Second Generation Engineering of Transketolase for Polar Aromatic Aldehyde Substrates

Panwajee Payongsri[a], David Steadman[b], Helen C. Hailes[b], Paul A. Dalby[a,*]

[a]Department of Biochemical Engineering, University College London, Gordon Street, London, WC1H 0AH, UK.

[b]Department of Chemistry, University College London, 20 Gordon Street, London WC1H 0AJ, UK.

[*]Corresponding author

Professor Paul A. Dalby

Department of Biochemical Engineering, University College London, Gordon Street, London, WC1H 0AH, UK

Email: p.dalby@ucl.ac.uk

Abstract

Transketolase has significant industrial potential for the asymmetric synthesis of carbon-carbon bonds with new chiral centres. Variants evolved on propanal were found previously with nascent activity on polar aromatic aldehydes 3-formylbenzoic acid (3-FBA), 4-formylbenzoic acid (4-FBA), and 3-hydroxybenzaldehyde (3-HBA), suggesting a potential novel route to analogues of chloramphenicol. Here we evolved improved transketolase activities towards aromatic aldehydes, by saturation mutagenesis of two active-site residues (R358 and S385), predicted to interact with the aromatic substituents. S385 variants selectively controlled the aromatic substrate preference, with up to 13-fold enhanced activities, and K_M values comparable to those of natural substrates with wild-type transketolase. S385E even completely removed the substrate inhibition for 3-FBA, observed in all previous variants. The mechanisms of catalytic improvement were both mutation type and substrate dependent. S385E improved 3-FBA activity via k_{cat}, but reduced 4-FBA activity via K_M. Conversely, S385Y/T improved 3-FBA activity via K_M and 4-FBA activity via k_{cat}. This suggested that both substrate proximity and active-site orientation are very sensitive to mutation. Comparison of all variant activities on each substrate indicated different binding modes for the three aromatic substrates, supported by computational docking. This highlights a potential divergence in the evolution of different substrate specificities, with implications for enzyme engineering.

Keywords

Biocatalysis, transketolase, enzyme engineering, benzaldehyde, directed evolution

Abbreviations:

3-FBA, 3-formylbenzoic acid; 3-HBA, 3-hydroxybenzaldehyde; 4-FBA, 4-formylbenzoic acid; HPA, hydroxypyruvic acid; TFA, trifluoroacetic acid; ThDP, thiamine diphosphate; TK, transketolase

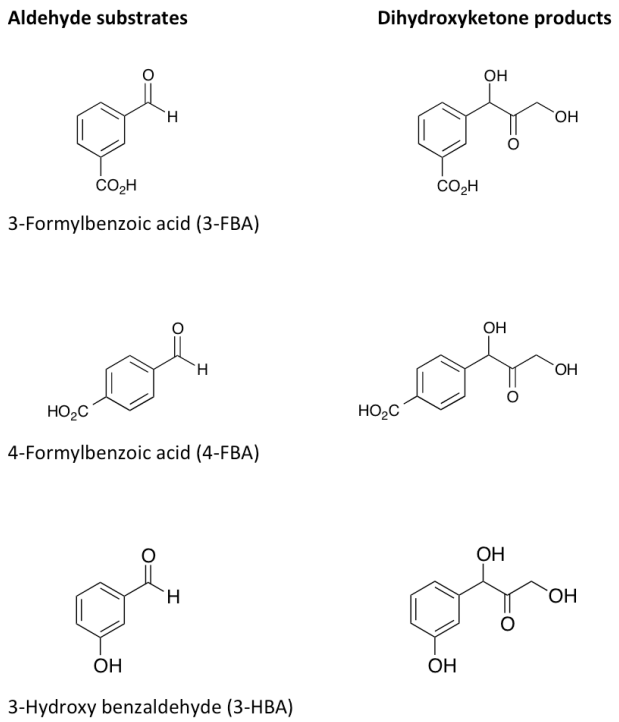
26 **1. Introduction**

27 Asymmetric carbon-carbon bond formation is a powerful method in organic synthesis [1-8]
28 and new routes could provide access to a wide range of novel and natural compounds that serve as
29 building blocks for further synthesis [2]. Enzymes such as transketolase (TK) (EC 2.2.1.1) have been
30 shown to catalyse asymmetric carbon-carbon bond formation with considerable synthetic potential
31 due to their high selectivity and specificity [1-4, 9-13]. TK belongs to the thiamine diphosphate
32 (ThDP) dependent enzyme family, and plays two crucial roles in the non-oxidative pentose
33 phosphate pathway and Calvin cycle [14]. It catalyses the reversible transfer of a C_2-hydroxyketone
34 group from a ketol donor such as ketose sugars, to an aldehyde [15-17]. TK-catalysed reactions
35 produce an α,α'-dihydroxy ketone group which is found in a wide range of natural compounds such
36 as carbohydrates and corticosteroids [18]. Hydroxypyruvate has been extensively used as a ketol
37 donor instead of the natural ketose sugar substrates, due to the production of carbon dioxide, which
38 can drive the reaction to completion. The use of TK could provide advantages over chemical
39 synthesis routes which tend to suffer from low yields due to multistep procedures [18] and racemic
40 products or low stereoselectivities when a single step method is applied [19, 20]. So far,
41 transketolases from several organisms have been exploited in new synthetic routes to carbohydrates
42 and sugar analogues [21-24], (+)-*exo*-brevicomin [25], N- hydroxypyrrolidine [26], furaneol [27],and
43 the glycosidase inhibitor 1,4-dideoxy-1,4-imino-D-arabinitol [28]. By coupling a transaminase (TAm)
44 step as a TK-TAm pathway, chiral amino-alcohols can also be synthesised enantioselectively [29].
45 Extending the acceptance of this pathway to include aromatic aldehydes would open up routes to
46 chiral aromatic amino-alcohols such as chloramphenicol antibiotics, norephedrine, and also their
47 analogues with alternative aromatic substituents.

48 Although wild-type transketolases can accept a wide range of aliphatic, aromatic, and cyclic
49 aldehydes [8, 17, 21, 23, 30], short chain aliphatic aldehydes tend to give faster reactions than more
50 sterically-challenging substrates [21, 23, 30]. In addition, most of the synthetic products previously
51 reported used aliphatic hydroxylated aldehydes as starting materials. Expansion of the substrate

range accepted by transketolases will increase their potential use for the synthesis of novel dihydroxy ketone compounds. Indeed, protein engineering of WT *E. coli* TK has been successfully used to improve its activity towards aliphatic aldehydes [31], with both enhanced and reversed stereospecificity [32-35], and also improved activity towards D-ribose and D-glucose [36]. By contrast, aromatic aldehydes still suffer from low reaction yields [33], except 3-formylbenzoic acid (3-FBA) and 4-formylbenzoic acid (4-FBA) which introduce a beneficial binding interaction between their carboxylic acid group and the transketolase phosphate-binding residues [37]. The *E. coli* TK variant D469T was found previously to have the highest activity towards the carboxylated aromatic substrates 3-FBA and 4-FBA, from all TK variants tested [37]. However, the best activity was still low compared to any natural substrates, and for 3-FBA, substrate inhibition was also always present.

Saturation mutagenesis at residues that interact directly with substrates have the greatest potential to improve enzyme activity and stereospecificity in a single round of evolution, compared to mutation of residues at more distant locations [38]. We have therefore applied a targeted saturation mutagenesis strategy to further improve the activity and yield towards the substituted benzaldehydes 3-FBA, 4-FBA and additionally, 3-hydroxybenzaldehyde (3-HBA) (Scheme 1).



Scheme 1. Aromatic aldehyde substrates and their respective products.

