



Cite this: *Green Chem.*, 2017, **19**, 481

## Engineering a thermostable transketolase for arylated substrates†‡

Thangavelu Saravanan,<sup>a</sup> Marie-Luise Reif,<sup>a</sup> Dong Yi,<sup>a</sup> Marion Lorillière,<sup>b,c</sup> Franck Charmantray,<sup>b,c</sup> Laurence Hecquet<sup>b,c</sup> and Wolf-Dieter Fessner<sup>\*a</sup>

Aromatic components are difficult substrates for enzymes catalyzing stereoselective carbonylation reactions. We have engineered transketolase from *Geobacillus stearothermophilus* by directed evolution to utilize arylalkanal and benzaldehyde as the electrophilic substrate in highly stereoselective C–C bond forming conversions. Enzyme variants were discovered with rate accelerations up to 28-fold that convert 2-phenylethanal, 3-phenylpropanal, phenyloxyethanal, benzyloxyethanal, and (*N*-Cbz)-3-aminopropanal with formation of the corresponding aryl-substituted 1,3-dihydroxyketones in good yields (60–72%) and virtually complete (*3S*)-stereoselectivity (>99% ee). Novel double-site variants were also found for the conversion of benzaldehyde.

Received 22nd July 2016,  
Accepted 3rd October 2016  
DOI: 10.1039/c6gc02017h

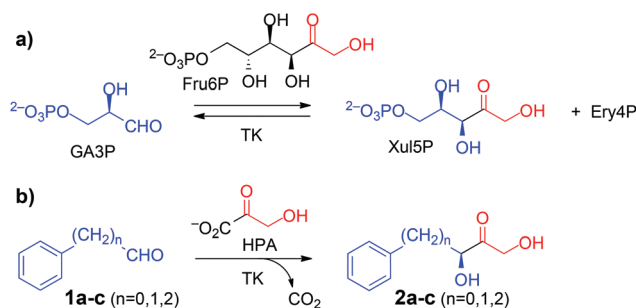
www.rsc.org/greenchem

## Introduction

Non-covalent interactions involving aromatic rings such as  $\pi$ - $\pi$  stacking, O–H/ $\pi$ , and cation– $\pi$  interactions are key stabilizing elements in both chemical and biological recognition. Because of their hydrophobic nature and low chemical reactivity, aromatic ring systems are of paramount importance as constituents of synthetic building blocks for rational drug design and lead optimization in medicinal chemistry.<sup>1</sup> Sustainability issues are progressively fostering the use of biocatalysis as a green, eco-friendly technology with significant potential for the chemical and pharmaceutical industries due to the advantages of their mild reaction conditions and intrinsically high chemo- and stereoselectivities.<sup>2</sup> Enzyme-catalysed asymmetric carbonylation is an emerging tool, yet aromatic components are of limited general utility for synthetic applications, because such compounds not only have low solubility in aqueous media but also low reactivity as electrophiles for aldolases.<sup>3,4</sup> Aryl-substituted chiral building blocks are more suitable as substrates for thiamine diphosphate (ThDP) dependent enzymes,<sup>5</sup> exemplified in the technical synthesis of (*R*)-phenylacetylcarbinol [(*R*)-PAC] as precursor to ephedrine and other

alkaloids catalysed by pyruvate decarboxylase and related biocatalysts.<sup>6</sup>

Transketolase (TK, EC 2.2.1.1), a key enzyme in metabolic regulation that serves to connect the pentose phosphate pathway to glycolysis, also is a member of the ThDP dependent enzymes toolbox.<sup>3,7</sup> *In vivo*, it performs the interconversion of various phosphorylated sugars by catalyzing a reversible transfer of a two-carbon ketol fragment from a ketose phosphate to an aldose phosphate (Scheme 1a). To render synthetic reactions irreversible *in vitro*, hydroxypyruvate (HPA) is used to generate the ketol nucleophile with release of CO<sub>2</sub>.<sup>8</sup> Carbonylation is both stereospecific, in that the new asymmetric centre formed has the (*S*) configuration, and stereoselective, in that the enzyme shows a strong kinetic preference for  $\alpha$ -hydroxyaldehydes having a (2*R*)-configuration. This makes TK a potent biocatalyst for the asymmetric synthesis of chiral ketotriol products that have the final *D*-threo or (3*S*,4*R*) configuration. TK shows high specificity towards its nucleophi-



**Scheme 1** (a) Natural versus (b) desired catalytic reactions of transketolase.

<sup>a</sup>Institut für Organische Chemie und Biochemie, Technische Universität Darmstadt, Alarich-Weiss-Str. 4, 64287 Darmstadt, Germany. E-mail: fessner@tu-darmstadt.de

<sup>b</sup>Clermont Université, Université Blaise Pascal, Institut de Chimie de Clermont-Ferrand, BP10448, 63000 Clermont-Ferrand, France

<sup>c</sup>CNRS, UMR 6296, ICCF, 63177 Aubière, France

† Dedicated to Romas Kazlauskas on the occasion of his 60th birthday.

‡ Electronic supplementary information (ESI) available: Enzyme kinetic and stereoselectivity data, HPLC profiles, copies of original NMR spectra and NMR shift experiments. See DOI: 10.1039/c6gc02017h



lic ketol substrates but is more tolerant towards structural modification in the electrophilic aldehyde substrate.<sup>3,9,10</sup> However, specific activity towards aliphatic aldehydes lacking a 2-hydroxylation is about 40-fold lower than for the corresponding hydroxylated derivatives.<sup>11</sup> Also, the wild-type TK from *E. coli* has been reported not to convert benzaldehyde (**1a**) and the related heteroaromatic 2-furaldehyde and 2-thiophene carboxaldehyde,<sup>12,13</sup> and to accept only a few arylated aldehydes such as phenylethanal (**1b**) with very low productivity.<sup>12,14</sup>

Recently, we have identified the TK from *Geobacillus stearothermophilus* (TK<sub>gst</sub>) as a novel thermostable enzyme with high potential for practical applications and future biocatalyst development.<sup>15</sup> TK<sub>gst</sub> offers a temperature optimum around 65 °C, remains active for three days at 65 °C, shows improved tolerance against deactivation by aldehyde electrophiles and has a high tolerance towards non-conventional media such as organic solvents and ionic liquids.<sup>16</sup> Also, thermostability of the TK<sub>gst</sub> scaffold has been demonstrated to correlate well with mutational robustness.<sup>17,18</sup> TK<sub>gst</sub> has high protein sequence and structural similarity to TKs from *E. coli* (TK<sub>eco</sub>), *S. cerevisiae* and *B. anthracis*, for which protein crystal structures have been determined with bound substrates for a detailed mechanistic understanding.<sup>19</sup> Our previous structure-guided work to broaden the substrate scope of TK<sub>gst</sub> towards selected classes of non-natural substrates was targeted at two opposing residues Leu382 and Asp470 in the active-site that are responsible for substrate binding of the electrophilic reaction partner and for controlling the enantioselective binding of 2-hydroxyaldehydes by specific hydrogen bonding to the OH group (Fig. 1). Indeed, protein libraries L382X/D470X generated by site-specific saturation mutagenesis (SSM) furnished novel catalysts with significantly improved activity for the conversion of generic aliphatic substrates such as propanal,<sup>17</sup> or with relaxed enantiomer selectivity for the conversion of non-natural (2*S*)-configured hydroxyaldehydes including L-lactaldehyde and L-glyceraldehyde.<sup>18</sup> However, no aryl-substi-

tuted aldehydes (Scheme 1b) have been tested yet as potential substrates for TK<sub>gst</sub>.

