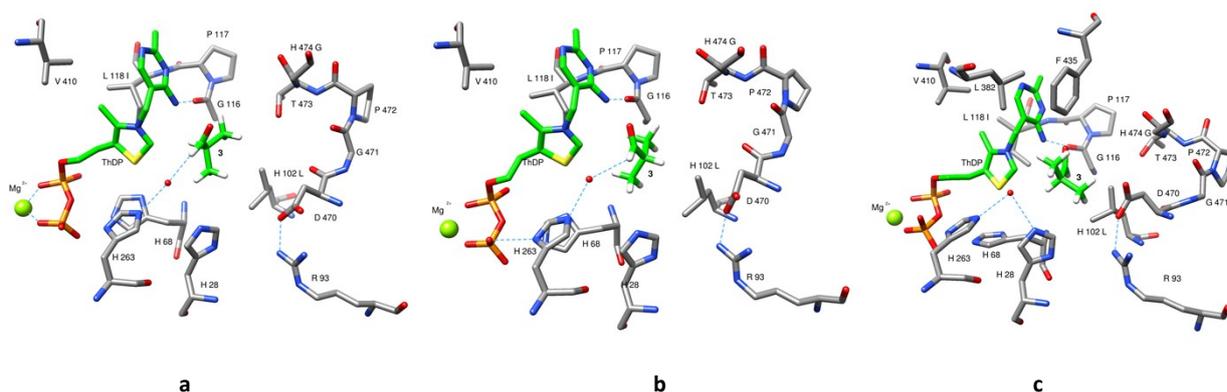


**Figure S16** : Rmsd of propanal in active site of H102L/L118I/H474G TK<sub>gst</sub> variant : far (A) or close (B) of the ThDP.

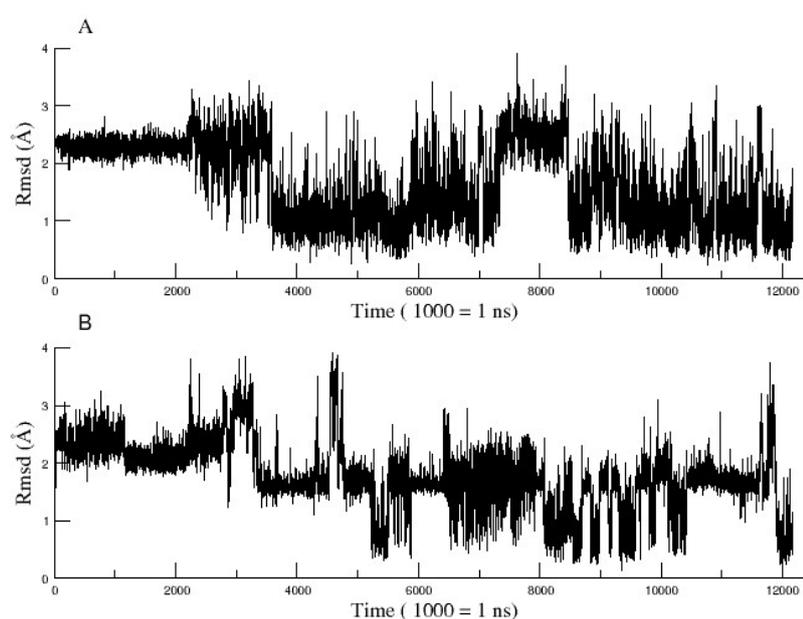
#### Molecular dynamics of *iso*-butanal 3 in H102L/L118I/H474G TK<sub>gst</sub> variant

The molecular dynamic study and Rmsd conducted with *iso*-butanal in both active sites showed several positions for *iso*-butanal with three preferred positions : two with the dimethyl group oriented towards H102L (**a** and **b**), and the third (**c**) with the dimethyl group pointing at the opposite direction towards the hydrophobic part of active site delimited by L191, L382 and F435. (Figure 3). The positions **a** and **b** were in accordance with the proposed reaction pathway (figure 4) showing that H263 was able to catch a proton. The third position of *iso*-butanal **c** did not allow a reaction pathway able to remove the proton, this one pointing in the direction of H102L.