69 Previously, the variant D469T/R520Q was found to retain most of the activity of D469T,
70 towards 3-FBA and 4-FBA. The saturation mutagenesis was targeted into D469T/R520Q, as it is also
71 more readily able to accommodate further mutations due to stabilisation of the enzyme in terms of
72 protection from the formation of insoluble aggregates [39]. Residue S385 is within one of two
73 cofactor-binding loops (residues 383-393) present in each active-site of *E. coli* TK, and not structured
74 in the apo-enzyme [40]. The side chain hydroxyl group of S385 can form a hydrogen bond with the
75 phosphate moiety of natural substrates, as determined from several *E. coli* TK crystal structures with
76 substrates (2R5N, 2R8O, 2R8P) [41]. Residue R358 is also involved in binding the phosphate moiety
77 of natural substrates [41], as confirmed by selected mutagenesis of the yeast TK at the equivalent
78 site, such that R359A resulted in a 2-fold and 38-fold increase in K_M for xylulose 5-phosphate and
79 ribose 5-phosphate, respectively [42]. Previous mutations, R358P and R358L also indicated an
80 analogous interaction with the carboxylate moieties of 3-FBA and 4-FBA [37]. In our previous work,
81 saturation mutagenesis of S385 and R358, along with 18 other sites within the wild-type enzyme,
82 aimed to improve activity towards non-aromatic aldehydes, glycolaldehyde and propanal [31, 43].
83 While several other sites gave significantly improved activities, S385 gave no variants of interest, and
84 R358 yielded only R358I with a 2-fold improvement of activity towards each substrate. Here, S385
85 and R358 have again been mutated to all possible amino acids, but within D469T/R520Q instead of
86 wild-type TK, resulting in improved activities towards substituted aromatic aldehydes.
87
88

89 **2. Materials and Methods**

90 **2.1. Chemicals and reagents**

91 All chemical reagents were purchased from Sigma-Aldrich (Aldrich® Chemistry, UK),
92 otherwise stated. Lithium hydroxypyruvate was prepared according to the previous protocol [10].

93

94 **2.2. Transketolase library**

95 Mutations were constructed on D469T/R520Q due to its stability [39], and initial activity on
96 the aromatic aldehydes [37]. The transketolase D469T/R520Q variant gene and all further variants
97 were expressed under the control of *tktA* gene promoter in the plasmid pQR791 in XL-10 Gold
98 (Stratagene) [40]. Target residues for saturation mutagenesis were from those determined to be in
99 the first shell [43]. Mutagenesis was targeted only to those involved in binding phosphate groups
100 within natural substrates, as determined by previous mutagenesis and crystal structure analysis [41].
101 These included S385, R358, H461 and R520. R520 was already mutated to R520Q in the parent
102 variant D469T/R520Q, and so not mutated further. H461 was also excluded from further mutation as
103 it was previously found to lead to stability issues [37,39]. Mutagenesis was carried out using
104 Quikchange site directed mutagenesis (Stratagene). The *dpnI*-digested PCR product was transformed
105 into XL-10 gold competent cells (Stratagene). The quality and diversity of the PCR was checked by
106 DNA sequencing prior to screening. The primers used for additional variant construction were listed
107 below (codon underlined, mutations in bold, and N is an equal mix of all four bases, H is 33% A,
108 33% C and 33% T, and S is 50% G and 50% C).

109 R358X GAAAATCGCCAGC**NHN**AAAGCGTCTCAGAATG

110 S385X GCTGACCTGGCGCCG**NHS**AACCTGACCCTGTGG

111 The NHN and NHS codons both exclude Cys, Trp, Arg and Gly amino-acids.

112

113 **2.3. Library Screening**

114 After the transformation, 24 colonies for each library were inoculated into 12 ml LB media
115 with 150 µg/ml ampicillin, incubated for 18 hours at 37 °C, shaking at 250 rpm and then 2 ml of the
116 culture was centrifuged at 13000 rpm for 10 minutes. The probability of coverage and possible
117 number of amino acid replacements was calculated using the online tool: GLUE-IT [44] (currently
118 maintained at http://guinevere.otago.ac.nz/cgi-bin/aef/glue-IT.pl). From each codon mix used, the

119 two libraries could produce 16 possible variants and so 11.6 distinct amino acids were expected
120 within 24 randomly picked colonies. The diversity of mutations was confirmed by DNA sequencing of
121 at least 6 random colonies from each library.

122 Cell pellets were resuspended in 200 µl of 2x cofactor solution (4.8 mM ThDP and 18 mM
123 $MgCl_2$ in 50 mM Tris.HCl, pH 7.0), incubated for 20 minutes, and 150 µl then transferred to a
124 borosilicate microplate (Radleys, Essex, UK). Screens were performed on whole cells to avoid the
125 influence of uneven cell disruption on TK release and the bioconversion rate. Final yields of products
126 using whole cells and clarified lysates, were found to correlate well (supplementary information Figs.
127 S1 & S2). Reactions were started by the addition of 150 µl of 2x substrate solution (100 mM 3-FBA
128 and 100 mM HPA, 80 mM 4-FBA and 80 mM HPA, or 30 mM 3-HBA and 60 mM HPA, each in 50 mM
129 Tris buffer, pH 7.0). All reactions were performed at 22 °C with shaking at 200 rpm, 3 mm amplitude,
130 to avoid cell sedimentation, and sealed to prevent evaporation (Thermo Scientific Nunc). Conversion
131 to product was determined in 3-FBA samples taken at 1 hour and 24 hours after the reaction was
132 started. Conversion at 1 hour provided a comparison of activities, whereas conversion at 24 hours
133 provided final product yields. For the slower reactions with 4-FBA and 3-HBA, the libraries were
134 screened for conversion after 18 hours to identify variants with improved activities.

135 At each time-point, 20 µl of the reaction sample was added to 380 µl 0.1% TFA, centrifuged
136 at 13000 rpm for 3 mins, and supernatants analysed by HPLC with an ACE5 C18 reverse phase
137 column (150x4.6 mm). The 3-FBA and 4-FBA reaction samples were analysed as previously described
138 [37] using two mobile phases 0.1% TFA and 100% acetonitrile, and a flow rate of 1 ml/min (see SI for
139 further details). The 3-HBA reaction samples were analysed by gradient elution as follows. Using a
140 mobile phase of 0.2M acetic acid and 80% methanol at the flow rate of 1ml/min, the flow profile was
141 separated into 3 phases: 1st phase 5 mins using 90% 0.2 M acetic acid/10% methanol (80%v/v); 2nd
142 phase ramped linearly to 40% 0.2 M acetic acid/60% methanol (80%v/v) over 14 mins, and the last
143 phase maintaining 40% 0.2 M acetic acid/60% methanol (80%v/v) for 6 mins. The column was then

144 re-equilibrated for 3 mins using 90% 0.2 M acetic acid/10% methanol (80%v/v). The retention time
145 of the product from the 3-HBA reaction was 4.7 min and that for 3-HBA was 14.5 min.
146
147 **2.4. Cell culture and protein quantification for detailed enzyme kinetics**
148 Glycerol stocks of selected variants were re-streaked on 150 µg/ml ampicillin-LB plates and
149 incubated at 37 °C for 18 hours. Single colonies were inoculated into 20 ml LB with 150 µg/ml
150 ampicillin in 250 ml shake flasks, incubated at 37 °C with shaking at 250 rpm for 18 hours, then
151 harvested by centrifugation. Supernatant-free cell pellets were resuspended in Tris buffer, pH 7.0
152 and sonicated on ice (MSE Soniprep 150 probe, Sanyo) with 10s on, 15s off for 10 cycles. Cell debris
153 was removed by centrifugation at 17,700 g for 10 minutes at 4 °C. The clarified lysate was aliquotted
154 and stored at -80 °C. Total protein concentration in the clarified lysate was determined using the
155 Bradford assay with BSA as a standard protein. The lysate was further analysed by SDS-PAGE and
156 densitometry as previously described [43] to determine the TK concentration which was always
157 over-expressed to above 20% of the total protein. The plasmids of these variants were extracted and
158 sequenced to identify the mutation. The whole procedure was performed in triplicate for each
159 variant.
160
161 **2.5. Detailed enzyme kinetics**
162 Detailed enzyme kinetic analyses were carried out on clarified lysates, as this is the form
163 most relevant to industrial biocatalytic processes, either as an isolated enzyme or within a *de novo*
164 pathway [45]. TK enzyme purification has also previously led to instabilities that would affect the
165 observed kinetics [39]. Clarified lysates of the variants with high, moderate, and low enzyme
166 activities or conversion yields, prepared as above, were used to determine their specific activities or
167 yield towards each substrate. Variants with high activities towards at least one substrate were
168 further studied to determine the K_M and k_{cat} for all the substrates. Kinetic parameters were obtained
169 at saturating 50 mM HPA levels and the aldehyde concentrations varied (3-FBA 3 to 90 mM, 4-FBA 6