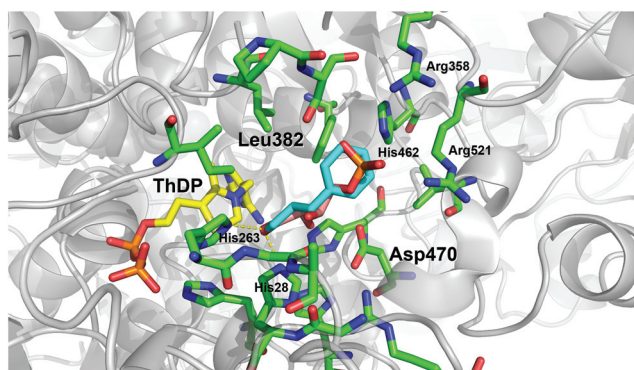
Similarly, protein engineering of TK<sub>eco</sub> has been successfully used to improve its activity by single-point saturation mutagenesis towards aliphatic aldehydes, with both enhanced and reversed stereospecificity,<sup>5f,20</sup> and also improved activity towards aldose substrates was found by rational mutagenesis.<sup>21</sup> Variants of TK<sub>eco</sub> evolved on propanal were found to show nascent activity on aromatic aldehydes<sup>12</sup> that later could be improved by rational engineering of protein and substrate to increase polar contacts, e.g. by using formylbenzoic acids that profit from better solubility and extra binding interaction between their carboxylic acid group and the phosphate-binding residues in the TK active-site.<sup>22</sup>

Still, a breakthrough is required for the conversion of generic aryl-containing compounds as TK substrates to broaden the scope of this enzyme class for novel applications within *in vitro* metabolic cascades along the general concept of *Systems Biocatalysis*.<sup>23</sup> Here we report on our approach toward novel TK<sub>gst</sub> variants that show rate accelerations up to 28-fold for the conversion of various non-functionalized phenylalkanal of the generic structure Ph(CH<sub>2</sub>)<sub>n</sub>CHO (**1a-c**, *n* = 0–2), as well as of structurally related arylated substrates, that allow the synthesis of the corresponding 1,3-dihydroxyketones in good yields (50–73%) and virtually complete (3*S*)-stereoselectivity (>99% ee).

## Results and discussion

### Library selection and pH shift screening for improved activity with 3-phenylpropanal

Arene-containing substrates in terms of their reactivity and structure are far distinct from the natural sugar phosphates used *in vivo*. Thus, TK<sub>gst</sub> variants able to accept phenylalkanal as substrates require more hydrophobic contribution to substrate binding and a sterically relaxed environment for an accommodation of the arene moiety. All TK enzymes have a highly conserved active-site environment to form a channel for entrance of the natural substrates (phosphorylated sugars) to the catalytic site at the ThDP cofactor, where residues that bind the acceptor substrate (an aldose phosphate such as D-erythrose 4-phosphate; Fig. 1) orient its electrophilic carbonyl group towards a nucleophilic attack. These residues are putative candidates to alter the TK substrate specificity. The original L382X/D470X library was constructed to allow for a cooperative adaptation of the TK<sub>gst</sub> active site towards various modifications in the electrophilic substrate, such as stereochemical inversion,<sup>18</sup> as well as lower oxygenation state, increasing steric bulk, and improved hydrophobic contact zone.<sup>17</sup> In particular, this library has proven useful for the successful identification of variants that are highly active and highly stereoselective for the conversion of generic alkanal,<sup>17</sup> and therefore it seemed to be appropriate also for binding of aryl-containing substrate analogues. Depending on the specific Ph(CH<sub>2</sub>)<sub>n</sub>CHO constitution the aryl unit might be expected to



**Fig. 1** Illustration of the potential binding orientation for arylated aldehydes by superimposing **1c** in minimum conformation (cyan) with the experimental binding mode of erythrose 4-phosphate (salmon) in the active site of the yeast TK (PDB entry 1NGS).<sup>25</sup> Residue numbering is for TK<sub>gst</sub>.<sup>15</sup> Graphic was produced with PyMOL.<sup>26</sup>



occupy a variable position in the substrate-binding channel from close to the reaction centre (for **1a**,  $n = 0$ ) up to close to the active site entrance (for **1c**,  $n = 2$ ). In the former case, the L382X/D470X library was expected to furnish suitable residue combinations that could accommodate even a bulky phenyl moiety, while for the latter case those residues forming the phosphate-binding site might reasonably be assumed to offer sufficient flexibility to also adapt a phenyl structure having a diameter similar to that of a phosphate ester moiety (Fig. 1) although its polarity is quite different.

Out of the many assay methods for measuring TK activity,<sup>24</sup> our recently developed colorimetric pH based assay method was chosen for library screening because of its simplicity, speed, universality, sensitivity, and low cost.<sup>11</sup> The TK catalysed condensation of HPA and an aldehyde consumes one proton per cycle owing to the release of carbon dioxide, which in low-buffer medium causes a pH shift to alkaline and results in a colour change in the presence of phenol red as a pH indicator. This assay principle is independent of the structure of an electrophilic substrate and can be applied in microtiter plate format for the rapid substrate screening of TK as well as for a reliable determination of its kinetic constants.<sup>11</sup>

Among the phenylalkanal series we selected 3-phenylpropanal (**1c**) as the most representative substrate for a primary screening, because 2-phenylethanal (**1b**) tends to form an enol tautomer in aqueous solution that reduces the available aldehyde concentration, and benzaldehyde (**1a**) suffers from lower reactivity as an electrophile owing to resonance stabilization. The assay conditions had to be optimized, however, for this specific substrate class because of their low solubility in the reaction buffer, even in the presence of 20% DMSO added as a co-solvent. Addition of 20% DMSO did not alter the activity of TK<sub>gst</sub> and also did not impact the assay fidelity. A homogenous assay solution could only be achieved at 10 mM substrate concentration, whereas at 25 mM incipient cloudiness of the solution precluded an accurate colorimetric measurement (Fig. 2). However, using substrate concentrations as low as 10 mM might be beneficial by enforcing more selective screening conditions that should favour enzyme variants having an improved substrate affinity.

