



Cite this: *Org. Biomol. Chem.*, 2021, **19**, 6493

## Characterisation of a hyperthermophilic transketolase from *Thermotoga maritima* DSM3109 as a biocatalyst for 7-keto-octuronic acid synthesis†

Max Cárdenas-Fernández,<sup>a</sup> Fabiana Subrizi,<sup>c</sup> Dragana Dobrijevic,<sup>a</sup> Helen C. Hailes<sup>c</sup> and John M. Ward<sup>a</sup>

Transketolase (TK) is a fundamentally important enzyme in industrial biocatalysis which carries out a stereospecific carbon–carbon bond formation, and is widely used in the synthesis of prochiral ketones. This study describes the biochemical and molecular characterisation of a novel and unusual hyperthermophilic TK from *Thermotoga maritima* DSM3109 (TK<sub>tmr</sub>). TK<sub>tmr</sub> has a low protein sequence homology compared to the already described TKs, with key amino acid residues in the active site highly conserved. TK<sub>tmr</sub> has a very high optimum temperature (>90 °C) and shows pronounced stability at high temperature (e.g.  $t_{1/2}$  99 and 9.3 h at 50 and 80 °C, respectively) and in presence of organic solvents commonly used in industry (DMSO, acetonitrile and methanol). Substrate screening showed activity towards several monosaccharides and aliphatic aldehydes. In addition, for the first time, TK specificity towards uronic acids was achieved with TK<sub>tmr</sub> catalysing the efficient conversion of D-galacturonic acid and lithium hydroxypyruvate into 7-keto-octuronic acid, a very rare C<sub>8</sub> uronic acid, in high yields (98%, 49 mM).

Received 25th June 2021,  
Accepted 5th July 2021

DOI: 10.1039/d1ob01237a

rsc.li/obc

## Introduction

Biocatalysis is considered as a green and sustainable technology based on the principles and metrics of green chemistry and sustainable development.<sup>1</sup> Application of enzymatic reactions in industrial chemical processes is very advantageous due to their high stereo- and regioselectivity, low energy demand and because enzymatic catalysts are renewable.<sup>2</sup> In order to enhance the contribution of biocatalysis to a more sustainable society, increasing the availability of different of enzyme classes through enzyme discovery and engineering is paramount.

Transketolase (TK, EC 2.2.1.1) is a thiamine diphosphate (ThDP) dication (usually Mg<sup>2+</sup>) dependent enzyme, that catalyses the *in vivo* reversible carbon–carbon bond formation between D-xylulose-5-phosphate and D-ribose-5-phosphate

yielding D-sedoheptulose-7-phosphate and glyceraldehyde-3-phosphate. TK follows a Ping-Pong Bi-Bi reaction kinetics where the donor and acceptor substrates are not able to bind to the protein simultaneously.<sup>3</sup> TK is a key enzyme linking the pentose phosphate and the glycolysis pathways.<sup>4,5</sup> From the synthetic point of view, wild-type (WT) TK adds a 2-carbon unit to an aldehyde creating a C–C bond between the aldehyde acceptor and a carbonyl or a keto acid donor substrate, and generating a new stereogenic centre. This forms a new asymmetric  $\alpha$ -hydroxyketone with an (S)-configuration, which is otherwise difficult to prepare chemically.<sup>6,7</sup> WT and engineered TKs from different microbial sources have been used in the synthesis of phosphorylated sugars,<sup>8</sup> non-phosphorylated aliphatic and aromatic acyloins,<sup>9–13</sup> and rare sugars such as L-glucoheptulose,<sup>14,15</sup> as well as in coupled cascade reactions.<sup>15–18</sup>

For synthetic applications, biocatalysis using enzymes such as TK allows a sustainable, one-step stereoselective way for preparing chiral building blocks and fine chemicals. In synthetic uses, the utilisation of lithium hydroxypyruvate (LiHPA) in TK-catalysed reactions is well established as a common ketol donor (Scheme 1). TKs can be deactivated during biocatalytic syntheses by different processes such as oxidation, substrate or product inhibition, or dissociation of the ThDP cofactor; all of this can be avoided by applying reaction engineering strat-

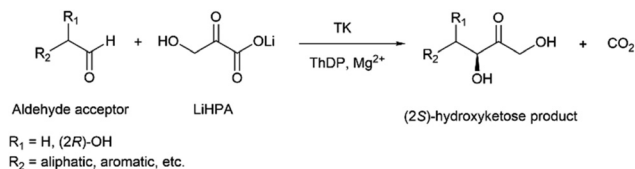
<sup>a</sup>Department of Biochemical Engineering, University College London, Gower Street, London WC1E 6BT, UK. E-mail: m.cardenas-fernandez@kent.ac.uk, j.ward@ucl.ac.uk

<sup>b</sup>School of Biosciences, University of Kent, Canterbury, Kent CT2 7NJ, UK

<sup>c</sup>Department of Chemistry, University College London, 20 Gordon Street, London WC1H 0AJ, UK

†Electronic supplementary information (ESI) available. See DOI: 10.1039/d1ob01237a





**Scheme 1** Transketolase reaction with aldehyde acceptors and lithium hydroxypyruvate (LiHPA) as carbon donor.

egies including the use of reducing agents, substrate feeding, or *in situ* product removal.<sup>19</sup> However, the use of thermostable or thermophilic enzymes also avoids some of these limitations.

Thermophilic enzymes are sourced mainly from microorganisms able to live in high temperature environments. These enzymes possess an optimum temperature above 50 °C and extended activity half-life ( $t_{1/2}$ ) at high temperatures, being more robust under demanding industrial applications such as high temperatures and the presence of organic solvents compared to their mesophilic counterparts.<sup>20,21</sup> *Thermotoga maritima* (DSM3109 or MSB8) is an anaerobic hyperthermophilic bacterium, which was first isolated from geothermally heated marine sediments near Vulcano Island in Italy; it has an optimum growth temperature of 80 °C and a G + C content of 46.2%.<sup>22</sup> *T. maritima* is able to grow on several monosaccharides and disaccharides as the sole carbon and energy source.<sup>23</sup> The genome sequence of *T. maritima* revealed that a substantial portion of its genes were inherited *via* horizontal or lateral gene transfer from Archaea, and from anaerobic members of the Firmicutes, the low GC Gram-positive bacteria.<sup>24</sup>

To date, despite their high demand in industrial biotechnology, only a few thermostable TKs have been described. Recently a “split-gene” TK from *Carboxydotherrmus hydrogenoformans* (TK<sub>chy</sub>) has been reported showing an optimum temperature of 70 °C and good stability in the presence of different organic solvents.<sup>25</sup> Also, Bawn *et al.*<sup>15</sup> isolated and characterised TKs from *Deinococcus geothermalis* (TK<sub>dge</sub>) and *Deinococcus radiodurans* (TK<sub>dra</sub>) which were used to produce L-glucose from L-arabinose; and both TKs have optimum temperatures of 50 °C. Notably, TK from *Geobacillus stearothermophilus* (TK<sub>gst</sub>) was the first thermostable TK reported also with an optimum temperature of 50 °C.<sup>26</sup> TK<sub>gst</sub> has been subjected to multiple protein engineering studies in order to expand its substrate scope.<sup>9,27–29</sup>