170 to 30 mM, 3-HBA 5 to 25 mM). Substrate solutions were prepared at 3x in 50 mM Tris.HCl, pH 7.0.
171 To 30 µl of clarified lysate, 30 µl of 10x cofactor solution (24 mM ThDP, 90 mM $MgCl_2$) was added
172 with 140 µl 50 mM Tris.HCl, pH 7.0, and incubated for 20 mins. Reactions were initiated by the
173 addition of 100 µl of substrate solution. All reactions were carried out in triplicate in glass vials at
174 22 °C. To 20 µl of reaction samples, 380 µl of 0.1% TFA was added and the mixture centrifuged at
175 13,000 rpm for 3 mins. The supernatants were analysed by HPLC as above. 3-FBA reaction samples
176 were taken at every 3 mins for 15 mins. The reaction samples of 4-FBA were taken at 15 minute
177 intervals for 90 mins. 3-HBA reaction samples were taken every 30 mins for 180 mins. Samples at
178 18 hours for the highest aldehyde concentrations were also taken to quantify the reaction yield
179 determined from product concentrations. Evaporation was not observed at any point, and control
180 samples without enzyme, or with lysates from non-TK expressing *E. coli* strains retained the initial
181 substrate concentrations throughout. The TK concentration in each reaction was between 0.07 and
182 0.3 mg/ml. Higher TK concentrations were used for the 3-HBA reaction due to the slower conversion.
183 All data were fitted by non-linear regression to the Michaelis-Menten equation to determine the K_M
184 and k_{cat} of all the variants with each substrate. Double-reciprocal Lineweaver-Burk plots were also
185 used to verify that the relationships between the velocity and the concentration were linear. The
186 racemic products from 3-HBA [33], 3-FBA and 4-FBA [37] were synthesised, purified, and
187 characterised as previously, as standards for HPLC calibration. None of the aromatic product
188 enantiomers could be directly resolved by chiral-HPLC and so derivatisation was attempted using
189 several methods. Only 3-FBA could be derivatised with any success by dibenzoylation, but a
190 significant degree of rearrangement via enolisation was noted, suggesting racemisation of the
191 product occurred. This means that all e.e. values were probably underestimated.
192

193 **2.6. Computational docking of 3-FBA, 4-FBA and 3-HBA in transketolase variant active-sites**

194 The structures of TK variants, aldehydes and ThDP enamine were modelled as previously
195 described [33] (see SI for further details) into TK variant structures prepared by replacing the target

196 residues of the *E. coli* wild type (PDB ID: 1QGD) [46]. We define "productive location/proximity" in
197 all cases simply as placing the aldehyde carbonyl and the enamine-cofactor intermediate at within
198 ca. 5 Å, provided that further rotation of the carbonyl is possible to achieve a suitable pi-stacking
199 trajectory for reaction. This simplification reflects the general expectations of using such a modelling
200 method.

201

202 **3. Results and discussion**

203

204 **3.1. Library screening**

205 **3.1.1. Screening of the S385X/D469T/R520Q library**

206 Saturation mutagenesis of S385 in wild-type *E. coli* TK was found previously to give no
207 variants with improved activity towards the aliphatic substrates glycolaldehyde or propanal [31, 43].
208 By contrast, the present screen of the S385X/D469T/R520Q library gave at least five unique variants
209 with markedly improved (2-5.5 fold) conversion of 3-HBA after 18 hours (relative to D469T), and two
210 variants with 2-2.5 fold improved conversion of 4-FBA after 18 hours. One variant with only
211 modestly improved (1.8 fold) conversion of 3-FBA after 1 hour (Fig. 1) was also obtained. Most other
212 variants lowered or maintained the conversion of 3-FBA after 1 hour. Strikingly, the conversions at
213 18 hours from 3-HBA were highly correlated with those from 4-FBA, and to a lesser degree with
214 conversions at 1 hour for 3-FBA. This suggested the same mechanism by which S385 mutants
215 improved the 3-HBA and 4-FBA reactions. Improvements for 3-FBA appear to arise through different
216 mechanisms, including substrate inhibition relief as discussed further below for at least one mutant.

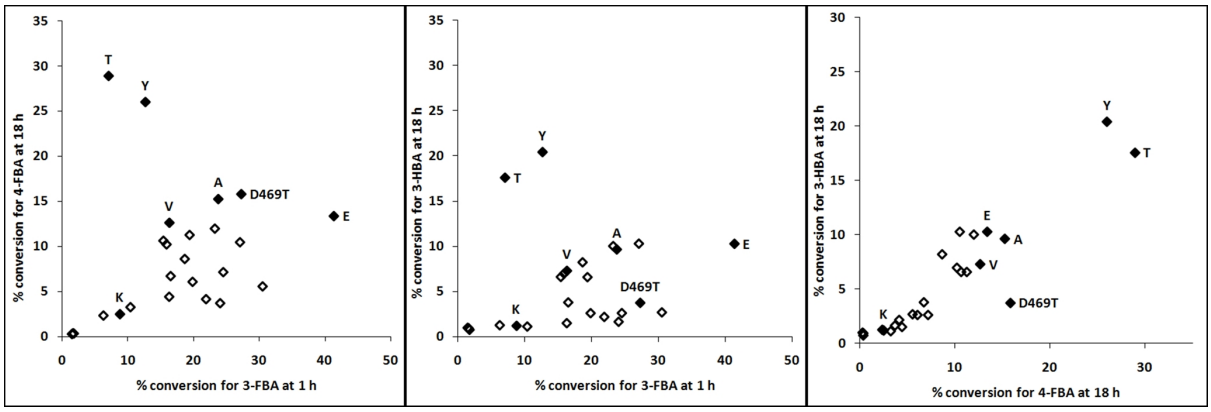


Fig. 1. Pairwise comparisons of activities for the S385X/D469T/R520Q library. Activity on 3-FBA was measured from the conversion after 1 hour, where the reactions were less than 35% complete. Conversion at 18 hours gave an indirect measure of activity with slower 4-FBA and 3-HBA reactions. All reactions were performed using freshly harvested whole cells on 50 mM 3-FBA and 50 mM HPA, 40 mM 4-FBA and 40 mM HPA, or 15 mM 3-HBA and 30 mM HPA, each with 2.4 mM ThDP and 9 mM $MgCl_2$ in 50 mM Tris.HCl, pH 7.0, at 22 °C with shaking at 200 rpm. Letters denote amino-acid substitutions observed at X in S385X/D469T/R520Q. D469T refers to the parent variant D469T/R520Q.