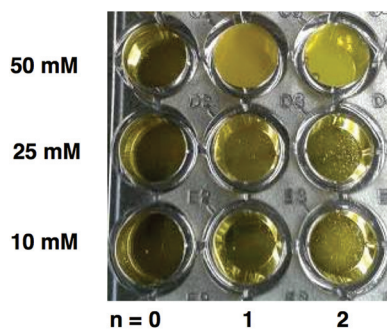


Fig. 2 Solubility limits of phenylalkanal substrates Ph(CH<sub>2</sub>)<sub>n</sub>CHO (**1a–c**) for assay development using buffer with 20% DMSO added as co-solvent.

As a short-cut we chose to screen the positive sub-set that was obtained by previous screening of the entire L382X/D470X library for activity with aliphatic aldehydes such as propanal.<sup>17</sup> The focus on this set of 198 clones confirmed as active hits with simple aliphatic substrates seemed plausible for the following reasons: firstly because of the structural similarity of the propanal moiety that offered high likelihood for activity also with the corresponding aryl-substituted substrate, secondly because inactive or incorrectly folded variants were already excluded from the library, and thirdly because sequences for this hit collection had already been determined, which offered the most economical interpretation for any new hits to be identified from a recurrent screening campaign.

Screening the active variant collection with the adapted high-throughput pH-assay method indeed yielded mostly active catalysts from which a set of 15 best hits was selected for further study that surpassed a threshold of more than 8-fold activity when compared to wild-type TK<sub>gst</sub>. Identified hits were confirmed by re-screening, then the protein variants were expressed and purified for individual kinetic characterization (Fig. 3, Table S1; see ESI†).

Interestingly, while 9 hits (60%) out of this set of 15 best catalysts for 3-phenylpropanal (**1c**) were also found among the best variants previously identified to be active with propanal and other aliphatic substrates,<sup>17</sup> the best hits overall for **1c** were identified as new candidates that were found to be less active with propanal. In particular, the best hit discovered for conversion of **1c** was the variant L382 M/D470L, which has an activity of 28-fold over wild-type TK<sub>gst</sub> but had shown only 3-fold higher activity when screened with propanal.<sup>17</sup> This improved activity could be due to increased hydrophobic binding affinity of the arylated aldehyde **1c** in the active site.

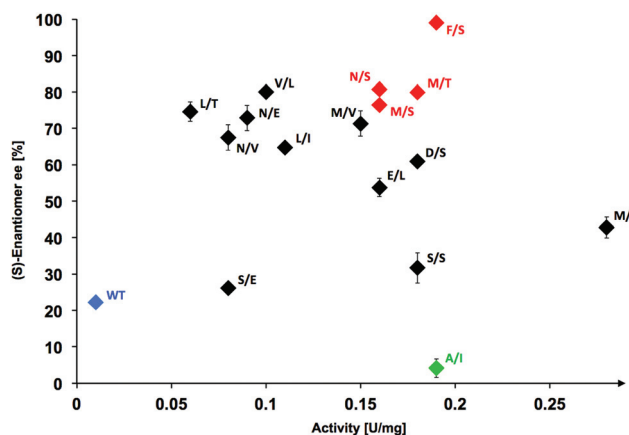


Fig. 3 Property spread for TK<sub>gst</sub> variants identified from L382X/D470X library that were identified for >8-fold higher activity with 3-phenylpropanal (**1c**) than wild-type enzyme (blue). Variants showing highest effectiveness (red) and lowest stereoselectivity (green) are highlighted in color. Variants are labeled with one-letter code for clarity. Assay solution (200  $\mu$ L) contained TK<sub>gst</sub> variant (25  $\mu$ g), Li-HPA (50 mM), **1c** (10 mM), ThDP (2.4 mM), MgCl<sub>2</sub> (9 mM), phenol red (0.028 mM), DMSO (20% v/v) and triethanolamine buffer (2 mM, pH 7.5). The ee of product **2c** was determined by chiral HPLC analysis.



Determination of stereoselectivity for active TK<sub>gst</sub> variants

The stereoselectivity of the active variants in the carboligation step with 3-phenylpropanal (**1c**) was determined by chiral HPLC analysis of the products formed upon analytical-scale synthesis reactions performed at room temperature. Wild-type TK<sub>gst</sub> produced the carboligation product **2c** rather sluggishly with only low 22% ee, which was rather unexpected in view of the much higher 79% ee observed for addition to propanal.<sup>17</sup> The most active variant L382 M/D470L showed only 43% ee in the carboligation step, which seems to indicate that the mutation probably improves the general binding affinity without significantly improving the carbonyl orientation. The majority of the top active variants (10/15 = 67%) catalysed the addition with 54–81% ee.

Replacement of Asp470 by non-polar residues mostly led to catalysts that convert **1c** with low selectivity, while small polar residues (Ser, Thr) seem to be a better compromise to support both good rate and selectivity; e.g., the single variant D470T showed 75% ee for **1c**. The least selective catalyst proved to be the L382A/D470I variant, which gave almost racemic product (4% ee). The most selective variant L382F/D470S showed a perfect product ee of >99% at a very high 19-fold rate acceleration. With propanal, the latter variant had been determined previously to show only 5-fold rate increase and 94% ee.<sup>17</sup> It is interesting to note that the unique L382F mutation apparently is critical for an optimally controlled orientation of the aldehyde carbonyl group in the carboligation transition state to induce maximum stereoselectivity, because we previously observed a similar preference for the L382F mutation when investigating the selectivity with aliphatic aldehyde substrates.<sup>17</sup>

## Analysis of substrate scope

In a subsequent stage, the most effective variants verified from the screening with **1c**, i.e. L382F/D470S, L382N/D470S, L382 M/D470S and L382 M/D470T as judged by rate and stereoselectivity, were also tested for their activity profile with structurally related substrates (Fig. 4, Table S2; see ESI<sup>†</sup>). Wild-type TK<sub>gst</sub> and the most active variant L382 M/D470L were included for comparison. Substrate modifications considered in this study were the lower homologue 2-phenylethanal (**1b**) and a few heteroatom containing substrate analogues such as phenyloxyethanal (equivalent in size to **1c**), benzyloxyethanal as well as (*N*-Cbz)-3-aminopropanal that were included to represent compounds of varying polarity and molecular dimensions. Substrate **1b** showed a rate profile very similar to that for **1c**, and the other aldehydes were also accepted by the variants tested with improved rate (up to 19-fold activity relative to wild-type). Interestingly, reactions of both the aryl ether components showed almost identical activity when compared to each other with all of the catalysts considered, but reactions were consistently somewhat less accelerated than those of pure arylkanals, despite the fact that the oxygenated compounds should profit from a higher electrophilic reactivity. Possibly, the oxygen atom is experiencing electrostatic repulsive inter-