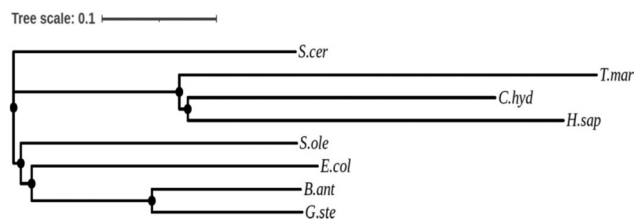
In this work, we report the biochemical characterisation of a hyperthermophilic TK from *T. maritima* DSM3109, with extreme thermostability, as well as investigations into substrates accepted including C<sub>5</sub> and C<sub>6</sub> monosaccharides, sugar derivatives (*e.g.* glucosamine and furfural) and aliphatic aldehydes (C<sub>3</sub>–C<sub>9</sub>) using LiHPA as the ketol donor. In addition, we describe for the first time, TK-catalysed conversions of uronic acids. In particular, the efficient bioconversion of D-galacturonic acid into a rare C<sub>8</sub> sugar 7-keto-octuronic acid is presented.

## Results and discussion

### TK<sub>tmar</sub> protein identification and sequence analysis

TK<sub>tmar</sub> protein and gene sequences were retrieved from UniProtKB (R4NSD7). The gene was 1908 bp long with a 49.6% G + C content. Although a high G + C content (>60%) is associated with thermal adaptation, this is not necessary the case for all thermophilic prokaryotes.<sup>30,31</sup> Several factors have been associated with protein thermostability such as certain amino acid substitutions, hydrophobic cores, buried polar contacts and ion pairs and interactions amongst subunits.<sup>32</sup> TK<sub>tmar</sub> is homodimeric, composed of 635 amino acid residues and has a molecular mass of 69.9 kDa per monomer (as shown in SDS-PAGE after Ni-NTA chromatography purification, ESI Fig. 1†) with a theoretical pI of 5.6 ([http://web.expasy.org/compute\\_pi/](http://web.expasy.org/compute_pi/)). The amino acid composition is considered as the primary factor for protein thermostability prediction. Comparing the protein sequence of TK<sub>tmar</sub> to other well-characterised TKs from *G. stearothermophilus*, *E. coli*, *B. anthracis*, *S. cerevisiae*, *H. sapiens*, *C. hydrogenoformans* and *S. oleracea*; TK<sub>tmar</sub> has a higher content in charged (Asp and Lys) and hydrophobic side chain (Val and Tyr) amino acid residues. TK<sub>tmar</sub> also has a lower frequency of polar uncharged amino acids residues (Gln, Ser and Thr). All of these properties are associated with an improved protein thermostability.<sup>32,33</sup> TK<sub>tmar</sub> has a very low Ala content (7% compared to the 11% average) and 5 Cys residues (compared to only 1 for *G. stearothermophilus* and *C. hydrogenoformans*) that could potentially be forming disulfide bonds as a mechanism for protein stabilisation (ESI Table 1†).<sup>34</sup> It has been suggested that Pro residues in the loop regions play an important role in enzyme thermostability;<sup>25,35</sup> however, TK<sub>tmar</sub> contains only 26 Pro residues (4.1%), which together with *B. anthracis* has the lowest number of Pro residues, amongst the analysed TKs.

Multiple sequence alignment (ESI Fig. 2†) and phylogenetic analysis (Fig. 1), showed a low homology between TK<sub>tmar</sub> and other TKs, with a higher similarity to TK<sub>chy</sub> (35.7%) and TK from *H. sapiens* (30.2%) (ESI Table 2†). In spite of their low sequence identity, 66 amino acid residues are highly conserved amongst all TKs, including the key ones in the active site,<sup>36</sup> except for (based on *E. coli* numbering) Gly156 (ThDP binding), Ile/Leu187 (metal ion binding) and His473 (substrate



**Fig. 1** Phylogenetic analysis of TKs were generated with Clustal Omega and formatted using the iTOL server (<http://itol.embl.de/index.shtml>). *T. maritima* (T.mar), *G. stearothermophilus* (G.ste), *E. coli* (E.col), *B. anthracis* (B.ant), *S. cerevisiae* (S.cer), *H. sapiens* (H.sap), *C. hydrogenoformans* (C.hyd) and *S. oleracea* (S.ole).



binding) which are replaced by Ala, Ala and Gln in TK<sub>tmar</sub>, respectively.

### Biochemical characterisation

TK<sub>tmar</sub> optima temperature and pH were determined by measuring the initial rates (for up to 20 min) of the reaction between glycolaldehyde (GA) and LiHPA to yield L-erythrulose (Ery). The optimum temperature was not possible to estimate based on our set of experiments (max 90 °C), inferring that the TK<sub>tmar</sub> optimal temperature might be around or above the boiling point of water. Notably, the initial rate at 90 °C was 21.7-fold higher compared to the one at 37 °C (Fig. 2A). Most hyperthermophilic enzymes are optimally active at temperatures close to the host organism's optimum growth temperature; however, it has also been reported that some hyperthermophilic enzyme have an optimum temperature above this value, or even above the boiling point of water.<sup>33,37</sup> In addition, the Ery yield after 20 min of reaction also increased with temperature, achieving around 76% at 80 and 90 °C, 8-fold higher compared to at 37 °C (ESI Fig. 3†).

The pH optimum was found between 6 and 6.4 which is below the optima pH reported for other thermophilic TKs (pH 7 and 8);<sup>15,25,26</sup> in fact, TK<sub>tmar</sub> lost 57% of its activity at pH 7.4 compared to pH 6.4 (Fig. 2B).