Specific activities of selected TK variants in clarified lysates were obtained for a more detailed and accurate comparison. D469T had a specific activity on 3-HBA of only 0.0071 µmol/mg/min, and a conversion to product of only 7% after 18 hours (Table 1). By contrast, the S385V/D469T/R520Q, S385Y/D469T/R520Q and S385T/D469T/R520Q variants achieved 15%, 25%, and 31% conversion, respectively. Their specific activities also increased proportionally, between 6- and 13-fold, to 0.03, 0.071 and 0.073 µmol/mg/min, respectively (see supplementary information, Fig. S3). The same correlation between specific activity and conversion after 18 hours was also observed for 4-FBA (Table 1). This correlation has been previously observed to result from competition against HPA substrate degradation for the slower reactions of 4-FBA and 3-HBA [37], rather than due to an incomplete reaction at 18 hours, or enzyme instability over the 18-hour reaction period. Indeed, enzyme activity typic ally decreased by less than 50% over 18 hours in the presence of aldehyde (see Fig. S4). For 3-FBA, the only variant with a higher (1.8-fold) activity in the whole-cell screen was S385E/D469T/R520Q. Using clarified lysate, this variant gave a 3.5-fold increase in specific activity with 3-FBA relative to that of D469T. The final conversion obtained with

239 3-FBA for all variants, which did not increase beyond 18 hours, was between 53% and 67% indicating
240 a persistent product inhibition.

241 Most of the beneficial mutations (except S385E) introduced hydrophobic (A, V, Y) or polar (T)
242 side chains, implicating modified hydrophobic interactions with each of the aromatic substrates,
243 rather than affecting any specific interactions to the hydroxyl group of 3-HBA or the carboxylate
244 moieties of 4-FBA or 3-FBA. This was slightly unexpected given that S385 interacts with the
245 phosphate group in natural sugar substrates. S385 is in one of the two cofactor-binding loops
246 present within each active site. While the B-factor of this residue and the neighbouring loop
247 residues are not significantly higher than the average B-factor for the structure (from PDB ID: 1QGD)
248 [46], an impact on K_M or even k_{cat} due to alteration of loop mobility upon mutation of S385 cannot
249 be ruled out.

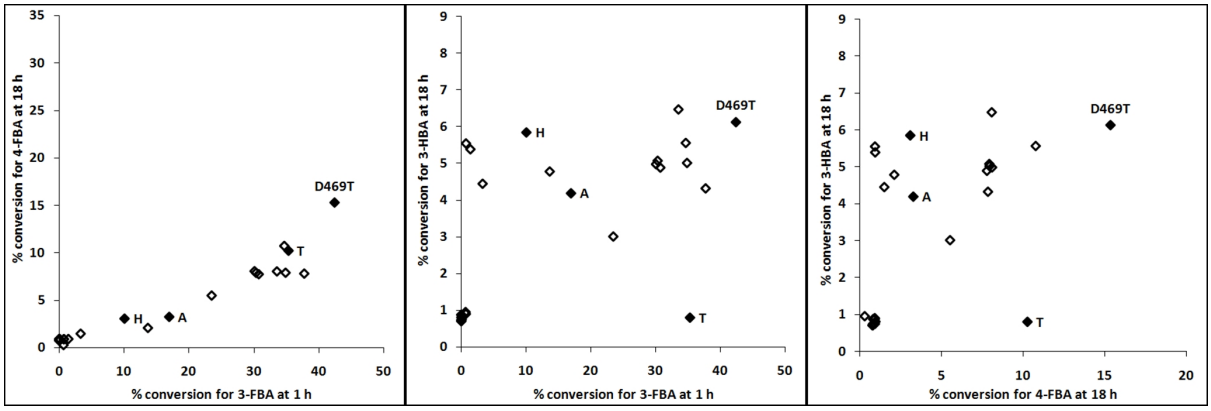
250

251 **3.1.2. Screening of the R358X/D469T/R520Q library**

252 Saturation mutagenesis of R358 in wild-type *E. coli* TK previously led to only one beneficial
253 variant R358I, which gave only a 2-fold improvement in the activity towards the aliphatic aldehydes
254 glycolaldehyde and propanal [31, 43]. Some triple variants of R358X/D469T/R520Q, constructed and
255 tested individually (data not shown), gave slightly reduced (R358H, and R358I) or severely
256 diminished (R358E, R358Q) activities towards 3-FBA. The R358X/D469T/R520Q library provided no
257 variants with improved activity towards any of the three aromatic substrates (Fig. 2). This is not
258 surprising for 3-FBA and 4-FBA, as we had previously identified charge stabilising interactions
259 between their carboxylate moieties and R358 [37]. The initial library screen revealed that the
260 conversions of 3-FBA and 4-FBA for all variants were highly correlated, and that most gave less than
261 50% of the activity observed for D469T towards 3-FBA and 4-FBA. This is consistent with a common
262 mechanism, in which each mutant affects similarly, the stabilising electrostatic interactions between
263 R358 and 3-FBA and 4-FBA. By contrast, half of the variants retained an activity towards 3-HBA
264 similar to that obtained with D469T. The conversions of 3-FBA and 4-FBA for all variants were also

265 not well correlated to those of 3-HBA. This suggested that removing the positively charged R358
266 side-chain was not as detrimental for activity towards 3-HBA as it was for 3-FBA or 4-FBA. However,
267 the lack of any improved R358 variants towards 3-HBA was still surprising, as new favourable
268 interactions with the uncharged hydroxyl moiety of 3-HBA might be expected.

269 A few variants, R358T in particular, maintained at least 60% of the activity to 3-FBA and 4-
270 FBA. By contrast, R358T significantly reduced the activity towards 3-HBA, which was particularly
271 unexpected given the potential to form a new hydrogen bond to this substrate. Conversely, 3-HBA
272 tolerated a histidine residue, whereas the negatively charged 4-FBA and 3-FBA did not. This suggests
273 that the hydroxyl moiety of bound 3-HBA might not be oriented towards R358 in the same way as
274 the carboxylate moieties of 3-FBA and 4-FBA.

275 

276 **Fig. 2. Pairwise comparisons of activities for the R358X/D469T/R520Q library.** Activity on 3-FBA was measured from the
277 conversion after 1 hour, where the reactions were less than 35% complete. Conversion at 18 hours gave an indirect
278 measure of activity with slower 4-FBA and 3-HBA reactions. All reactions were performed using freshly harvested whole
279 cells on 50 mM 3-FBA and 50 mM HPA, 40 mM 4-FBA and 40 mM HPA, or 15 mM 3-HBA and 30 mM HPA, each with
280 2.4 mM ThDP and 9 mM $MgCl_2$ in 50 mM Tris.HCl, pH 7.0, at 22 °C with shaking at 200 rpm. Letters denote amino-acid
281 substitutions observed at X in R358X/D469T/R520Q. D469T refers to the parent variant D469T/R520Q.

282

283 Overall, the S385X library appeared to be more tolerant to different mutations than the R358X
284 library when screening against 3-FBA and 4-FBA. However, the R358 residue was tolerant to most
285 mutations with the activity on 3-HBA. It appeared then that the interactions between natural TK
286 substrates and residue S385 were no longer critical or present with the acidic aromatic-aldehyde