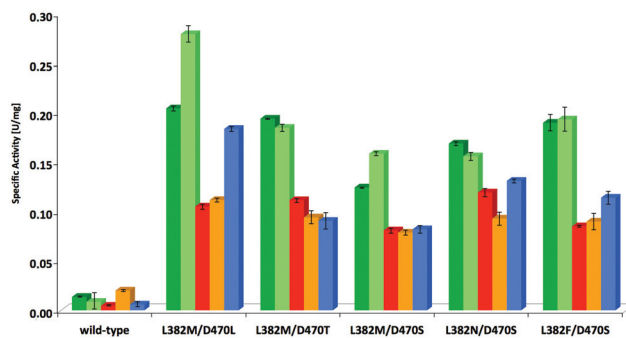
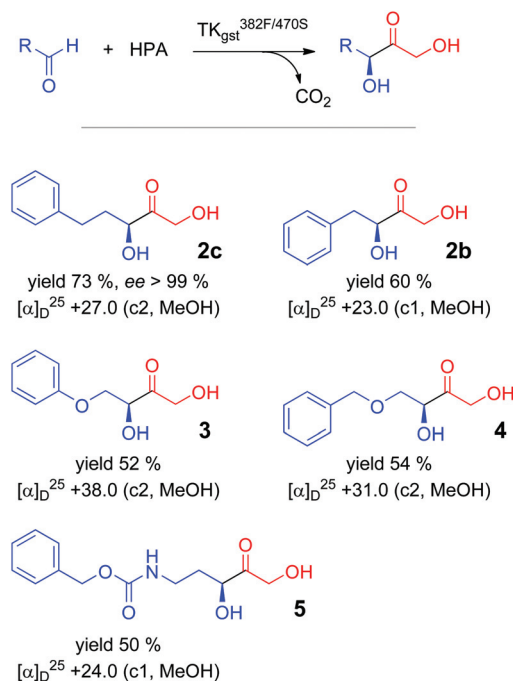


Fig. 4 Substrate scope of new TK<sub>gst</sub> variants for aryl-containing aldehydes. Substrates: ■ 2-phenylethanal, ■ 3-phenylpropanal, ■ phenoxyethanal, ■ benzyloxyethanal, ■ (*N*-Cbz)-3-aminopropanal. The assay solution (200 μL) contained TK<sub>gst</sub> variant (25 μg), Li-HPA (50 mM), aldehyde (10 mM), ThDP (2.4 mM), MgCl<sub>2</sub> (9.4 mM), phenol red (0.028 mM), DMSO (20% v/v) and triethanolamine buffer (2 mM, pH 7.5).



Scheme 2 Synthetic reactions performed with variant TK<sub>gst</sub> L382F/D470S.

actions not found for the aliphatic chains, or the extent of hydrate formation.

Preparative-scale experiments using best TK<sub>gst</sub> variant

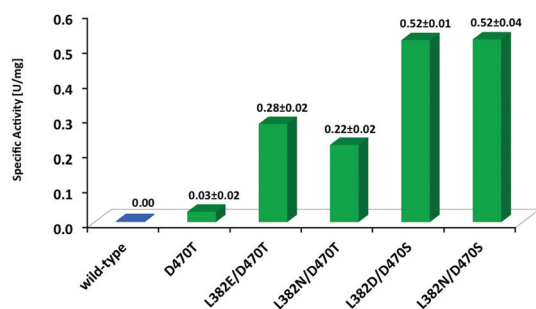
Preparative carboligation was carried out with all the tested aldehyde substrates and HPA as the nucleophile using the most effective (active and selective) variant L382F/D470S (Scheme 2). Reactions were conducted at 50 °C to profit from high overall reaction rates and the improved solubility of the arylated aldehyde substrates. The corresponding products were isolated with moderate to good yields (50–73%) and spectroscopically characterized to confirm their constitution. The



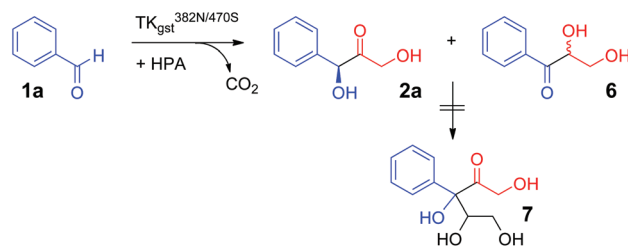
stereoselectivity was analysed by chiral HPLC, which gave high resolution for the racemic product from 3-phenylpropanal obtained by a chemical synthesis method<sup>27</sup> and showed an excellent enantiomeric purity (>99% ee) for the enzymatically produced material **2c**. In correlation with results obtained for the structurally related propanal adduct, it is assumed that the ketodiol must have the (3*S*) configuration. An earlier literature report had shown that the TK<sub>eco</sub> variant D469T also converts 2-phenylethanal into the corresponding (3*S*)-ketodiol with 96% ee.<sup>12</sup> Despite extensive experimental variation, the other products could not be directly resolved with available chiral HPLC columns, and derivatization (*e.g.*, mono- or dibenzoylation)<sup>12</sup> led to significant levels of racemized products.<sup>17</sup> However, we found that the isolated AB signal of the C1 protons, which is a characteristic of all the products, differentially shifts in the presence of Pirkle's alcohol, and thus allows for a sensitive determination of the enantiomeric purity. This method correlates well with the HPLC determination for **2c** and confirmed the optical purity (>98% ee) of the products **2b**, **3**, **4** and **5**, as was to be expected from the close structural relationship across the substrate series (Table S3; see ESI†).

### Identification of improved TK<sub>gst</sub> variants for benzaldehyde

Benzaldehyde (**1a**) is an attractive substrate for the synthesis of chiral building blocks but suffers from rather low electrophilicity relative to aliphatic aldehydes, and consequent lower reactivity in enzymatic carboligation reactions due to its resonance stabilization. On the other hand, aqueous solubility of **1a** is significantly higher than for the phenylalkanal studied (Fig. 2). Similar to TKs from other organisms,<sup>12</sup> wild-type TK<sub>gst</sub> showed very little activity against **1a** (Fig. 5), which clearly is of kinetic origin upon coupling to the practically irreversible decarboxylation of HPA. When the most stereoselective variant L382F/D470S showed only little improved reactivity with **1a** (<10-fold enhancement), the entire original hit library for propanal was screened again with **1a** in an attempt to find active clones that are better suited to adapt the increased bulkiness of this specific substrate type. The screening identified four



**Fig. 5** TK<sub>gst</sub> variants identified for improved activity against benzaldehyde (**1a**), in comparison to wild-type as a negative control and the TK<sub>gst</sub> D470T variant (matching the published TK<sub>eco</sub> D469T mutation) as a positive control. The assay solution (200  $\mu$ L) contained TK<sub>gst</sub> variant (20  $\mu$ g), Li-HPA (50 mM), **1a** (10 mM), ThDP (2.4 mM), MgCl<sub>2</sub> (9.4 mM), phenol red (0.028 mM), DMSO (20% v/v) and triethanolamine buffer (2 mM, pH 7.5).