Apparent kinetic parameters were calculated following the pH-based high-throughput assay with GA and LiHPA as substrates at pH 7 and 65 °C (maximum equipment temperature) (Table 1, ESI Fig. 4†).  $K_M$  calculations indicated that TK<sub>tmar</sub> has a higher affinity towards GA (the aldehyde acceptor) than

**Table 1** TK<sub>tmar</sub> apparent kinetic parameters towards LiHPA and glycolaldehyde (GA), were calculated following the pH-based high-throughput assay with pure TK<sub>tmar</sub> (0.25 mg mL<sup>-1</sup>), GA (final concentration 0.5 to 50 mM – LiHPA constant at 50 mM) or LiHPA (final concentration 5 to 100 mM – GA constant at 50 mM), phenol red (28 mM), ThDP (2.4 mM), MgCl<sub>2</sub> (9 mM). All reaction components were prepared in TEA 2 mM buffer pH 7 and reactions carried out 65 °C. The reactions were monitored at 560 nm for up to 30 min in a plate reader.  $K_M$  and  $V_{max}$  were determined with OriginPro 2018 software. All experiments were carried out in duplicate

	$V_{max}$ (mM min <sup>-1</sup> )	$K_M$ (mM)	$k_{cat}$ (s <sup>-1</sup> )	$k_{cat}/K_M$ (mM <sup>-1</sup> s <sup>-1</sup> )
LiHPA	0.33 ± 0.01	53.8 ± 4.1	0.77 ± 0.03	0.0144 ± 0.001
GA	0.19 ± 0.01	3.6 ± 0.4	0.45 ± 0.03	0.1257 ± 0.006

towards LiHPA, in contrast to other TKs that show more affinity towards the ketol donor (LiHPA) as reported using the same activity assay.<sup>38</sup> Likewise,  $k_{cat}$  and  $k_{cat}/K_M$  values showed a similar trend and they were significantly lower compared to that reported for other TKs.<sup>38</sup>

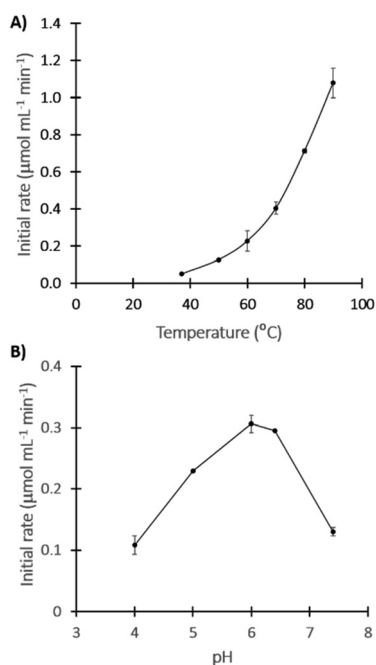
### Stability of TK<sub>tmar</sub>

Enzyme stability was determined by incubating purified TK<sub>tmar</sub> under different conditions. Samples were taken periodically, cooled down and then initial rates were calculated following the TK activity assay with Ery and D-ribose-5-phosphate as substrates. TK<sub>tmar</sub> did not lose activity at 40 °C after 45 h. Despite the loss of activity, which was observed with temperature increase, TK<sub>tmar</sub> was shown to be highly thermostable with  $t_{1/2}$  of 99, 54.1, 21.2 and 9.3 h at 50, 60, 70 and 80 °C, respectively (Fig. 3A). This is a remarkable feature for TK<sub>tmar</sub> compared to TK<sub>gst</sub> that was fully deactivated at 70 °C after 1 h.

Thermophilic enzymes have been shown to be more robust and suited to demanding chemical processes such as those utilising organic solvents.<sup>20,21</sup> TK<sub>tmar</sub> stability in three organic solvents commonly used in industry (DMSO, methanol and acetonitrile) was tested. The data in Fig. 3B and C shows that TK<sub>tmar</sub> is highly stable at low (10% v/v) and high (50% v/v) organic solvent concentrations for up to 30 h, retaining on average 67% and 70% of the initial activity respectively. When the enzyme was incubated in presence of 10% (v/v) organic solvent, the loss of activity profile was similar to the control reaction; interestingly, a more rapid loss of activity during the first hours of incubation was observed at 50% (v/v) of organic solvent but followed by activity stabilisation. TK<sub>tmar</sub> showed better organic solvent stability properties compared to TK<sub>chy</sub> whose loss of activity was around 50% after 1 h incubation in the same organic solvents.<sup>25</sup>

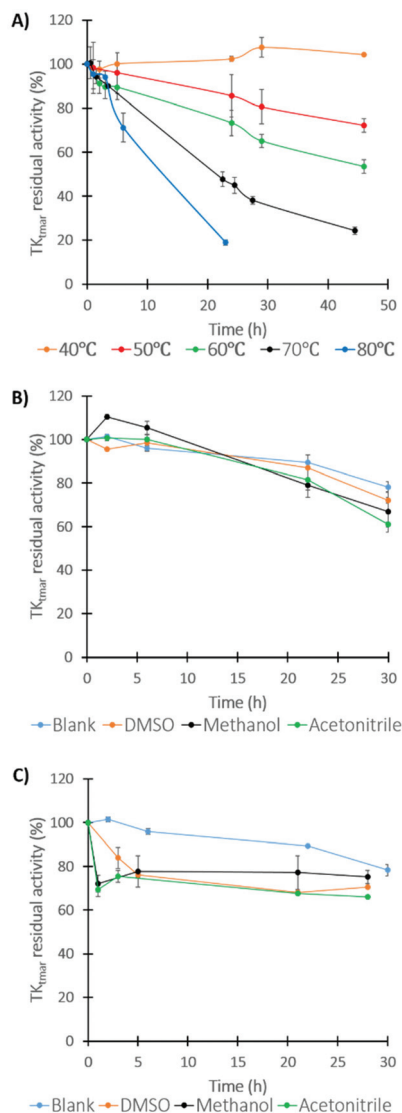
### Substrate scope screening

The TK pH-based high-throughput assay proved to be effective for initial substrate scope screening for a series of aldehydes substrates using LiHPA as ketol donor.<sup>9,38</sup> In this work several aldehyde donors were tested including monosaccharides: C<sub>5</sub> (L-rhamnose, L-arabinose, D-ribose, 2-deoxy-D-ribose and



**Fig. 2** Determination of TK<sub>tmar</sub> optima activity conditions. (A) Temperature, in TRIS-HCl 0.1 M buffer pH 7. (B) pH, performed at 60 °C in acetate 0.1 M buffer (pH 4, 5 and 6) or TRIS-HCl 0.1 M buffer (pH 6.4 and 7.4). All experiments were carried out in duplicate.





**Fig. 3** TK<sub>tmar</sub> stability characterisation. (A) Thermal stability, (B) organic solvent stability 10% (v/v) and (C) organic solvent stability 50% (v/v). All experiments were performed with TK<sub>tmar</sub> (0.7 U mL<sup>-1</sup>), ThDP (2.4 mM) and MgCl<sub>2</sub> (9 mM) in TRIS-HCl 0.1 M buffer pH 7. Organic solvent stability experiments were performed at 50 °C. Enzymatic activity was measured at pH 7.4 and 40 °C following TK<sub>tmar</sub> enzymatic assay. All experiments were carried out in duplicate.

D-xylose); C<sub>6</sub> (D-glucose, D-galactose and D-mannose); uronic acids (D-galacturonic acid and D-glucuronic acid); and monosaccharide derivatives (D-glucosamine and 5-hydroxyfurfural); as well as aliphatic aldehydes (C<sub>3</sub> to C<sub>9</sub>). The initial reaction rates were compared to that for the GA reaction.