287 substrates 3-FBA and 4-FBA, whereas the charge-charge interactions with R358, as determined
288 previously [37], were still important to retain.
289
290 **3.2. Kinetic parameters of the variants with 3-FBA, 4-FBA, and 3-HBA**
291 Any variant with improved activity towards at least one substrate was assessed in more
292 detail with all three substrates to understand whether any shifts in specificity were due to the loss of
293 affinity or increased catalytic efficiency (k_{cat}). All kinetic parameters are summarised in Table 1.
294 S385E/D469T/R520Q gave a 3.5-fold increase in specific activity with 3-FBA relative to that of D469T,
295 and it was found to significantly increase the k_{cat} to 34.3± 1.6 s^{-1}, which was 2.6 times greater than
296 that for D469T. Furthermore, the K_M decreased to 18.3 ± 2.6 mM, which was 3-fold lower than that
297 for D469T, although most of this substrate binding effect was already achieved in with the R520Q
298 mutation in D469T/R520Q. Therefore, most of the beneficial effect of the S385E mutation was due
299 to an increase in k_{cat}. The K_M achieved was also lower than the 55 mM previously determined for
300 D469T with propanal [31], and lower than the 24.5 mM K_M for the well-characterised glycolaldehyde
301 substrate and wild-type *E. coli* TK [37]. The k_{cat}/K_M of S385E/D469T/R520Q was increased to
302 1900±270 s^{-1} M^{-1}, a 4-fold improvement over D469T/R520Q, and a 9-fold increase over D469T.
303 Surprisingly, the substrate inhibition observed previously in all TK variants with 3-FBA [37], was also
304 completely removed. No inhibition was observed up to the maximum concentration tested of
305 90 mM 3-FBA.
306 The specific activity of S385E/D469T/R520Q towards 3-HBA was also improved 6-fold
307 relative to D469T, although a K_M of 245 mM and k_{cat} of 0.59 s^{-1} indicated that this activity was still
308 relatively modest. The low activities made it too difficult to obtain kinetic parameters for D469T and
309 D469T/R520Q. Interestingly, the same variant lost 80% of the specific activity towards 4-FBA relative
310 to that of D469T. In fact, the K_M (72 ± 30 mM) for 4-FBA improved 3.5-fold relative to that with
311 D469T, but increased 3-fold relative to that with D469T/R520Q. However, the k_{cat} was 16-fold lower
312 than for D469T, and only marginally improved relative to D469T/R520Q.

313 It is interesting that S385E significantly increases the k_{cat} for 3-FBA, but has no effect on the
314 k_{cat} for 4-FBA. This suggests that k_{cat} for aromatic aldehyde substrates is dependent on the aromatic
315 substitution type and position, and is not a general catalytic effect of the enzyme for all substrates.
316 Indeed k_{cat} can be affected by substrate orientation and proximity to the cofactor and catalytic
317 residues. The data suggest that while the carboxylate moieties of 3-FBA and 4-FBA both interact with
318 R358, the reactive aldehyde moieties are oriented differently in the active site due to the different
319 spatial constraints of the two molecules, giving rise to their respective k_{cat} values. S385 mutations
320 appear to influence this molecular orientation and hence k_{cat}. While electronic effects on the
321 aldehyde moiety, due to different positions of the aromatic substitutions in 3-FBA and 4-FBA, may
322 also potentially play a role, this appears to be a minor factor as 4-FBA would be expected to have
323 greater reactivity than 3-FBA based on electronic effects alone.
324 The highest values of k_{cat}/K_M were found for S385Y/D469T/R520Q and S385T/D469T/R520Q
325 towards 3-FBA (Table 1). Most of this improvement was due to an almost 10-fold decrease in K_M
326 relative to D469T/R520Q, to the lowest value for all variant-substrate combinations of just 1.3 mM
327 3-FBA. Notably, this K_M value is comparable to those of the natural TK substrates, which range from
328 2.1 mM on D-glyceraldehyde-3-phosphate, down to 0.008 mM on D-Ribose-5-phosphate [14].
329 However, the activities of the S385Y/D469T/R520Q and S385T/D469T/R520Q variants towards 3-FBA
330 were significantly lower than most of the other variants, and only around 50% of D469T. This was
331 due to significant inhibition of both variants by 3-FBA at concentrations above 20 mM, which is in
332 marked contrast to the removal of substrate inhibition by the S385E mutation.
333 S385Y/D469T/R520Q and S385T/D46T/R520Q were also the only two variants to give at
334 least 40% greater conversion of 4-FBA during screening, than for D469T. Previously it was observed
335 that the R520Q mutation in D469T/R520Q decreased the K_M 10-fold, but unfortunately also
336 decreased the k_{cat} 25-fold with 4-FBA [37]. The two triple variants S385Y/D469T/R520Q and
337 S385T/D469T/R520Q retained similar affinities towards 4-FBA as for D469T/R520Q, but their k_{cat}
338 values were considerably improved. The restored k_{cat} and good substrate affinities of these two

339 variants resulted in a dramatic increase in their activities towards 4-FBA, a 4.5-5 fold improvement in
340 k_{cat}/K_M relative to that of D469T, and over 10-fold improvement in k_{cat}/K_M compared to
341 D469T/R520Q. The latter is attributable to the effect of the S385Y and S385T mutations upon k_{cat}
342 alone.

343 S385Y/D469T/R520Q and S385T/D469T/R520Q were also found to have greater activities
344 towards 3-HBA compared to D469T. The k_{cat} and K_M values for 3-HBA could not be readily obtained
345 for D469T or D469T/R520Q due to low activities. The k_{cat} values obtained for S385Y/D469T/R520Q
346 and S385T/D469T/R520Q with 3-HBA were 2.1 and 1.0 s^{-1} respectively, which were similar to those
347 for 4-FBA but 3 to 4 fold lower than those for 3-FBA. Furthermore, their K_M values were 100-300
348 times higher than those for 3-FBA, and 8-25 fold higher than for 4-FBA. Therefore, their k_{cat}/K_M
349 values were up to 1000 times lower than for S385Y/D469T/R520Q with 3-FBA.

350 These data support our previous hypothesis that enzyme-substrate affinity and positioning
351 of the substrate are competing major factors to be overcome when attempting to improve the
352 bioconversion rate of poorly accepted aromatic aldehydes in transketolase [37]. The specific
353 activities of all variants towards 3-HBA are significantly lower than with 3-FBA or 4-FBA, as their K_M
354 values are all still much higher than the substrate concentration. However, different k_{cat} effects
355 observed suggest positioning of the substrate relative to catalytic groups can have a significant
356 impact. This underlies the different shifts in substrate preference for the S385E mutation within
357 D469T/R520Q compared to those of S385Y and S385T. S385E mostly improved the k_{cat} towards 3-
358 FBA, whereas for 4-FBA the greatest impact was an increase in K_M. S385Y and S385T both improved
359 the K_M for 3-FBA, and yet increased the k_{cat} towards 4-FBA. This suggests that the position of the
360 substrate relative to the catalytic groups is very sensitive to mutation, and that while 3-FBA and
361 4-FBA are have favourable electrostatic interactions to R358, they orient the aldehyde moiety
362 differently towards the active-site cofactor. This presumably reflects the different relative positions
363 of their constrained aldehyde and carboxylate moieties.
364

	Substrate	Specific activity (relative to D469T)	Conversion at 18hrs (%)	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M (s^{-1} M^{-1})
D469T	3-FBA	1.0[a]	67 (1.0)	56 (10)[b]	13.2 (1.5)	240 (50)
D469T/R520Q	3-FBA	0.8	67 (0.1)	13 (4)[b]	6.0 (0.75)	470 (170)
S385E/D469T/R520Q	3-FBA	3.5	63 (1.0)	18(3)	34 (1.6)	1900 (270)
S385Y/D469T/R520Q	3-FBA	0.5	53 (0.8)	1.3 (0.4)	7.0 (0.3)	5400 (1490)
S385T/D469T/R520Q	3-FBA	0.5	59 (0.5)	1.7 (0.25)	7.3 (0.2)	4250 (620)
D469T	4-FBA	1.0[a]	30 (1.7)	251 (240)[b]	5.0 (4.4)	20 (26)
D469T/R520Q	4-FBA	0.5	13 (0.4)	25 (8)[b]	0.20 (0.03)	8.1 (2.8)
S385E/D469T/R520Q	4-FBA	0.2	<5	72 (30)	0.3 (0.1)	4.6 (2.4)
S385Y/D469T/R520Q	4-FBA	3.4	48 (0.4)	15 (3)	1.7 (0.1)	110 (21)
S385T/D469T/R520Q	4-FBA	2.1	48 (1.7)	23 (6)	2.1 (0.3)	90 (26)
D469T	3-HBA	1.0[a]	7	n.d.[c]	n.d.	n.d.
D469T/R520Q	3-HBA	n.d.	6	n.d.	n.d.	n.d.
S385E/D469T/R520Q	3-HBA	6.0	15 (1)	245 (15)	0.6 (0.1)	2.4 (0.2)
S385Y/D469T/R520Q	3-HBA	12.4	25 (1)	390 (10)	2.1 (0.2)	5.4 (0.1)
S385T/D469T/R520Q	3-HBA	12.7	31 (2)	180 (10)	1.0 (0.1)	5.5 (0.2)

365 **Table 1.The kinetic parameters of all the variants towards 3-FBA, 4-FBA and 3-HBA.** [a] Specific activities of D469T at 50 mM
366 3-FBA/50 mM HPA, 30 mM 4-FBA/50 mM HPA and 15 mM 3-HBA/ 30 mM HPA were 4.63 µmol/mg/min,
367 0.45 µmol/mg/min, and 0.0071 µmol/mg/min, respectively. [b] The kinetic data of D469T and D469T/R520Q were from [37].
368 [c] Values were not determined (n.d.) due to very low activity. Standard deviations are given in parentheses. All mutants
369 have been determined to achieve e.e.s in the range 45-62% with 3-FBA, but due to partial product racemization in the
370 derivatisation steps, these values are probably considerably underestimated.