*iso*-butanal, with two carbons more compared to propanal, led to several possible positions (with two preferred orientations) while propanal gave one preferred position (Figures S17 and S18).



**Figure S17** : three main positions (a, b, c) of *iso*-butanal obtain in the two active sites of H102L/L118I/H474G TK<sub>gst</sub>.variant by molecular dynamics.



**Figure S18** : Rmsd with *iso*-butanal in active sites of H102L/L118I/H474G TK<sub>gst</sub>.variant : A) and B) corresponds to two different initial orientations of *iso*-butanal in active sites of TK<sub>gst</sub>.

### QMMM study to determine the reaction pathway

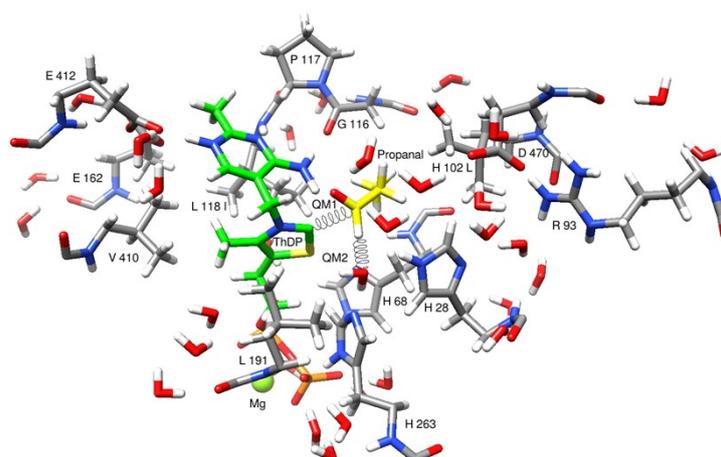
The objective was to determine by QMMM study if the water molecule trapped by H28 and H263 can serve as a relay in the removal of the proton from propanal during the reaction with ThDP in H102L/L118I/H474G TK<sub>gst</sub>.variant.

**Method.** From the previous molecular dynamics, we use the QwikMD module from VMD to prepare a QMMM molecular dynamics. The QM part will be treated by MOPAC with the semi-empirical method PM7, while the MM part will be treated with NAMD and the Charmm force field. To define the QM part we used ThDP ; Mg<sup>2+</sup> ; H28 ; H68 ; R93 ; H263 ; H102L ; G116 ; P117 ; L118I ; E162 ; L191 ; V410 ; E418 ; D470, propanal and all water molecules within a distance of 2.5 Å of the residues selection (total number of 361 atoms). QMMM protocol was a minimization with restraints on the backbone followed by a ramp in temperature to reach 300 K in NPT ensemble, a volume equilibration in NPT with restraints on backbone and then a production QMMM dynamics without any restraints (PM7 XYZ T=2M 1SCF MOZYME CUTOFF=9.0 AUX LET GRAD). In a second time two consecutive QMMM production molecular dynamics were made with qmcSMD constraints