It has been described that TKs can accept both non-phosphorylated and phosphorylated substrates, showing higher affinity for the latter (natural substrates).<sup>18</sup> Overall, TK<sub>tmar</sub> exhibited low activity towards both C<sub>5</sub> and C<sub>6</sub> monosaccharides with a relative activity below 30% (Fig. 4). No activity was observed towards D-mannose and the C<sub>6</sub> derivatives. Our results are in agreement with those ones published by Yi *et al.*

where TK<sub>gst</sub> was tested with similar non-phosphorylated monosaccharides.<sup>38</sup> However, Bawn *et al.* showed that TK<sub>dge</sub> and TK<sub>dra</sub> accepted L-arabinose which was upgraded to L-glucuheptulose in high yields.<sup>15</sup> TK<sub>tmar</sub> also showed activity towards C<sub>3</sub> to C<sub>8</sub> aliphatic aldehydes, with the highest relative activity for pentanal (49%). No activity was detected towards nonanal. Likewise, TK<sub>gst</sub> had a very similar profile activity towards these substrates and its activity towards heptanal and octanal was very low.<sup>9</sup>

More interestingly, our results showed that TK<sub>tmar</sub> is active towards uronic acids with relative activities towards D-galacturonic acid (GalAc) and D-glucuronic acid of 52% and 53% respectively. This TK-catalysed reaction would lead to the formation of the respective 7-keto-octuronic acids.

Eight-carbon monosaccharides or octuloses are rare in nature being naturally produced by some plant species and by a few microorganisms but in low concentrations.<sup>39</sup> However, TK catalyses the C-C bond formation between D-xylulose-5-phosphate and D-ribose-5-phosphate yielding a C<sub>7</sub> monosaccharide D-sedoheptulose-7-phosphate.<sup>40,41</sup> In addition, it has been suggested that octulose-phosphates are synthesised in the non-oxidative pentose phosphate pathway catalysed by a TK; for example, TK from spinach catalysed D-glycero-D-ido-octulose-8-phosphate synthesis from fructose-6-phosphate and glucose-6-phosphate *in vivo*.<sup>42</sup> To the best of our knowledge, this is the first study where the synthesis of 7-keto-octuronic acids catalysed by a TK has been definitively reported.

### 7-Keto-octuronic acid preparative synthesis and identification

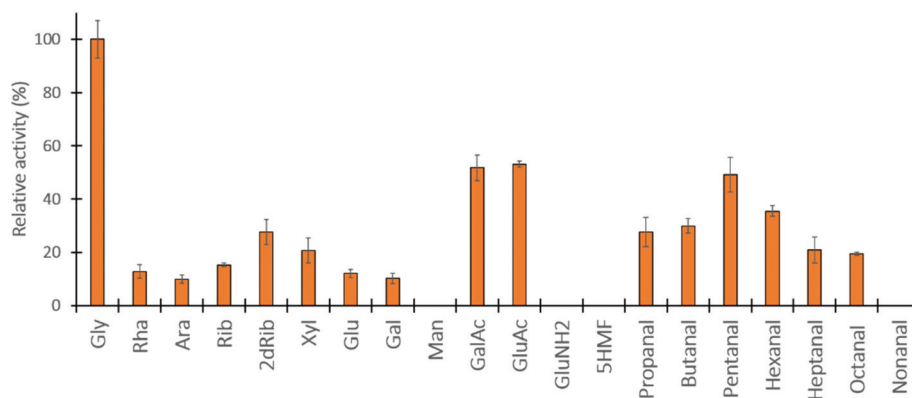
To demonstrate the utility of this new reactivity towards uronic acids, the preparative scale synthesis of (2S,3R,4R,5R)-2,3,4,5,6,8-hexahydroxy-7-oxo-octanoic acid (7-keto-octuronic acid, OctAc) was carried out on a 25 mL reaction volume using GalAc and LiHPA as starting materials (Scheme 2). GalAc is a relatively inexpensive chemical, which is mainly obtained from the hydrolysis of pectin renewable biomass, for example sugar beet pulp.<sup>43,44</sup> The reaction was performed at 50 °C to avoid LiHPA degradation that occurs at higher temperatures, particularly for longer reaction times.<sup>16</sup> By ICS analysis, 84% of GalAc was converted in to OctAc after a 1 h reaction. In order to increase GalAc conversions, 1 mL of LiHPA 50 mM (final concentration 20 mM) was added after 4.5 h reaction and this strategy increased the GalAc conversion to 98%, equivalent to 49 mM of OctAc being produced.

OctAc was then isolated by a Dowex 50WX8 anion exchange column first, to remove excess of TRIS buffer, to give OctAc with a final isolated yield of 90%. <sup>1</sup>H NMR spectroscopy confirmed that OctAc was formed with the major product in the pyranose form (ESI Fig. 5†). In the TK reaction, a new stereogenic centre was formed and considering the 2 possible anomers of the cyclic pyranose form, a mixture of 4 different isomers can be generated.

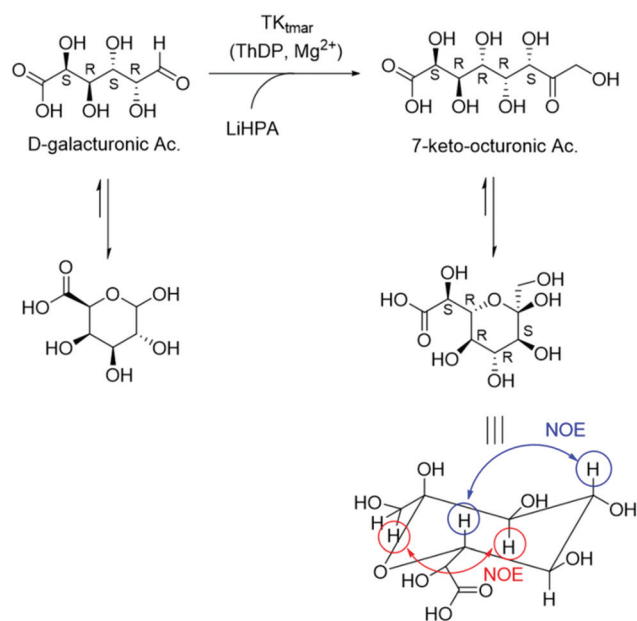
Remarkably, NMR spectroscopy of a freshly prepared sample identified the presence of one major isomer in the form of pyranose. The large *trans* coupling constant between







**Fig. 4**  $TK_{tmar}$  substrate scope screening. Relative activity towards several aldehyde substrates (with glyceraldehyde (GA) as 100%) and LiHPA as carbon donor was calculated following the pH-based high-throughput assay with pure  $TK_{tmar}$  ( $0.25 \text{ mg mL}^{-1}$ ) and pH 7.2 at  $50^\circ\text{C}$ . Reactions were performed with 200 mM of monosaccharides: L-rhamnose (Rha), L-arabinose (Ara), D-ribose (Rib), 2-deoxy-D-ribose (2dRib), D-xylose (Xyl), D-glucose (Glu), D-galactose (Gal), D-mannose (Man), D-galacturonic acid (GalAc), D-glucuronic acid (GluAc), D-glucosamine (GluNH<sub>2</sub>), 5-hydroxyfurfural (5HMF). Aliphatic aldehydes (propanal, butanal, pentanal, hexanal, heptanal, octanal and nonanal) were tested at 100 mM with addition of DMSO (20%). All experiments were carried out in duplicate.