371

372 **3.3. Computational modelling of S385E/Y/D469T/R520Q structures and substrate binding**

373 The shifts in k_{cat} are not the same for all three substrates, suggesting that they are due to
374 modified binding orientation or proximity to the catalytic residues, rather than general effects on
375 catalytic enzyme residues or the cofactors. For 3-HBA, the proximity to R358 also appears to be less
376 important than that of 3-FBA and 4-FBA. Computational modelling was used to begin to rationalise
377 these observations structurally, and provide testable hypotheses for future structural studies. We
378 have previously used the same approach with transketolase, and shown that it predicts
379 experimental K_M values with an R^2 of 0.82 [37], partly because the active-site residues do not move
380 significantly upon binding of substrates. This approach was therefore shown to be effective at
381 rationalising the effect of TK mutations on various substrates and so was used here to extend the
382 models to S385 mutations within the previously modelled D469T/R520Q active-site.

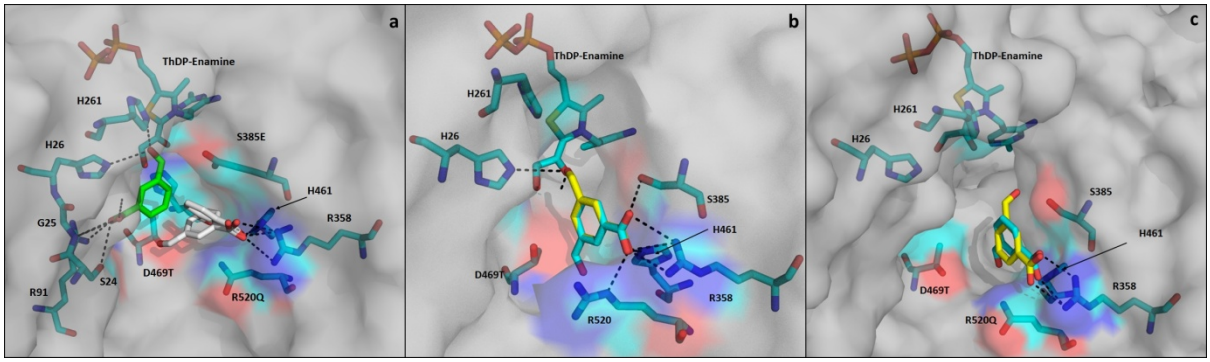
383 3-FBA was expected to bind in the S385E/D469T/R520Q active-site in a rather different
384 conformation than for D469T and D469T/R520Q. S385E/D469T/R520Q bound to 3-FBA was
385 modelled in AUTODOCK and compared in Fig. 3 to those obtained previously for D469T and
386 D469T/R520Q [37, 39]. Docking of 3-FBA in S385E/D469T/R520Q revealed two possible binding
387 clusters, with different binding energies and populations (Fig. 3a). The first binding cluster had a
388 lower population (38%) than the second (62%), and resulted from interactions between R358 and
389 H461, and the carboxylate moiety on 3-FBA. It was the only binding cluster found previously in the
390 models for D469T (Fig. 3b) and D469T/R520Q (Fig. 3c). However, this cluster is poorly oriented for
391 catalysis with the aldehyde at 5-7 Å away from the cofactor-enamine intermediate. The second
392 cluster appeared only with S385E/D469T/R520Q, and placed the aldehyde in a more catalytically
393 productive location. Interestingly, the 3-FBA in this cluster was bound very differently, with no
394 interaction between R358 and the carboxylate moiety of 3-FBA. The carboxylate moiety of 3-FBA
395 instead formed hydrogen bonds with the side-chains of R91 and H26, and with the backbone NH of
396 G25. The aldehyde was held in place by forming hydrogen bonds with H26 and H261, placing the
397 aldehyde 4.6 Å away from the reactive enamine intermediate. This distance is slightly shorter than

398 the closest predicted previously with D469T (4.9Å) [37], and could contribute to the improved k_{cat}

399 observed experimentally. The new binding mode found only in S385E/D469T/R520Q could also

400 potentially explain the alleviation of substrate inhibition.

401



402

403 **Fig. 3. The computational docking of 3-FBA into a) S385E/D469T/R520Q, b) D469T, and c) D469T/R520Q active-sites.** For

404 S385E/D469T/R520Q the first cluster is shown with white sticks, and the second cluster is shown with green sticks. For

405 D469T and D469T/R520Q, the two conformations within the single identified cluster are shown with yellow (catalytically

406 productive) and cyan (catalytically unproductive) sticks. Calculated hydrogen bonds are shown as black dotted lines.

407

408 Computational docking was also used to investigate the significantly improved k_{cat} found

409 with the S385Y mutation in S385Y/D469T/R520Q for both 4-FBA and 3-HBA. Docking revealed only

410 one binding cluster and conformation of 4-FBA in S385Y/D469T/R520Q (Fig. 4a), which provides a

411 possible entropic explanation for the improved k_{cat} relative to that of D469T/R520Q, for which many

412 clusters were predicted [37]. Autodock also previously predicted a single binding cluster for 4-FBA in

413 D469T [37], which has a 3-fold higher k_{cat} than S385Y/D469T/R520Q (Table 1). The distance from the

414 enamine intermediate to the aldehyde in 4-FBA was predicted to be 5 Å in S385Y/D469T/R520Q,

415 which is the same as for the closest conformer predicted with D469T [37]. The predicted binding of

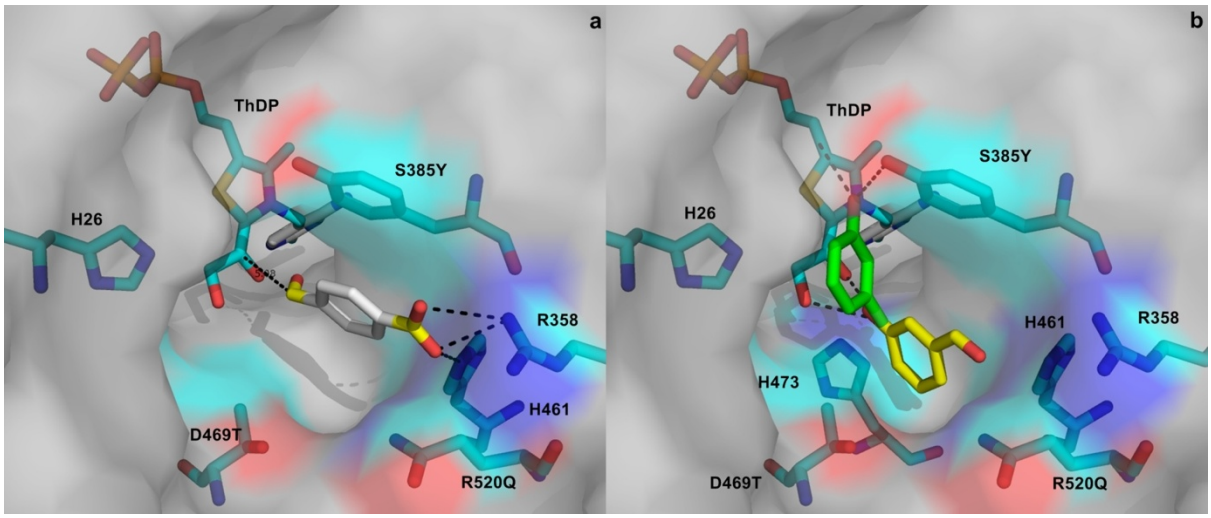
416 4-FBA in S385Y/D469T/R520Q also showed possible pi-stacking between 4-FBA and the tyrosine ring

417 of S385Y, along with hydrogen bonding / ionic interaction between the 4-FBA carboxylate moiety

418 and R358. However, a similar K_m to those of D469T/R520Q, S385T/D469T/R520Q and similar gains in

419 k_{cat} for both mutations to S385, suggest that pi stacking is not a key factor.