**Scheme 3** Conversion of benzaldehyde performed with variant TK<sub>gst</sub> L382N/D470S.

unique TK variants L382E/D470T, L382N/D470T, L382N/D470S and L382D/D470S (Fig. 5).

Apparently, an exchange of the native negatively charged Asp470 with a polar but uncharged Ser or Thr residue seems to improve the productive binding of **1a** for the carboligation step. Our recent studies also revealed Ser or Thr replacements for Asp470 to be instrumental for activity and stereoselectivity of TK<sub>gst</sub> with non-hydroxylated substrates,<sup>17</sup> as well as with unnaturally (2*S*)-configured hydroxyaldehydes.<sup>18</sup> Similar observations were previously also reported for analogous variants of TK<sub>eco</sub> created at its corresponding Asp469 site, among which the D469T variant turned out to show low activity towards unsubstituted **1a** and related hetero-analogous compounds that was not observed with the wild-type.<sup>12</sup> Thus, it seems clear that at least one hydrogen bond donating residue at position 470 (in TK<sub>gst</sub>) is critical to maintain sufficiently high activity with substrates that do not carry any, or at least no (2*R*)-hydroxylation. In order to evaluate this hypothesis, we further compared the activity of the single-point variant TK<sub>gst</sub> D470T, which confirmed its activity with benzaldehyde, albeit at a rather low level like for the TK<sub>eco</sub> homologue (Fig. 5). On the other hand, this result also revealed that much of the higher rate acceleration by more than one order of magnitude observed for the double-site hits is caused by the accompanying replacement of the second residue, which hints to a strong cooperative effect. §

Quite to our surprise, all hits for **1a** include an exchange of the hydrophobic Leu382 by a polar or even anionic residue, which by its location should be positioned in direct contact with the large hydrophobic aromatic ring system. Although O-H/ $\pi$  bonding in a suitable substrate orientation may be involved, as well as new hydrogen bonding contacts to neighbouring residues, a more detailed interpretation is currently difficult in the absence of an experimental structure or reliable computational model for these mutants.

Stereoselectivity determination for the reaction with **1a** (Scheme 3) was investigated using the variant L382N/D470S, which displayed the highest overall rates. An accurate ee

§ Notably, the specific activity of *ca.* 0.5 U mg<sup>-1</sup> our best TK variants identified for benzaldehyde conversion compares quite favourably with the activity of 2.5 U mg<sup>-1</sup> of the so far best variant of the pyruvate decarboxylase from *Zymomonas mobilis* that recently had been engineered for optimized carboligation activity with benzaldehyde for (*R*)-PAC synthesis.<sup>6d</sup>



analysis<sup>28</sup> proved difficult, though, because of the instability of the chiral centre in **2a** under the slightly basic reaction conditions (pH 7.5) due to C–H acidity of the ketone and the benzylic activation. The tendency for *in situ* racemization *via* enolate formation may also be inferred from a facile carbonyl rearrangement to the resonance-stabilized benzylic position, which gave rise to the formation of the corresponding *rac*-dihydroxypropiophenone **6** upon extended reaction times, as has been noted elsewhere.<sup>12</sup> However, no side products from consecutive double carbonylation events (*i.e.*, **7**) were observed as had been reported for TK<sub>eco</sub> catalysts (Scheme 3).<sup>12</sup>

## Conclusions

In this study, we have identified various double-site variants of the TK<sub>gst</sub> with significant improvement of activity and stereoselectivity towards arylated aldehydes (**1a–c**), which expand the toolbox for asymmetric carbonylation towards an important class of new products.¶ It is appropriate to note that the reactions presented above lead to novel chiral products that hitherto were inaccessible by biocatalytic carbonylation activity. So far, these structures cannot be produced by using other ThDP-dependent enzymes and, because of their unique acyloin nature of the ketodiol moiety, they are no substrates for aldolases either.<sup>3,4,9</sup>

Our study confirms that the stable TK<sub>gst</sub> is an excellent catalytic scaffold for challenging reaction conditions such as the presence of organic co-solvents at elevated temperatures needed to support the conversion of non-physiological substrates having low aqueous solubility. The intrinsic stereoselectivity of TK<sub>gst</sub> and its engineered variants render them valuable candidates for elaborate synthetic processes that can be operated efficiently under mild conditions with low environmental footprint.

## Experimental

### Materials and methods

All chemicals were purchased from Sigma-Aldrich. Li-HPA,<sup>30</sup> 2-benzyloxyethanal,<sup>31</sup> 2-phenoxyethanal and (*N*-Cbz)-3-amino-propanal<sup>32</sup> were synthesized according to literature procedures. NMR spectra were recorded on Bruker AR-300 and DRX-500 spectrometers. IR spectra were recorded on a FT-IR spectrometer Paragon 1000. Column chromatography was performed on Merck 60 silica gel (0.063–0.200 mesh; Millipore); analytical thin layer chromatography was performed on Merck silica gel plates 60 GF254 with anisaldehyde stain for detection. Reaction pH for preparative synthesis was maintained using a Schott TitroLine alpha TL autotitrator. Chiral HPLC analysis was performed on a Shimadzu 20A system using a

Daicel Chiralpak IB column (4 × 250 mm) and *n*-hexane with 10% isopropanol was used as a mobile phase at a flow rate of 1.0 mL min<sup>-1</sup> with UV detection at 215 nm. Optical rotation was measured using a Perkin Elmer Model 141 polarimeter. Kinetic measurements were carried out on a SpectraMax 190 microplate reader at 560 nm.

### Library screening

**Transketolase library.** For the construction of the TK<sub>gst</sub> Leu382X/Asp470X double-site mutagenesis libraries, the QuikChange kit from Agilent (USA) was used. The complete protocol for the library creation was already described.<sup>17</sup>

**Preparation of cell-free lysate.** 2 µL of each clone from the subset library (198 clones; active hits from propanal screening) was transferred to individual wells in 96-well plates containing 150 µL per well LB growth media comprised with 50 µg mL<sup>-1</sup> kanamycin. BL21(DE3) host cells were used as a control. The plates were sealed with plastic lids to avoid evaporation, and then incubated (Heidolph Titramax 1000 microplate incubator) at 37 °C, 900 rpm overnight. After that, 20 µL was transferred from each well to 96-deepwell plates containing 400 µL per well of LB-kanamycin (50 µg mL<sup>-1</sup>) growth medium with 0.1 mM IPTG. The plates were covered with sealing foil, then incubated at 30 °C, 1200 rpm overnight. Cells were harvested by centrifugation at 4000g for 30 min. The culture medium was removed and the cell pellets stored at –80 °C. The cell pellets were suspended in 150 µL re-suspension buffer (1/10 BugBuster solution (Novagen, Germany)), 0.5 mg mL<sup>-1</sup> lysozyme (Roth, Germany) and 4 U mL<sup>-1</sup> Benzonase endonuclease (Novagen, Germany) and incubated in a shaker for 1 hour at room temperature. Then, the cell lysate was centrifuged at 4000g for 30 min.