**Scheme 2**  $TK_{tmar}$  catalysed the high-yield synthesis of 7-keto-octuronic acid from D-galacturonic acid and lithium hydroxypyruvate (LiHPA) as carbon donor.

the axial protons at 5-H and 6-H also confirmed that the hydroxyl group at C-6 was in an equatorial orientation giving an *S*-configuration to the newly generated stereocentre at this position. The TK catalysed reaction showed exceptional diastereoselectivity leading to a ketose with the *D*-threo configuration, which is consistent with the stereoselectivity observed in TK reactions with other aldose sugars.<sup>6</sup> The major anomer was identified to be the more stable  $\alpha$ -anomer; however, OctAc exhibited mutarotation behaviour; a solution of the acid in D<sub>2</sub>O was analysed by NMR spectroscopy after several days showing an increasing amount of the other

isomers (see ESI Fig. 6† for comparison <sup>1</sup>H NMR analyses after 2 and 4 weeks).

## Experimental

### Materials

All chemicals were purchased from Merck KGaA (Sigma-Aldrich) and molecular biology reagents were from New England Biolabs, unless otherwise stated. LiHPA was prepared as previously reported.<sup>45</sup>

### Transketolase cloning, expression and purification

The protein sequence of the putative transketolase from *Thermotoga maritima* DSM3109 ( $TK_{tmar}$ ) was retrieved from UniProtKB (ID number: R4NSD7).  $TK_{tmar}$  cloning was carried out with the modified plasmid pET-29: SacB-SapI as previously described.<sup>46</sup> PCR primers were designed by adding overhangs homologous to the pET-29: SacB-SapI vector; forward: 5'-AAAGCTCTTCGATGGAAAGGTTTCTATGA-3' ( $T_m = 60.3^\circ\text{C}$ ) and reverse: 5'-AAAGCTCTTCGGTGGAGCATCTCTCTGAGTCTGG-3' ( $T_m = 57.7^\circ\text{C}$ ).  $TK_{tmar}$  gene was amplified using Phusion® High-Fidelity PCR Master Mix, 100 ng of genomic *T. maritima* DNA, 0.5  $\mu\text{M}$  of each primer and DMSO 5% in 50  $\mu\text{L}$  reaction. PCR was carried out as follow: initial denaturation,  $98^\circ\text{C}$  for 3 min; 30 cycles of denaturation  $98^\circ\text{C}$  for 10 s, annealing  $58^\circ\text{C}$  for 30 s, extension  $72^\circ\text{C}$  for 45 s; and final extension,  $72^\circ\text{C}$  for 10 min. The amplicon was confirmed by agarose gel electrophoresis (1% in TBE buffer) and PCR products were then recovered from the gel following Monarch® DNA Gel Extraction Kit protocol.  $TK_{tmar}$  cloning was performed with a one-pot restriction-ligation reaction method;<sup>46</sup> transformed into chemically competent *E. coli* NovaBlue that were plated on LB agar supplemented with sucrose 20% and kanamycin  $50 \mu\text{g mL}^{-1}$  and grown at  $37^\circ\text{C}$  overnight. Colonies were picked and then grown overnight in 10 mL of LB broth with kanamycin  $50 \mu\text{g mL}^{-1}$  at  $37^\circ\text{C}$  and 250



rpm, and plasmids were extracted using QIAprep Spin Miniprep kit (Qiagen) and sent for sequencing (Eurofins Genomics) to confirm gene sequences. Positive plasmid construct was then transformed into expression host *E. coli* BL21(DE3).

For expression, recombinant TK<sub>tmr</sub> (containing a C-terminal His<sub>6</sub>Tag) was grown in a 2 L shake-flask with 200 mL of TB broth with glycerol 8 mL L<sup>-1</sup>, kanamycin 50 µg mL<sup>-1</sup> at 37 °C and 250 rpm. IPTG was added (0.1 mM final concentration) when the OD<sub>600 nm</sub> reached ~1.5, then kept at 25 °C and 250 rpm for 15 h. Cells were recovered by centrifugation (10 000 rpm at 4 °C for 30 min) and the cell pellet resuspended in TRIS-HCl 50 mM buffer pH 7 (cell concentration 0.5 g mL<sup>-1</sup>) and then lysed by sonication on ice (10 s on and 15 s off for 25 cycles) followed by centrifugation (12 500 rpm at 4 °C for 30 min). The clarified lysate was kept at 4 °C (or -20 °C for up to one month) for further analysis.

TK<sub>tmr</sub> was purified by IMAC methodology with a 5 mL Ni-NTA agarose column (Qiagen) following the manufacturer's instructions. Clarified lysate (prepared in 10 mM imidazole buffer) was loaded into pre-equilibrated column (10 mM imidazole buffer), then the column washed (50 mM imidazole buffer) and finally the enzyme was eluted with 0.5 M imidazole buffer. TK<sub>tmr</sub> was finally precipitated with ammonium sulfate 70% saturation and kept at 4 °C. All imidazole solution buffers were prepared in TRIS-HCl 50 mM with NaCl 0.3 M and pH adjusted to 7.

Enzyme expression and purification was confirmed with SDS-PAGE analysis with Novex® TBE 10% gels (Invitrogen) and NuPAGE™ MOPS-SDS running buffer (ThermoFisher Scientific). Protein quantification was carried with Quick Start Bradford Protein Assay (Bio-Rad) using bovine serum albumin as protein standard.

### Transketolases phylogenetic analysis

TKs sequences from different sources were retrieved from UniProtK: *Geobacillus stearothermophilus* (A0A0I9QGZ2), *Escherichia coli* (P27302.5), *Bacillus anthracis* (Q81Y15), *Saccharomyces cerevisiae* (P23254), *Homo sapiens* (P29401), *Carboxydotherrhus hydrogeoformans* (Q3AFP7 and Q3AFP6) and *Spinacia oleracea* (O20250). Multiple sequence alignment and phylogenetic analysis were carried out using Clustal Omega, Jalview and MEGA 10. iTOL server was used for phylogenetic tree visualisation and formatting.

### Transketolase characterisation

All pHs values for TRIS-HCl and triethanolamine (TEA) buffers were adjusted for 37 °C. Final experimental pH value at higher temperatures were estimated by considering a temperature coefficient of -0.028 and -0.02 per °C, respectively.