420



421

422 **Fig. 4. The computational docking of a) 4-FBA and b) 3-HBA into the S385Y/D469T/R520Q active site.** For 4-FBA a single
423 docking cluster was observed (white sticks). For 3-HBA two binding clusters were observed and shown in green sticks
424 (catalytically productive) and yellow sticks (catalytically unproductive). Calculated hydrogen bonds are shown as black
425 dotted lines.

426

427 Computational docking was also used to investigate the similarities and differences between
428 4-FBA and 3-HBA activity in S385Y/D469T/R520Q (Figs. 4a and 4b). This variant improved the
429 specific activity towards both 4-FBA and 3-HBA, but the activity towards 3-HBA was still much lower
430 than for 4-FBA. The k_{cat} values for 4-FBA and 3-HBA were similar but their K_M values were at least
431 20-fold different. Consistent with this experimental observation, only two out of the four major
432 docking clusters were found to be deep within the active site, and only one (highlighted in green in
433 Fig. 4b) was in a catalytically productive proximity, indicating a greater prevalence of non-specific
434 binding for 3-HBA relative to that of 4-FBA. 3-HBA did not form the pi-stacking interaction with the
435 new tyrosine ring in S385Y that was observed with 4-FBA. It is also not possible for 3-HBA to form
436 any favourable electrostatic interaction with R358 such as that observed with 4-FBA. The hydroxyl
437 moiety of 3-HBA was found to hydrogen bond with the S385Y tyrosine side-chain 4-hydroxyl group,
438 but not with R358. This is consistent with the experimental data in which R358 mutations did not

439 lead to improved activity towards 3-HBA, and also in particular why the R358T mutation was not
440 beneficial.
441

442 **4. Conclusions**

443 Random mutation of one residue in the first shell of the transketolase active-site was shown
444 to improve the enzyme activity and yield towards both carboxylated and hydroxylated aromatic
445 aldehyde substrates. Remarkably, these new TK variants gave K_M values comparable to that found
446 with wild-type TK for the natural substrate glyceraldehyde-3-phosphate, and k_{cat}/K_M values of 1900-
447 5400 s^{-1} M^{-1}. The new TK activities, which generate aromatic dihydroxyketones, open up a novel
448 route to chloramphenicol analogues via subsquent amination of the product ketone-moieties, either
449 chemically or using transaminases. Triple variants of TK were created with up to 13-fold greater
450 activity than D469T towards aromatic aldehydes. The slow bioconversion of 3-HBA relative to the
451 carboxylated substrates was predominantly due to a lower affinity between enzyme and substrate.
452 Interestingly, a comparison of the kinetics for all variants towards the three aromatic substrates
453 indicated that the binding modes for each substrate in their best respective variant are different,
454 suggesting one mechanism by which natural enzyme specificities could diverge. For those aiming to
455 evolve multiple substrate specificities using target libraries and substrate walking approaches [47],
456 such divergence of substrate binding modes could have important implications for the choice of
457 directed evolution strategies in subsequent steps, as different substrates may require distinctly
458 different libraries to be screened. Computational docking was able to rationalise the different
459 behaviours for the three substrates in terms of their different response to mutations upon k_{cat}, K_M
460 and substrate inhibition, and supported the hypothesis for divergent binding modes that could be
461 tested in future crystallographic studies.

Acknowledgement

The authors would like to thank the support from the Royal Thai Government for Panwajee Payongsri and UCL Chemistry Department for David Steadman.

References

[1] Breuer M, Hauer B. Carbon–carbon coupling in biotransformation. Curr Opin Biotechnol 2003;14:570-6.

[2] Enders D, Narine A. Lessons from nature: biomimetic organocatalytic carbon-carbon bond formations. J Org Chem 2008;73:7857-70.

[3] Fessner WD. Enzyme mediated C-C bond formation. Curr Opin Chem Biol 1998;2:85-97.

[4] Resch V, Schrittwieser JH, Siirola E, Kroutil W. Novel carbon-carbon bond formations for biocatalysis. Curr Opin Biotechnol 2011;22:793-9.

[5] Wohlgemuth R. C2-Ketol elongation by transketolase-catalyzed asymmetric synthesis. J Mol Catal B: Enzym 2009;61:23-9.

[6] Brovetto M, Gamenara D, Saenz Mendez P, Seoane GA. C-C bond-forming lyases in organic synthesis. Chem Rev 2011;111:4346-403.

[7] Fesko K, Gruber-Khadjawi M. Biocatalytic methods for C-C bond formation. ChemCatChem 2013;5:1248-72.

481 [8] Hoyos P, Sinisterra JV, Molinari F, Alcantara AR, Dominguez de Maria P. Biocatalytic
482 strategies for the asymmetric synthesis of α-hydroxy ketones. Acc Chem Res 2010;43:288-
483 99.

484 [9] Fessner WD, Helaine V. Biocatalytic synthesis of hydroxylated natural products using
485 aldolases and related enzymes. Curr Opin Biotechnol 2001;12:574-86.

486 [10] Morris KG, Smith MEB, Turner NJ, Lilly MD, Mitra RK, Woodley JM. Transketolase from
487 *Escherichia coli*: A practical procedure for using the biocatalyst for asymmetric carbon-
488 carbon bond synthesis. Tet Asymm 1996;7:2185-8.

489 [11] Samland AK, Sprenger GA. Microbial aldolases as C-C bonding enzymes: unknown
490 treasures and new developments. Appl Microbiol Biotechnol 2006;71:253-64.

491 [12] Takayama S, McGarvey GJ, Wong C-H. Enzymes in organic synthesis: recent
492 developments in aldol reactions and glycosylations. Chem Soc Rev 1997;26:407.

493 [13] Toone EJ, Whitesides GM. Enzymes as catalysts in carbohydrate synthesis. In: Bednarski
494 MD, Simon ES, editors. ACS Symposium Series 1991; 466:1-22.

495 [14] Sprenger GA, Schorken U, Sprenger G, Sahm H. Transketolase A of *Escherichia coli* K12.
496 Purification and properties of the enzyme from recombinant strains. Eur J biochem
497 1995;230:525-32.

498 [15] Schenk G, Duggleby RG, Nixon PF. Properties and functions of the thiamin diphosphate
499 dependent enzyme transketolase. Int J Biochem Cell Biol 1998;30:1297-318.

500 [16] Schneider G, Lindqvist Y. Crystallography and mutagenesis of transketolase:
501 mechanistic implications for enzymatic thiamin catalysis. Biochim Biophys Acta
502 1998;1385:387-98.

503 [17] Schörken U, Sprenger GA. Thiamin-dependent enzymes as catalysts in chemoenzymatic
504 syntheses. Biochim Biophys Acta 1998;1385:229-43.

505 [18] Hailes HC, Dalby PA, Lye GJ, Baganz F, Micheletti M, Szita N, Ward JM. alpha,alpha '-
506 Dihydroxy ketones and 2-amino-1,3-diols: Synthetic and process strategies using
507 biocatalysts. Curr Org Chem 2010;14:1883-93.