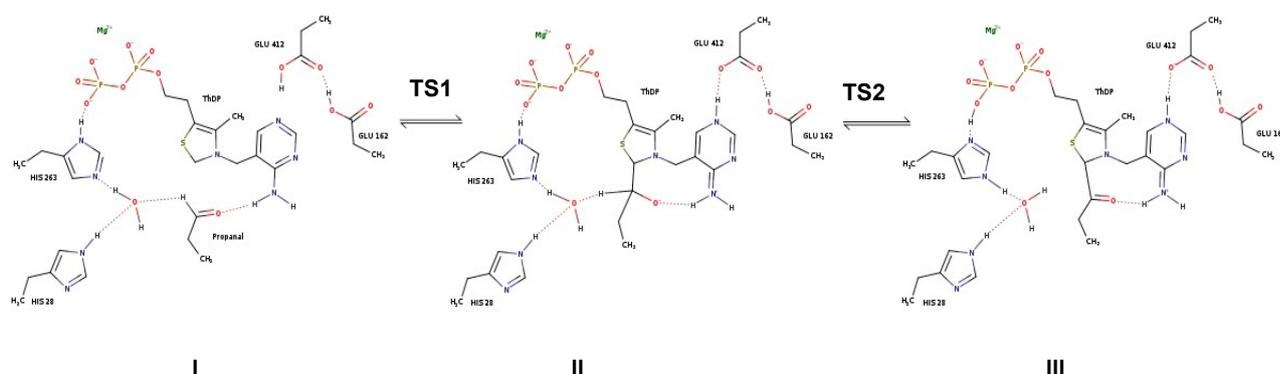
i) between C2 ThDP and the carbon atom of CHO of propanal ( QM1 : 100 0.01 1.7 ) for the first one and ii) qmcSMD constraints between the hydrogen atom of CHO of propanal and a water molecule (QM2 : 100 0.01 1.5 ) for the second.

**Analysis** : the reaction pathway proposed by dynamic QMMM, showed that the reaction is electronically possible and led to three stable entities I, II, III (Figure S20). The analysis of analyzes of distances (Figure S21) and the evolution of the energy during QMMM molecular dynamics (Figure S22) showing two maximum energies TS1 and TS2 were in accordance with our proposed two steps reaction pathway.

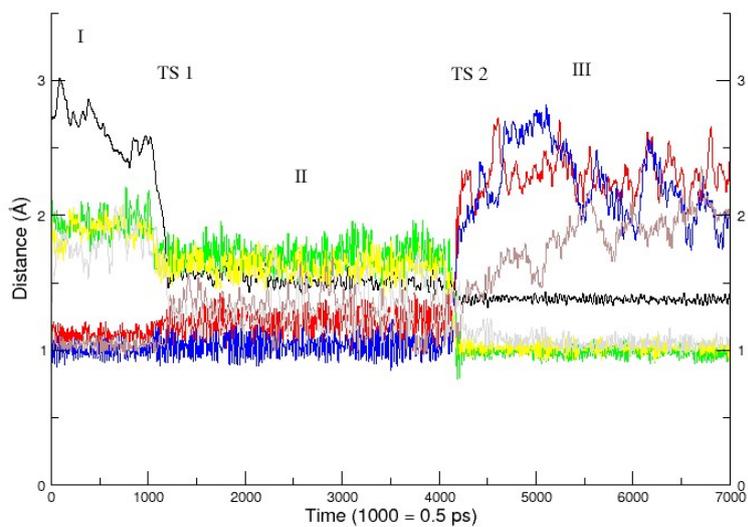
At the semi empirical precision level PM7, we will not discuss on the energy values, the imprecision is too great). we cannot go further in the precision of our model by DFT type methods with more precise electronic bases, type 6-31g(d,p), our experience has taught us that for such calculation we must start from high-resolution crystallographic structures to be able to hope to obtain a stable and fully justified state, the starting point of a molecular dynamics in classical mechanics for this type of calculation is not possible, except to drastically reduce the system which will have the consequence of reduce the effects of the environment on the reaction and therefore to reduce the overall precision.



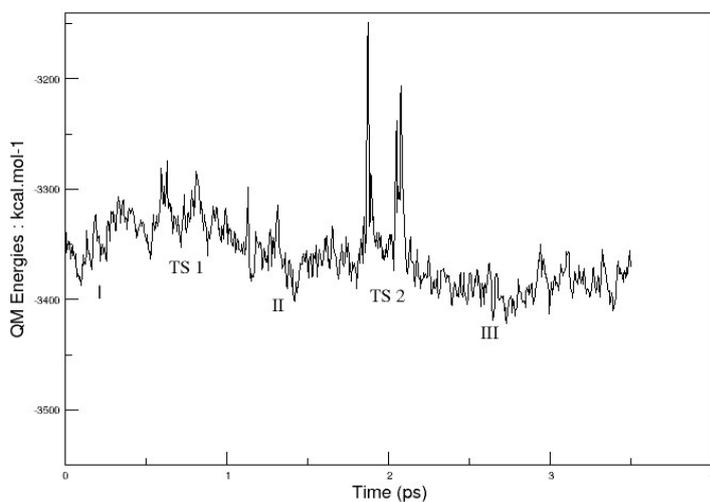
**Figure S19** : QM part treated by the semi-empirical method PM7, in QMMM molecular dynamics with propanal in H102L/L118I/H474G TK<sub>gst.</sub> variant. The springs represent the constraints applied to the system to accelerate the reactivity (qmcSMD method).