**Biochemical characterisation.** TK<sub>tmr</sub> optima temperature and pH were determined by measuring the initial rates (for up to 20 min) of the reaction with GA and LiHPA to yield Ery. The reaction contained GA (10 mM), LiHPA (10 mM), ThDP (2.4 mM), MgCl<sub>2</sub> (9 mM) and TK<sub>tmr</sub> (0.05 mg mL<sup>-1</sup>). For optimum temperature experiments, the reaction was carried out in TRIS-HCl 50 mM buffer pH 7 and incubated at 37, 50,

60, 70, 80 or 90 °C. Optimum pH experiments were performed at 60 °C in 0.1 M buffer acetate (pH 4, 5 and 6) or TRIS-HCl (pH 6.4 and 7.4). All experiments were carried out in duplicate, 1 mL reaction volume and in a Eppendorf Thermomixer™ C fitted with a Thermotop (to avoid condensation on tube lid) with agitation at 750 rpm. Samples were taken periodically, mixed with one volume of TFA 0.5% (v/v) and then analysed by HPLC.

**Determination of apparent kinetics parameters.** Kinetic parameters were calculated following the pH-based high-throughput assay described by Yi *et al.*<sup>38</sup> 10 µL of pure TK<sub>tmr</sub> (stock 5 mg mL<sup>-1</sup>) was added to 150 µL of reaction mixture containing phenol red (28 µM), ThDP (2.4 mM), MgCl<sub>2</sub> (9 mM), and GA (final concentration 50, 25, 10, 5, 2.5, 1 or 0.5 mM), this mixture was preincubated at 65 °C for 10 min followed by the addition of 40 µL of preincubated LiHPA (final concentration 50 mM). A similar procedure was performed keeping GA (50 mM) concentration constant and varying LiHPA concentration (final concentration 100, 75, 50, 25, 12.5 or 5 mM). All reaction components were prepared in TEA 2 mM buffer pH 7. The reactions were carried out in duplicates and monitored at 560 nm for up to 30 min in a plate reader. Linear regression ( $R^2 > 0.95$ ) was used to measure the slope of the initial rates. NaHCO<sub>3</sub> (1, 0.75, 0.5, 0.25, 0.1 and 0.05 mM) standard curve was constructed and used for enzyme activity calculation expressed as µmol mL<sup>-1</sup> min<sup>-1</sup>.  $K_M$  and  $V_{max}$  were determined with OriginPro 2018 software.

**Stability studies.** TK<sub>tmr</sub> thermostability was evaluated by incubating pure TK<sub>tmr</sub> (0.7 U mL<sup>-1</sup>), ThDP (2.4 mM) and MgCl<sub>2</sub> (9 mM) in TRIS-HCl 0.1 M buffer pH 7 at 40, 50, 60, 70 and 80 °C for up to 45 h. Likewise, co-solvent stability was performed at similar conditions and in presence of 10 or 50% (v/v) of either DMSO, acetonitrile or methanol and incubated at 50 °C for up to 30 h. Sample were taken periodically, cooled down in an ice bath for 5 min, and the enzymatic activity was measured following the TK<sub>tmr</sub> enzymatic assay.

**Substrate scope screening.** Activity towards several monosaccharides and aliphatic aldehydes as carbon acceptors was screened using the pH-based assay as described above and in buffer TEA 2 mM pH 7.2 at 50 °C. Reactions were performed with 200 mM of monosaccharides: L-rhamnose, L-arabinose, D-ribose, 2-deoxy-D-ribose, D-xylose, D-glucose, D-galactose, D-mannose, D-galacturonic acid (GalAc), D-glucuronic acid; sugar derivatives: D-glucosamine, 5-hydroxyfurfural. Aliphatic aldehydes (propanal, butanal, pentanal, hexanal, heptanal, octanal and nonanal) were also tested with this method at 100 mM final concentration with addition of DMSO 20% (v/v).

### Preparative synthesis of 7-keto-octuronic acid and characterisation

7-Keto-octuronic acid (OctAc) synthesis was carried out with GalAc (50 mM), LiHPA (10 mM), ThDP (2.4 mM), MgCl<sub>2</sub> (9 mM) and TK<sub>tmr</sub> clarified lysate 20% (v/v, 11.5 mg total protein per mL, TK<sub>tmr</sub> represents approximately the 25% of total protein based in densitometry analysis) in TRIS-HCl 50 mM buffer pH 6.6 in a 25 mL total reaction volume at 50 °C



and with moderate agitation with a magnetic stirrer. 1 mL of LiHPA 0.5 M was added after 4.5 h and the reaction was left at same conditions for up to 20 h. The reaction mixture was centrifuged and the supernatant kept at  $-20\text{ }^{\circ}\text{C}$  until analysed. For product purification, an aliquot was taken (7 mL) and the pH was adjusted to 4 with HCl 1 M, then the solution was loaded on to a Dowex 50WX8 ion exchange resin (8 mL) and gently mixed for 20 min in order to remove the TRIS buffer. After this time, the eluent was collected, and the resin washed with further 3 mL of MilliQ water. All the fractions were collected and loaded onto an Ambersep 900 (OH) ion exchange resin (6 mL) and gently mixed for 30 min. The resin was washed several times with MilliQ water ( $2 \times 7\text{ mL}$ ) to remove salt excess. The OctAc product was then eluted with HCl 2 M ( $2 \times 5\text{ mL}$ ) and the resin washed with MilliQ water (5 mL). For each treatment, the mixture was agitated for 5 min prior to elution. The eluent was concentrated under reduced pressure and freeze-dried to give the OctAc as a pale brown residue (77.1 mg, 90%). Mp  $73 \pm 2\text{ }^{\circ}\text{C}$ ;  $^1\text{H}$  NMR ( $\text{CD}_3\text{OD}$ ; 600 MHz) 4.48 (1H, d,  $J = 1.5\text{ Hz}$ , 2-H), 4.11 (1H, dd,  $J = 10.0, 1.5\text{ Hz}$ , 3-H), 3.71 (1H, t,  $J = 9.4\text{ Hz}$ , 5-H), 3.66 (1H, d,  $J = 11.2\text{ Hz}$ , 8-HH), 3.53 (1H, dd,  $J = 10.0, 9.1\text{ Hz}$ , 4-H), 3.50 (1H, d,  $J = 9.6\text{ Hz}$ , 6-H), 3.35 (1H, d,  $J = 11.2\text{ Hz}$ , 8-HH);  $^{13}\text{C}$  NMR ( $\text{CD}_3\text{OD}$ ; 126 MHz) 175.1 (C-1), 99.2 (C-7), 75.6 (C-5), 74.6 (C-3), 71.3 (C-6), 70.8 (C-2), 70.4 (C-4), 65.0 (C-8);  $m/z$  HRMS (ES+)  $[\text{MNa}]^+$  277.0528,  $\text{C}_8\text{H}_{14}\text{O}_9\text{Na}$  requires 277.0536.