**Mutant screening method.** A 40 µL aliquot of the supernatant, containing *ca.* 40 µg of TK protein, was transferred to a new 96-well plate. Substrate solution [140 µL; containing 9 mM MgCl<sub>2</sub>, 2.4 mM ThDP, 0.028 mM phenol red, 10 mM 3-phenylpropanal (20% DMSO), 2 mM TEA at pH 7.5] was added to each well. The reaction was initiated by the addition of 50 mM HPA as the lithium salt (Li-HPA). The OD increase was read by plate-reader at 560 nm. Positive colonies with greater than 8-fold activity from all plates were picked into a new plate for a second round of screening.

### Expression and purification of positive mutants

Each clone was raised in 2 L auto-induction medium (without preculture)<sup>33</sup> with addition of a few drops anti-foam suspension. After 20 h incubation at 200 rpm and 37 °C, the culture was harvested at 4000 rpm for 30 minutes. The collected cells were re-suspended in 100 mL of 20 mM phosphate buffer (pH 7.4), and the suspension was frozen at –20 °C. After thawing, lysozyme (1 mg mL<sup>-1</sup>) and DNase I (1 µL mL<sup>-1</sup> from stock of 1 mg mL<sup>-1</sup> in 50% glycerol (w/v); 20 mM Tris HCl, pH 7.5; 1 mM MgCl<sub>2</sub>) were added and the mixture stirred for 1 h at 37 °C. After addition of 100 mL 0.5 M NaCl, 20 mM phosphate buffer (pH 7.4) the suspension was centrifuged at 12 000g for 30 min. The supernatant was removed and 30 mM

¶ Interestingly, all hits active with **1a** (except for the L382F/D470S variant) were also found to be active with (*E*)-cinnamaldehyde,<sup>29</sup> as could be gathered by TLC monitoring of small analytical-scale experiments.



imidazole (5 M stock solution) was added. The Ni-NTA column (GE FastFlow resin with gravity flow) was equilibrated with 0.5 M NaCl, 20 mM phosphate buffer pH 7.4. Then the supernatant was applied and the column washed with 10 column volumes of buffer with 30 mM imidazole to remove unbound and unspecifically bound proteins. Protein was eluted with 10 column volumes of 0.5 M NaCl, 20 mM phosphate buffer containing 50 mM EDTA. The buffer in the eluted protein solution was exchanged by ultra-filtration against 2 mM TEA buffer (pH 7.0) and this solution was lyophilized. The purity of these samples was analyzed by SDS-PAGE and their protein concentration was measured by the Bradford method.

#### General procedure for preparative scale biocatalytic synthesis

ThDP (28 mg, 2.4 mM) and MgCl<sub>2</sub>·6H<sub>2</sub>O (48 mg, 9.4 mM) were dissolved in H<sub>2</sub>O (25 mL, total volume) and the pH adjusted to 7.5 using 0.1 M NaOH. To this, the lyophilized TK<sub>gst</sub> (L382F/D470S or L382N/D470S) enzyme (12 mg) was added and the mixture stirred for 30 min. After 30 min Li-HPA (152 mg, 50 mM) and 3-phenylpropanal (50 mM) with 5% DMSO as co-solvent were added and then stirring continued at 50 °C. The pH was automatically maintained throughout at 7.5 by addition of 0.1 M HCl using a pH stat. Reactions were monitored by TLC, and after 12 h the reaction mixture was concentrated under vacuum. The crude material was dry loaded on a silica column and purified using cyclohexane–ethyl acetate (1 : 1) as eluent.

**(R)-1,3-Dihydroxy-1-phenylpropan-2-one (2a).**<sup>12</sup> Yield 103 mg (50%); <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>): δ 7.41–7.32 (m, 5H), 5.24 (s, 1H), 4.34 (d, *J* = 19.4, 1H), 4.23 (d, *J* = 19.4 Hz, 1H); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>): δ 209.1, 137.5, 129.4, 129.3, 127.1, 77.9, 65.3.

**(S)-1,3-Dihydroxy-4-phenylbutan-2-one (2b).**<sup>12</sup> Yield 135 mg (60%); <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>): δ 7.34–7.31 (m, 2H), 7.29–7.26 (m, 1H), 7.22–7.19 (m, 2H), 4.50–4.46 (m, 1H), 4.41 (d, *J* = 19.5, 1H), 4.31 (d, *J* = 19.5 Hz, 1H), 3.12 (dd, *J* = 14 Hz & 4.5 Hz, 1H), 2.88 (dd, *J* = 14 Hz & 8 Hz, 1H); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>): δ 211.4, 135.8, 129.5, 129.0, 127.4, 76.1, 66.4, 40.6.

**(S)-1,3-Dihydroxy-5-phenylpentan-2-one (2c).** Yield 177 mg (73%); [α]<sub>D</sub><sup>25</sup> = +27.0 (*c* 2, CH<sub>3</sub>OH); <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>): δ 7.27–7.33 (m, 3H), 7.18–7.24 (m, 2H), 4.44 (d, *J* = 19.5, 1H), 4.35 (d, *J* = 19.5, 1H), 4.26 (dd, *J* = 8.7 & 3.9 Hz, 1H), 2.78 (t, *J* = 7.8, 2H), 2.03–2.14 (m, 1H), 1.82–1.95 (m, 1H); <sup>13</sup>C NMR (75 MHz, CDCl<sub>3</sub>): δ 211.8, 140.7, 128.7, 128.6, 126.4, 74.3, 65.6, 35.8, 31.1; EI-MS: *m/z* = 217.08357, calcd for C<sub>11</sub>H<sub>14</sub>O<sub>3</sub>[M + Na], 217.08352; HPLC: racemic product gave retention times of 29.9 (3*S*) and 32.4 (3*R*) at 1 mL min<sup>-1</sup> flow rate with *n*-hexane and iso-propanol (95 : 05). The enzymatic product was obtained with ee >99% (3*S* isomer).

**(S)-1,3-Dihydroxy-4-phenoxybutan-2-one (3).** Yield 127 mg (52%); [α]<sub>D</sub><sup>25</sup> = +38.0 (*c* 2, CH<sub>3</sub>OH); <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>): δ 7.31–7.28 (m, 2H), 7.02–6.99 (m, 1H), 6.91–6.88 (m, 2H), 4.67–4.52 (m, 3H), 4.28 (dd, *J* = 10 Hz & 4.5 Hz, 1H), 4.22 (dd, *J* = 10 Hz & 4.5 Hz, 1H); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>): δ 210.0, 157.9, 129.8, 122.1, 114.8, 74.4, 69.0, 66.9; ν<sub>max</sub>(neat)/cm<sup>-1</sup>

3441, 2925, 1730, 1242, 1081; ESI-MS: *m/z* calcd for C<sub>11</sub>H<sub>14</sub>O<sub>3</sub>[M + Na] 219.06, found 219.07.