508 [19] Smith MEB, Smithies K, Senussi T, Dalby PA, Hailes HC. The first mimetic of the
509 transketolase reaction. Eur J Org Chem 2006;2006:1121-3.

510 [20] Galman JL, Steadman D, Haigh LD, Hailes HC. Investigating the reaction mechanism and
511 organocatalytic synthesis of alpha,alpha'-dihydroxy ketones. Org Biomol Chem
512 2012;10:2621-8.

513 [21] Bolte J, Demuynck C, Samaki H. Utilization of enzymes in organic chemistry:
514 Transketolase catalyzed synthesis of ketoses. Tet Letts 1987;28:5525-8.

515 [22] Dalmas V, Demuynck C. An efficient synthesis of sedoheptulose catalyzed by spinach
516 transketolase. Tet Asymm 1993;4:1169-72.

517 [23] Demuynck C, Bolte J, Hecquet L, Dalmas V. Enzyme-catalyzed synthesis of
518 carbohydrates: synthetic potential of transketolase. Tet Letts 1991;32:5085-8.

519 [24] Kobori Y, Myles DC, Whitesides GM. Substrate specificity and carbohydrate synthesis
520 using transketolase. J Org Chem 1992;57:5899-907.

521 [25] Myles DC, Andrulis PJ, Whitesides GM. A transketolase-based synthesis of (+)-exo-
522 brevicomin. Tet Letts 1991;32:4835-8.

523 [26] Humphrey AJ, Parsons SF, Smith MEB, Turner NJ. Synthesis of a novel N-
524 hydroxypyrrolidine using enzyme catalysed asymmetric carbon–carbon bond synthesis. Tet
525 Letts 2000;41:4481-5.

526 [27] Hecquet L, Bolte J, Demuynck C. Enzymatic synthesis of “natural-labeled” 6-deoxy-L-
527 sorbose precursor of an important food flavor. Tetrahedron 1996;52:8223-32.

528 [28] Ziegler T, Straub A, Effenberger F. Enzyme-catalyzed synthesis of 1-
529 deoxymannojirimycin, 1-deoxynojirimycin, and 1,4-dideoxy-1,4-imino-D-arabinitol. Angew
530 Chem Intl Ed 1988;27:716-7.

531 [29] Ingram CU, Bommer M, Smith ME, Dalby PA, Ward JM, Hailes HC, Lye GJ. One-pot
532 synthesis of amino-alcohols using a de-novo transketolase and beta-alanine: pyruvate
533 transaminase pathway in *Escherichia coli*. Biotechnol Bioeng 2007;96:559-69.

534 [30] Chen BH, Hibbert EG, Dalby PA, Woodley JM. A new approach to bioconversion reaction
535 kinetic parameter identification. AIChE J 2008;54:2155-63.

536 [31] Hibbert EG, Senussi T, Smith ME, Costelloe SJ, Ward JM, Hailes HC, Dalby PA. Directed
537 evolution of transketolase substrate specificity towards an aliphatic aldehyde. J Biotechnol
538 2008;134:240-5.

539 [32] Cazares A, Galman JL, Crago LG, Smith ME, Strafford J, Rios-Solis L, Lye GJ, Dalby PA,
540 Hailes HC. Non-alpha-hydroxylated aldehydes with evolved transketolase enzymes. Org
541 Biomol Chem 2010;8:1301-9.

542 [33] Galman JL, Steadman D, Bacon S, Morris P, Smith ME, Ward JM, Dalby PA, Hailes HC.
543 alpha,alpha'-Dihydroxyketone formation using aromatic and heteroaromatic aldehydes with
544 evolved transketolase enzymes. Chem Commun 2010;46:7608-10.

545 [34] Smith MEB, Hibbert EG, Jones AB, Dalby PA, Hailes HC. Enhancing and reversing the
546 stereoselectivity of *Escherichia coli* transketolase via single-point mutations. Adv Syn Catal
547 2008;350:2631-8.

548 [35] Yi D, Devamani T, Abdoul-Zabar J, Charmantray F, Helaine V, Hecquet L, Fessner WD. A
549 pH-based high-throughput assay for transketolase: fingerprinting of substrate tolerance and
550 quantitative kinetics. Chembiochem 2012;13:2290-300.

551 [36] Ranoux A, Hanefeld U. Improving transketolase. Top Catal 2013;56:750–764.

552 [37] Payongsri P, Steadman D, Strafford J, MacMurray A, Hailes HC, Dalby PA. Rational
553 substrate and enzyme engineering of transketolase for aromatics. Org Biomol Chem
554 2012;10:9021-9.

555 [38] Paramesvaran J, Hibbert EG, Russell AJ, Dalby PA. Distributions of enzyme residues
556 yielding mutants with improved substrate specificities from two different directed evolution
557 strategies. Prot Eng Des Sel 2009;22:401-11.

558 [39] Strafford J, Payongsri P, Hibbert EG, Morris P, Batth SS, Steadman D, Smith ME, Ward
559 JM, Hailes HC, Dalby PA. Directed evolution to re-adapt a co-evolved network within an
560 enzyme. J Biotechnol 2012;157:237-45.

561 [40] Martinez-Torres RJ, Aucamp JP, George R, Dalby PA. Structural stability of *E. coli*
562 transketolase to urea denaturation. Enzym Microb Technol 2007;41:653-62.

563 [41] Asztalos P, Parthier C, Golbik R, Kleinschmidt M, Hubner G, Weiss MS, Friedemann R,
564 Wille G, Tittmann K. Strain and near attack conformers in enzymic thiamin catalysis: X-ray
565 crystallographic snapshots of bacterial transketolase in covalent complex with donor
566 ketoses xylulose 5-phosphate and fructose 6-phosphate, and in noncovalent complex with
567 acceptor aldose ribose 5-phosphate. Biochem 2007;46:12037-52.

568 [42] Nilsson U, Meshalkina L, Lindqvist Y, Schneider G. Examination of substrate binding in
569 thiamin diphosphate-dependent transketolase by protein crystallography and site-directed
570 mutagenesis. J Biol Chem 1997;272:1864-9.

571 [43] Hibbert EG, Senussi T, Costelloe SJ, Lei W, Smith ME, Ward JM, Hailes HC, Dalby PA.
572 Directed evolution of transketolase activity on non-phosphorylated substrates. J Biotechnol
573 2007;131:425-32.

574 [44] Patrick WM, Firth AE, Blackburn JM. User-friendly algorithms for estimating
575 completeness and diversity in randomized protein-encoding libraries. Protein Eng
576 2003;16:451-7.

577 [45] Rios-Solis L, Halim M, Cázares A, Morris P, Ward JM, Hailes HC, Dalby PA, Baganz F, Lye
578 GJ. A toolbox approach for the rapid evaluation of multi-step enzymatic syntheses

579 comprising a ‘mix and match’ *E. coli* expression system with microscale experimentation.
580 Biocatal Biotrans 2011;29:192-203.

581 [46] Littlechild J, Turner N, Hobbs G, Lilly M, Rawas A, Watson H. Crystallization and
582 preliminary X-ray crystallographic data with *Escherichia coli* transketolase. Acta Crystallog
583 Sect D: Biol Crystallog 1995;51:1074-6.

584 [47] Savile CK, Janey JM, Mundorff EC, Moore JC, Tam S, Jarvis WR, Colbeck JC, Krebber A,
585 Fleitz FJ, Brands J, Devine PN, Huisman GW, Hughes GJ. Biocatalytic asymmetric synthesis of
586 chiral amines from ketones applied to sitagliptin manufacture. Science 2010;329:305-9.

587 [48] Goodsell DS, Morris GM, Olson AJ. Automated docking of flexible ligands: applications
588 of AutoDock. J Mol Recognit 1996;9:1-5.