### Enzyme activity assay

$\text{TK}_{\text{tmar}}$  activity assay was adapted to a microplate format (200  $\mu\text{L}$  reaction volume per well) following a procedure previously reported by Hecquet *et al.*<sup>47</sup> The reaction mixture contained Ery 100 mM, D-ribose-5-phosphate (9.1 mM), ThDP (0.1 mM),  $\text{MgCl}_2$  (0.5 mM), NADH (0.2 mM), alcohol dehydrogenase from *Saccharomyces cerevisiae* (25 U  $\text{mL}^{-1}$ ) and the  $\text{TK}_{\text{tmar}}$  suspension (10  $\mu\text{L}$ , diluted when necessary). The reaction was carried out in TRIS-HCl 0.1 M buffer pH 7.4 at  $40\text{ }^{\circ}\text{C}$ , and monitored at 340 nm for up to 10 min in a plate reader (CLARIOstar Plus, BMG Labtech) and the initial rate was then calculated. One unit (U) is defined as the amount of  $\text{TK}_{\text{tmar}}$  that catalyses the formation of 1  $\mu\text{mol}$  of glyceraldehyde per minute at  $40\text{ }^{\circ}\text{C}$  at pH 7.4 (value  $\epsilon_{\text{NADH } 340\text{ nm}} = 6220\text{ M}^{-1}\text{ cm}^{-1}$ ).

### Analytical HPLC

LiHPA and Ery were analysed by HPLC with a Ultimate 3000 + HPLC (ThermoFisher Scientific) fitted with a Aminex HPX-87H column (BioRad), TFA 0.1% as mobile phase at  $0.6\text{ mL min}^{-1}$  for 20 min and compounds detected by RI (RefractoMax 520); retention times were 8.3 and 11.5 min respectively. GalAc was analysed using an Ion Chromatography System (ICS 5000+, Thermo Scientific) using 5% (v/v) of 1 M sodium acetate (electrochemical detection grade, Fisher Scientific) with 3.8 min retention time as described by Ward *et al.*<sup>48</sup> All quantitative analyses were performed measuring peak area using the external standard method.

## Conclusions

$\text{TK}_{\text{tmar}}$  is a hyperthermophilic enzyme with unique features compared to previously reported TKs. It has a low protein homology related to well-known TKs, with 10% of amino acid residues highly conserved especially the ones in the active site. It has a high content of charged and hydrophobic side chain amino acid residues both associated with improved protein thermostability.  $\text{TK}_{\text{tmar}}$  characterisation confirmed its remarkable stability at very high temperature (*e.g.*  $t_{1/2}$  of 9.3 h at  $80\text{ }^{\circ}\text{C}$ ) as well as in presence of high concentrations of organic solvents (up to 50% v/v) commonly used in the chemical industry.  $\text{TK}_{\text{tmar}}$  is able to accept a broad range of aldehyde substrates including  $\text{C}_5$  and  $\text{C}_6$  non-phosphorylated monosaccharides and aliphatic aldehydes ( $\text{C}_3$  to  $\text{C}_8$ ). For the first time, we have also reported that a TK can readily accept uronic acids. Preparative scale reaction from GalAc and LiHPA enabled the synthesis of 7-keto-octuronic acid with a high reaction yield (98%) and concentration of 49 mM. This novel  $\text{TK}_{\text{tmar}}$  is an excellent alternative to the current mesophilic and thermostable TKs reported. However, more studies regarding the molecular structure are needed to better understand the  $\text{TK}_{\text{tmar}}$  mechanism and how it withstands very high temperatures, as well as to expand its substrate scope further.

## Conflicts of interest

The authors declare that they have no conflict of interests.

## Acknowledgements

M. Cárdenas-Fernández and J. M. Ward would like to thank the UK Biotechnology and Biological Sciences Research Council (BBSRC) for financial support of this work with BB/R021627/1 (ERA CoBioTech Grant). F. Subrizi and H. C. Hailes would like to thank the financial support from BBSRC BB/R021643/1 (ERACoBioTech Grant). Input and advice from the project collaborators from the ERABioCat European Project TRALAMINOL is also acknowledged.

## References

- 1 R. A. Sheldon and J. M. Woodley, *Chem. Rev.*, 2018, **118**, 801–838.
- 2 J. M. Woodley, *Curr. Opin. Green Sustain. Chem.*, 2020, **21**, 22–26.
- 3 E. Fullam, F. Pojer, T. Bergfors, T. A. Jones and S. T. Cole, *Open Biol.*, 2012, **2**, 1–12.
- 4 G. Schenk, R. G. Duggleby and P. F. Nixon, *Int. J. Biochem. Cell Biol.*, 1998, **30**, 1297–1318.
- 5 J. Zhao and C.-J. Zhong, *Neurosci. Bull.*, 2009, **25**, 94–99.
- 6 U. Schörken and G. A. Sprenger, *Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol.*, 1998, **1385**, 229–243.