**(S)-4-(Benzyloxy)-1,3-dihydroxybutan-2-one (4).** Yield 142 mg (54%); [α]<sub>D</sub><sup>25</sup> = +31.0 (*c* 2, CH<sub>3</sub>OH); <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>): δ 7.38–7.27 (m, 5H), 4.59–4.49 (m, 3H), 4.45–4.37 (m, 2H), 3.77 (dd, *J* = 10 Hz & 4.5 Hz, 1H), 3.69 (dd, *J* = 10 Hz & 4.5 Hz, 1H); <sup>13</sup>C NMR (75 MHz, CDCl<sub>3</sub>): δ 210.8, 137.1, 128.7, 128.3, 128.0, 75.0, 73.9, 70.9, 66.8; ν<sub>max</sub>(neat)/cm<sup>-1</sup> 3446, 2926, 1717, 1098; ESI-MS: *m/z* calcd for C<sub>11</sub>H<sub>14</sub>O<sub>3</sub>[M + Na] 233.08, found 233.09.

**(S)-(*N*-Cbz)-5-amino-1,3-dihydropentan-2-one (5).** Yield 158 mg (50%); [α]<sub>D</sub><sup>25</sup> = +24.0 (*c* 2, CH<sub>3</sub>OH); a complex NMR spectrum was recorded that showed two cyclic diastereomeric hemi-aminal compounds and their conformational equilibrium: <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>): δ 7.36–7.28 (m), 5.25 (t, *J* = 4.5 Hz) 5.13–5.05 (m), 4.49 (dd, *J* = 19.5 Hz & 9.5 Hz), 4.27 (dd, *J* = 9.5 Hz & 3.5 Hz), 4.18 (t, *J* = 4.5 Hz) 4.12–4.02 (m, 1H), 3.81–3.71 (m), 3.64–3.59 (m), 2.32–2.10 (m), 2.01–1.67 (m); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>): δ 212.6, 156.2, 136.3, 136.0, 128.7, 128.6, 128.6, 128.4, 128.3, 128.2, 128.1, 128.0, 88.0, 77.4, 73.1, 72.8, 67.6, 67.3, 66.8, 66.0, 64.4, 43.3, 40.2, 37.0, 36.1, 34.1, 29.3, 24.2; ν<sub>max</sub>(neat)/cm<sup>-1</sup> 3440, 2927, 2361, 1717, 1695, 1264, 1047; ESI-MS: *m/z* calcd for C<sub>11</sub>H<sub>14</sub>O<sub>3</sub>[M + Na] 290.10, found 290.10.

## Acknowledgements

This work was funded by the Deutsche Forschungsgemeinschaft (grant Fe244/9-1 to W. D. F.) and the Agence Nationale de la Recherche (grant ANR-09-BLAN-0424-CSD3 to L. H.) within the framework of Programme Blanc International, as well as by ESF project COST CM1303 *Systems Biocatalysis*.

## Notes and references

- (a) E. A. Meyer, R. K. Castellano and F. Diederich, *Angew. Chem., Int. Ed.*, 2003, **42**, 1210–1250; (b) E. Persch, O. Dumele and F. Diederich, *Angew. Chem., Int. Ed.*, 2015, **54**, 3290–3327.
- U. T. Bornscheuer, G. W. Huisman, R. J. Kazlauskas, S. Lutz, J. C. Moore and K. Robins, *Nature*, 2012, **485**, 185–194.
- W.-D. Fessner, in *Asymmetric Organic Synthesis with Enzymes*, ed. V. Gotor, I. Alfonso and E. Garcia-Urdiales, Wiley-VCH, Weinheim, 2008, pp. 275–318.
- (a) W.-D. Fessner and C. Walter, *Top. Curr. Chem.*, 1996, **184**, 97–194; (b) M. Brovetto, D. Gamenara, P. Saenz Mendez and G. A. Seoane, *Chem. Rev.*, 2011, **111**, 4346–4403.
- (a) A. J. Humphrey, N. J. Turner, R. McCague and S. J. C. Taylor, *J. Chem. Soc., Chem. Commun.*, 1995, 2475–2476; (b) E. Gonzalez-Garcia, V. Helaine, G. Klein, M. Schuermann, G. A. Sprenger, W.-D. Fessner and J.-L. Reymond, *Chem. – Eur. J.*, 2003, **9**, 893–899;