**Figure S20**: Reaction pathway with propanal in H102L/L118I/H474G TK<sub>gst.</sub> variant proposed and validated by the semi-empirical method PM7



**Figure S21** : analyzes of distances during QMMM molecular dynamics obtained by the semi-empirical method PM7 with propanal in H102L/L118I/H474G TK<sub>gst</sub>.variant. Black ThDP:C2 to 3:C1 ; Red 3:C1 to 3:H1 ; Green 3:H1 to HOH:O ; Blue HOH:O to HOH:H1 ; Yellow H263:Nε2 to HOH:H1 ; Brown H263:Nδ1 to H263:Hδ1 ; Grey H263:Hδ1 to ThDP:O3β).



**Figure S22** : evolution of the energy of the QM part of the QMMM molecular dynamics with propanal in H102L/L118I/H474G TK<sub>gst</sub>.variant. The QM energies showed two maximum energies TS1 and TS2 required for our proposed two steps reaction pathway (the two higher pics correspond to artifacts due to proton transfers independent of the reaction pathway).

ES 7 : Determination of mass metrics<sup>31</sup>

Table S6 : comparison of some metrics for pathway A and B

Hydroxyketone products	Pathway A : decarboxylation of $\alpha$ -ketoacids by TK <sub>gst</sub> <sup>17h</sup>				Pathway B : cross acyloin catalysed by TK <sub>gst</sub> (present work)			
	E factor	Atom economy	Carbon efficiency	Relative mass intensity	E factor	Atom economy	Carbon efficiency	Process mass intensity
<b>9</b>					0.63	100 %	61 %	1.63
<b>10</b>	1.15	69 %	58 %	46 %	1.96	100 %	34 %	2.96
<b>11</b>	0.82	61 %	61 %	55 %	0.98	100 %	50 %	1.98
<b>12</b>					0.67	100 %	60 %	1.67
<b>13</b>					1.58	100 %	31 %	2.58

Table S7 : comparison of *E* factors for TK<sub>gst</sub>- and NHCs- catalyzed reactions obtained with the same nucleophiles **2** or **3** but different electrophiles (electrophiles **4**, **5** and **6** not used with NHCs).

E factor : TK <sub>gst</sub> -catalyzed cross acyloin reaction (present work)		
Electrophiles	Nucleophile : propanal <b>2</b>	Nucleophile : <i>iso</i> -butanal <b>3</b>
CH <sub>2</sub> OH-CHO <b>4</b>	0.63	0.67
CH <sub>2</sub> OH-CHOH-CHO <b>5</b>	1.96	1.58
CH <sub>2</sub> OH-CHOH-CHOH-CHO <b>6</b>	0.98	
E factor : NHCs/Cs <sub>2</sub> CO <sub>3</sub> or Rb <sub>2</sub> CO <sub>3</sub> -catalyzed cross acyloin reaction <sup>32,33</sup>		
Electrophiles	Nucleophile : propanal <b>2</b>	Nucleophile : <i>iso</i> -butanal <b>3</b>
C <sub>6</sub> H <sub>5</sub> -CHO	7.8	
CF <sub>3</sub> C <sub>6</sub> H <sub>4</sub> -CHO	8	
ClC <sub>6</sub> H <sub>4</sub> -CHO		1.1
3-Me-C <sub>6</sub> H <sub>4</sub> -CHO		8.8