- 7 H. C. Hailes, D. Rother, M. Müller, R. Westphal, J. M. Ward, J. Pleiss, C. Vogel and M. Pohl, *FEBS J.*, 2013, **280**, 6374–6394.
- 8 J. Shaeri, I. Wright, E. B. Rathbone, R. Wohlgemuth and J. M. Woodley, *Biotechnol. Bioeng.*, 2008, **101**, 761–767.
- 9 H. Casajus, A. Lagarde, M. Leremboure, T. De Dios Miguel, L. Nauton, V. Thery, W.-D. Fessner, N. Duguet, F. Charmantray and L. Hecquet, *ChemCatChem*, 2020, **12**, 5772–5779.
- 10 H. Yu, R. I. Hernández López, D. Steadman, D. Méndez-Sánchez, S. Higson, A. Cázares-Körner, T. D. Sheppard, J. M. Ward, H. C. Hailes and P. A. Dalby, *FEBS J.*, 2020, **287**, 1758–1776.
- 11 P. Payongsri, D. Steadman, J. Strafford, A. MacMurray, H. C. Hailes and P. A. Dalby, *Org. Biomol. Chem.*, 2012, **10**, 9021–9029.
- 12 A. Ranoux, S. K. Karmee, J. Jin, A. Bhaduri, A. Caiazza, I. W. C. E. Arends and U. Hanefeld, *ChemBioChem*, 2012, **13**, 1921–1931.
- 13 J. L. Galman, D. Steadman, S. Bacon, P. Morris, M. E. B. Smith, J. M. Ward, P. A. Dalby and H. C. Hailes, *Chem. Commun.*, 2010, **46**, 7608–7610.
- 14 F. Subrizi, M. Cárdenas-Fernández, G. J. Lye, J. M. Ward, P. A. Dalby, T. D. Sheppard and H. C. Hailes, *Green Chem.*, 2016, **18**, 3158–3165.
- 15 M. Bawn, F. Subrizi, G. J. Lye, T. D. Sheppard, H. C. Hailes and J. M. Ward, *Enzyme Microb. Technol.*, 2018, **116**, 16–22.
- 16 M. Lorillière, M. De Sousa, F. Bruna, E. Heuson, T. Gefflaut, V. de Berardinis, T. Saravanan, D. Yi, W.-D. Fessner, F. Charmantray and L. Hecquet, *Green Chem.*, 2017, **19**, 425–435.
- 17 M. F. Villegas-Torres, J. M. Ward and F. Bagan, *J. Biotechnol.*, 2018, **281**, 150–160.
- 18 N. J. Turner, *Curr. Opin. Biotechnol.*, 2000, **11**, 527–531.
- 19 A. Ranoux and U. Hanefeld, *Top. Catal.*, 2013, **56**, 750–764.
- 20 J. Atalah, P. Cáceres-Moreno, G. Espina and J. M. Blamey, *Bioresour. Technol.*, 2019, **280**, 478–488.
- 21 J. Littlechild, *Front. Bioeng. Biotechnol.*, 2015, **3**, 1–9.
- 22 R. Huber, T. A. Langworthy, H. König, M. Thomm, C. R. Woese, U. B. Sleytr and K. O. Stetter, *Arch. Microbiol.*, 1986, **144**, 324–333.
- 23 D. Rodionov, I. Rodionova, X. Li, D. Ravcheev, Y. Tarasova, V. Portnoy, K. Zengler and A. Osterman, *Front. Microbiol.*, 2013, **4**, 1–14.
- 24 K. M. Noll, in *Brenner's Encyclopedia of Genetics (Second Edition)*, ed. S. Maloy and K. Hughes, Academic Press, San Diego, 2013, pp. 63–65. DOI: 10.1016/B978-0-12-374984-0.01535-7..
- 25 P. James, M. N. Isupov, S. A. De Rose, C. Sayer, I. S. Cole and J. A. Littlechild, *Front. Microbiol.*, 2020, **11**, 1–13.
- 26 J. Abdoul-Zabar, I. Sorel, V. Hélaine, F. Charmantray, T. Devamani, D. Yi, V. de Berardinis, D. Louis, P. Marlière, W.-D. Fessner and L. Hecquet, *Adv. Synth. Catal.*, 2013, **355**, 116–128.
- 27 T. Saravanan, M.-L. Reif, D. Yi, M. Lorillière, F. Charmantray, L. Hecquet and W.-D. Fessner, *Green Chem.*, 2017, **19**, 481–489.
- 28 T. Saravanan, S. Junker, M. Kickstein, S. Hein, M.-K. Link, J. Ranglack, S. Witt, M. Lorillière, L. Hecquet and W.-D. Fessner, *Angew. Chem., Int. Ed.*, 2017, **56**, 5358–5362.
- 29 C. Zhou, T. Saravanan, M. Lorillière, D. Wei, F. Charmantray, L. Hecquet, W.-D. Fessner and D. Yi, *ChemBioChem*, 2017, **18**, 455–459.
- 30 L. D. Hurst and A. R. Merchant, *Proc. R. Soc. Lond. B Biol. Sci.*, 2001, **268**, 493–497.
- 31 K. B. Zeldovich, I. N. Berezovsky and E. I. Shakhnovich, *PLoS Comput. Biol.*, 2007, **3**, 62–72.
- 32 Q. Wang, Z. Cen and J. Zhao, *Physiology*, 2015, **30**, 97–106.
- 33 C. Vieille and G. J. Zeikus, *Microbiol. Mol. Biol. Rev.*, 2001, **65**, 1–43.
- 34 D. R. Boutz, D. Cascio, J. Whitelegge, L. J. Perry and T. O. Yeates, *J. Mol. Biol.*, 2007, **368**, 1332–1344.
- 35 Y. Suzuki, K. Oishi, H. Nakano and T. Nagayama, *Appl. Microbiol. Biotechnol.*, 1987, **26**, 546–551.
- 36 P. Asztalos, C. Parthier, R. Golbik, M. Kleinschmidt, G. Hübner, M. S. Weiss, R. Friedemann, G. Wille and K. Tittmann, *Biochemistry*, 2007, **46**, 12037–12052.
- 37 C. Vieille, D. S. Burdette and J. G. Zeikus, in *Biotechnology Annual Review*, ed. M. R. El-Gewely, Elsevier, 1996, vol. 2, pp. 1–83.
- 38 D. Yi, T. Devamani, J. Abdoul-Zabar, F. Charmantray, V. Helaine, L. Hecquet and W.-D. Fessner, *ChemBioChem*, 2012, **13**, 2290–2300.
- 39 Q. Zhang and D. Bartels, *J. Exp. Bot.*, 2017, **68**, 5689–5694.
- 40 G. Schenk, R. G. Duggleby and P. F. Nixon, *Int. J. Biochem. Cell Biol.*, 1998, **30**, 1297–1318.
- 41 J. Zhao and C. J. Zhong, *Neurosci. Bull.*, 2009, **25**, 94–99.
- 42 J. F. Williams, P. F. Blackmore and M. G. Clark, *Biochem. J.*, 1978, **176**, 257–282.
- 43 M. Cárdenas-Fernández, C. Hamley-Bennett, D. J. Leak and G. J. Lye, *Bioresour. Technol.*, 2018, **269**, 195–202.
- 44 M. Cárdenas-Fernández, M. Bawn, C. Hamley-Bennett, P. K. V. Bharat, F. Subrizi, N. Suhaili, D. P. Ward, S. Bourdin, P. A. Dalby, H. C. Hailes, P. Hewitson, S. Ignatova, C. Kontoravdi, D. J. Leak, N. Shah, T. D. Sheppard, J. M. Ward and G. J. Lye, *Faraday Discuss.*, 2017, **202**, 415–431.
- 45 K. G. Morris, M. E. B. Smith, N. J. Turner, M. D. Lilly, R. K. Mitra and J. M. Woodley, *Tetrahedron: Asymmetry*, 1996, **7**, 2185–2188.
- 46 D. Dobrijevic, L. A. Nematollahi, H. H. Hailes and J. M. Ward, *BioTechniques*, 2020, **69**, 384–387.
- 47 L. Hecquet, J. Bolte and C. Demuynck, *Biosci., Biotechnol., Biochem.*, 1993, **57**, 2174–2176.
- 48 D. P. Ward, P. Hewitson, M. Cárdenas-Fernández, C. Hamley-Bennett, A. Díaz-Rodríguez, N. Douillet, J. P. Adams, D. J. Leak, S. Ignatova and G. J. Lye, *J. Chromatogr., A*, 2017, **1497**, 56–63.