- (c) P. Hoyos, J.-V. Sinisterra, F. Molinari, A. R. Alcantara and P. Dominguez de Maria, *Acc. Chem. Res.*, 2009, **43**, 288–299; (d) M. Mueller, D. Gocke and M. Pohl, *FEBS J.*, 2009, **276**, 2894–2904; (e) M. Mueller, G. A. Sprenger and M. Pohl, *Curr. Opin. Chem. Biol.*, 2013, **17**, 261–270; (f) H. C. Hailes, D. Rother, M. Mueller, R. Westphal, J. M. Ward, J. Pleiss, C. Vogel and M. Pohl, *FEBS J.*, 2013, **280**, 6374–6394.
- 6 (a) V. B. Shukla and P. R. Kulkarni, *World J. Microbiol. Biotechnol.*, 2000, **16**, 499–506; (b) B. Rosche, M. Breuer, B. Hauer and P. L. Rogers, *J. Biotechnol.*, 2005, **115**, 91–99; (c) V. Sandford, M. Breuer, B. Hauer, P. Rogers and B. Rosche, *Biotechnol. Bioeng.*, 2005, **91**, 190–198; (d) D. Meyer, L. Walter, G. Kolter, M. Pohl, M. Muller and K. Tittmann, *J. Am. Chem. Soc.*, 2011, **133**, 3609–3616; (e) T. Sehl, H. C. Hailes, J. M. Ward, R. Wardenga, E. V. Lieres, H. Offermann, R. Westphal, M. Pohl and D. Rother, *Angew. Chem., Int. Ed.*, 2013, **52**, 6772–6775.
- 7 (a) G. Schenk, R. G. Duggleby and P. F. Nixon, *Int. J. Biochem. Cell Biol.*, 1998, **30**, 1297–1318; (b) R. A. W. Frank, F. J. Leeper and B. F. Luisi, *Cell. Mol. Life Sci.*, 2007, **64**, 892–905; (c) S. J. Costelloe, J. M. Ward and P. A. Dalby, *J. Mol. Evol.*, 2008, **66**, 36–49; (d) M. Widmann, R. Radloff and J. Pleiss, *BMC Biochem.*, 2010, **11**, 9.
- 8 P. Srere, J. R. Cooper, M. Tabachnik and E. Racker, *Arch. Biochem. Biophys.*, 1958, **74**, 295–305.
- 9 W.-D. Fessner and V. Helaine, *Curr. Opin. Biotechnol.*, 2001, **12**, 574–586.
- 10 (a) Y. Kobori, D. C. Myles and G. M. Whitesides, *J. Org. Chem.*, 1992, **57**, 5899–5907; (b) N. J. Turner, *Curr. Opin. Biotechnol.*, 2000, **11**, 527–531; (c) R. Wohlgemuth, *J. Mol. Catal. B: Enzym.*, 2009, **61**, 23–29.
- 11 D. Yi, T. Devamani, J. Abdoul-Zabar, F. Charmantray, V. Helaine, L. Hecquet and W.-D. Fessner, *ChemBioChem*, 2012, **13**, 2290–2300.
- 12 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.
- 13 G. R. Hobbs, M. D. Lilly, N. J. Turner, J. M. Ward, A. J. Willets and J. M. Woodley, *J. Chem. Soc., Perkin Trans. 1*, 1993, 165–166.
- 14 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.
- 15 J. Abdoul-Zabar, I. Sorel, V. Helaine, F. Charmantray, T. Devamani, D. Yi, V. de Berardinis, D. Louis, P. Marliere, W.-D. Fessner and L. Hecquet, *Adv. Synth. Catal.*, 2013, **355**, 116–128.
- 16 G. Ali, T. Moreau, C. Forano, C. Mousty, V. Prevot, F. Charmantray and L. Hecquet, *ChemCatChem*, 2015, **7**, 3163–3170.
- 17 D. Yi, T. Saravanan, T. Devamani, F. Charmantray, L. Hecquet and W.-D. Fessner, *Chem. Commun.*, 2015, **51**, 480–483.
- 18 J. Abdoul Zabar, M. Lorilliere, D. Yi, S. Thangavelu, T. Devamani, L. Nauton, F. Charmantray, V. Helaine, W.-D. Fessner and L. Hecquet, *Adv. Synth. Catal.*, 2015, **357**, 1715–1720.
- 19 (a) Y. Lindqvist, G. Schneider, U. Ermler and M. Sundström, *EMBO J.*, 1992, **11**, 2373–2379; (b) M. Nikkola, Y. Lindqvist and G. Schneider, *J. Mol. Biol.*, 1994, **238**, 387–404; (c) G. Schneider and Y. Lindqvist, *Biochim. Biophys. Acta*, 1998, **1385**, 387–398; (d) P. Asztalos, C. Parthier, R. Golbik, M. Kleinschmidt, G. Huebner, M. S. Weiss, R. Friedemann, G. Wille and K. Tittmann, *Biochemistry*, 2007, **46**, 12037–12052; (e) N. Maltseva, Y. Kim, K. Kwon, A. Joachimiak and W. F. Anderson, *RCSB protein data bank*, 2010, entry 3M49; (f) P. Neumann and K. Tittmann, *Curr. Opin. Struct. Biol.*, 2014, **29**, 122–133; (g) L. Nauton, V. Helaine, V. They and L. Hecquet, *Biochemistry*, 2016, **55**, 2144–2152.
- 20 (a) E. G. Hibbert, T. Senussi, M. E. B. Smith, S. J. Costelloe, J. M. Ward, H. C. Hailes and P. A. Dalby, *J. Biotechnol.*, 2008, **134**, 240–245; (b) M. E. B. Smith, E. G. Hibbert, A. B. Jones, P. A. Dalby and H. C. Hailes, *Adv. Synth. Catal.*, 2008, **350**, 2631–2638; (c) A. Cazares, J. L. Galman, L. G. Crago, M. E. B. Smith, J. Strafford, L. Rios-Solis, G. J. Lye, P. A. Dalby and H. C. Hailes, *Org. Biomol. Chem.*, 2010, **8**, 1301–1309.
- 21 (a) A. Ranoux, S. K. Karmee, J. Jin, A. Bhaduri, A. Caiazzo, I. W. C. E. Arends and U. Hanefeld, *ChemBioChem*, 2012, **13**, 1921–1931; (b) F. Subrizi, M. Cardenas-Fernandez, G. J. Lye, J. M. Ward, P. A. Dalby, T. D. Sheppard and H. C. Hailes, *Green Chem.*, 2016, **18**, 3158–3165.
- 22 (a) P. Payongsri, D. Steadman, J. Strafford, A. MacMurray, H. C. Hailes and P. A. Dalby, *Org. Biomol. Chem.*, 2012, **10**, 9021–9029; (b) P. Payongsri, D. Steadman, H. C. Hailes and P. A. Dalby, *Enzyme Microb. Technol.*, 2015, **71**, 45–52.
- 23 (a) W.-D. Fessner, *New Biotechnol.*, 2015, **32**, 658–664; (b) W.-D. Fessner and C. Walter, *Angew. Chem., Int. Ed. Engl.*, 1992, **31**, 614–616.
- 24 (a) G. Simon, T. Eljezi, B. Legeret, F. Charmantray, J. Castillo, C. Guerard-Helaine, M. Lemaire, M. Bouzon-Bloch, P. Marliere, V. Helaine and L. Hecquet, *ChemCatChem*, 2013, **5**, 784–795; (b) L. Hecquet, W.-D. Fessner, V. Helaine and F. Charmantray, in *Cascade Biocatalysis*, ed. S. Riva and W.-D. Fessner, Wiley-VCH, Weinheim, 2014, pp. 315–337.
- 25 U. Nilsson, L. Meshalkina, Y. Lindqvist and G. Schneider, *J. Biol. Chem.*, 1997, **272**, 1864–1869.
- 26 W. L. DeLano, *The PyMOL Molecular Graphics System, v1.3*, Schrödinger LLC, 2002, <http://www.pymol.org>.
- 27 (a) M. E. B. Smith, K. Smithies, T. Senussi, P. A. Dalby and H. C. Hailes, *Eur. J. Org. Chem.*, 2006, 1121–1123; (b) K. Rohr and R. Mahrwald, *Org. Lett.*, 2011, **13**, 1878–1880.
- 28 J. L. Galman and H. C. Hailes, *Tetrahedron: Asymmetry*, 2009, **20**, 1828–1831.
- 29 A. Cosp, C. Dresen, M. Pohl, L. Walter, C. Roehr and M. Mueller, *Adv. Synth. Catal.*, 2008, **350**, 759–771.
- 30 (a) F. Dickens and D. H. Williamson, *Biochem. J.*, 1958, **68**, 74–81; (b) A. W. Holldorf, in *Methods of Enzymatic Analysis*, (3rd ed.), ed. H. U. Bergmeyer, VCH Weinheim, 1984, vol VI, pp. 578–582.



- 31 R. K. Haynes, W. W.-L. Lam, L.-L. Yeung, I. D. Williams, A. C. Ridley, S. M. Starling, S. C. Vonwiller, T. W. Hambley and P. Lelandais, *J. Org. Chem.*, 1997, **62**, 4552–4553.
- 32 L. Espelt, T. Parella, J. Bujons, C. Solans, J. Joglar, A. Delgado and P. Clapes, *Chem. – Eur. J.*, 2003, **9**, 4887–4899.
- 33 F. W. Studier, *Protein Expression Purif.*, 2005, **41**, 207–234.

