Rational substrate and enzyme engineering of transketolase for aromatics

Panwajee Payongsri\textsuperscript{a}, David Steadman\textsuperscript{b}, John Strafford\textsuperscript{a}, Andrew MacMurray\textsuperscript{a}, Helen C. Hailes\textsuperscript{b}, and Paul A. Dalby\textsuperscript{a}

Received (in XXX, XXX) XthXXXXXX 20XX, Accepted Xth XXXXXXXX 20XX
DOI: 10.1039/b000000x

The uses of 3-formylbenzoic acid and 4-formylbenzoic acid as molecular probes along with previous and new transketolase mutants revealed the factors governing the rate of reaction between transketolase and aromatic aldehydes. The novel \(\alpha,\alpha\)-dihydroxyketones were produced at 15 to 30-fold higher yields and up to 250-fold higher specific activities with D469T TK when compared to those obtained for benzaldehyde.

Introduction

Stereospecifically controlled carbon-carbon bond formation is a key goal of organic synthesis. While many carbon-carbon bond forming chemical reactions can be achieved in a single step, several challenges still exist,\textsuperscript{1} and enzymes such as transketolase (TK) offer an attractive alternative due to their high stereospecificity and selectivity.\textsuperscript{2–8} Transketolase (E.C. 2.2.1.1) is a thiamine diphosphate (ThDP) dependent enzyme which plays a crucial role in the non-oxidative phase of the pentose phosphate pathway and Calvin cycle,\textsuperscript{9} where it catalyses a reversible transfer of a two-carbon ketol group from a ketose sugar to aldose.\textsuperscript{10, 11} The use of hydroxypyruvate as the ketol donor leads to the liberation of carbon dioxide and hence an effectively irreversible reaction, which is synthetically attractive due to the higher yields achieved. The dihydroxyketone functionality generated in these reactions is present in a wide range of natural products,\textsuperscript{4, 12, 13} and TK has been used in multi-enzyme cascades to synthesise several chiral sugars and their analogues.\textsuperscript{2, 14} (+)-exo-brevicomin,\textsuperscript{15} and chiral amino-diols.\textsuperscript{16, 17}

Several residues around the TK active-site are highly conserved across species and also throughout the ThDP-dependent enzyme family.\textsuperscript{18} Many are directly involved in substrate recognition and binding, which dictate the substrate preference. Transketolases from different species can accept a wide range of hydroxylated and non-hydroxylated aliphatics, as well as cyclic and aromatic aldehydes.\textsuperscript{10, 11, 19, 20} and nitroso aromatics.\textsuperscript{21} However, the hydroxylated substrates are far more readily accepted due to highly conserved histidine and aspartate residues, which directly interact with the hydroxyl groups of acceptor substrates. D469 (\textit{E. coli} TK) also governs the stereospecificity and stereoselectivity of transketolase.\textsuperscript{22–23} Directed evolution of TK, especially with mutations at D469, has been shown to enhance and reverse the stereoselectivity of TK and improve the reaction towards non-hydroxylated substrates.\textsuperscript{23–25} This has increased the spectrum of aliphatic substrates accepted by transketolase and its potential in organic synthesis applications.

Improving the acceptance of aromatic aldehydes by TK would pave the way for the biocatalytic synthesis of novel aromatic dihydroxyketones, and via subsequent transamination, access aromatic aminodiol analogues as the basis of novel routes to antibiotics such as chloramphenicol analogues. Wild-type (WT) \textit{E. coli} TK has been reported to accept only a few aromatic aldehydes such as phenylacetaldheydes, and with low product yields from overnight reactions, often containing by-products.\textsuperscript{20} More challenging are benzaldehyde and the heteroaromatic aldehydes such as 2-furaldehyde and 2-thiophene-carboxaldehyde which are not accepted by wild-type \textit{E. coli} TK before the hydroxypyruvate substrate degrades,\textsuperscript{20, 26} although some limited activity has been reported for yeast TK.\textsuperscript{10} Recent mutants of \textit{E. coli} TK: D469E; D469T; D469K and F434A, each resulted in the acceptance of benzaldehyde, 2-furaldehyde and 2-thiophene-carboxaldehyde, but again with very limited activity and product yields of just 1-10%.\textsuperscript{26}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{Scheme1}
\caption{The reaction between hydroxypyruvate and aldehyde catalysed by TK to generate dihydroxyketones.}
\end{figure}

This journal is © The Royal Society of Chemistry [year]
Table 1. Aldehyde substrates used and products formed.

<table>
<thead>
<tr>
<th>Aldehyde TK substrate</th>
<th>Dihydroxyketone product</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benaldehyde 2a</td>
<td>3a</td>
</tr>
<tr>
<td>3-Formylbenzoic acid (3-FBA) 2b</td>
<td>3b</td>
</tr>
<tr>
<td>4-Formylbenzoic acid (4-FBA) 2c</td>
<td>3c</td>
</tr>
<tr>
<td>Propanal 2d</td>
<td>3d</td>
</tr>
</tbody>
</table>

The current work aimed to identify the factors governing the bioconversion of aromatic aldehydes by transketolase, beginning with the known D469 mutants of E. coli TK. As the poorly accepted benzaldehyde 2a has few features that would allow the enzyme to form specificity enhancing interactions, carboxylated analogues 3-formylbenzoic acid (3-FBA) 2b and 4-formylbenzoic acid (4-FBA) 2c were chosen to re-establish electrostatic and hydrogen bonding interactions similar to those found with the natural phosphorylated substrates. To confirm that these substrates were interacting with the phosphate binding residues in TK, a series of rationally designed mutants were created to probe the importance of each residue for the observed activity with aromatic substrates. Computational docking of 3-FBA 2b and 4-FBA 2c into energy-minimised mutant structure models provided further insights into the impact of each mutation upon activity and substrate binding.

Results and Discussion

Screening for novel activities

TK crystal structures and earlier mutagenesis studies revealed several highly conserved residues that contribute to substrate specificity, including interactions with the phosphate moiety in natural substrates. Furthermore, ribose-5-phosphate (R5P) can bind the TK active site in its cyclic six-membered ring form (cR5P), as shown in Figure 1a. This suggested that some previously generated E. coli TK mutants might accept the aromatic ring of benzaldehyde in a similar manner and that presentation of negatively charged 3- or 4- substituents towards the phosphate binding residues could improve substrate affinity. This hypothesis was supported by the docking of 3-FBA 2b into D469T containing the ThDP-enamine intermediate (Figure 1b), and comparing this to the crystal structure for E. coli TK bound to cR5P (Figure 1a).

In the crystal structure residues R358, S385, H461, and R520 interact with the phosphate group of cR5P, via a mixture of hydrogen bonds and charge-charge interactions, similar to those found in the yeast TK structure complexed with erythrose-4-phosphate. Similar interactions were found between the negatively-charged carboxylate group and residues R358, S385, H461, and R520 when 3-FBA 2b was docked into D469T. The orientation of the six-membered ring in cR5P was also found to be similar to the aromatic ring of 3-FBA 2b although the latter was pulled further away from the ThDP cofactor. The new substrates 3-FBA 2b and 4-FBA 2c, therefore provide negatively charged moieties for potential electrostatic interactions with the phosphate binding residues of TK, with carboxylate pKₐs of 3.8 and 3.7, for 3-FBA 2b and 4-FBA 2c respectively.

Table 2. Performance of TK mutants towards benzaldehyde derivatives.

<table>
<thead>
<tr>
<th>Mutant</th>
<th>Reported isolated yield a</th>
<th>Conversion yield b</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2a</td>
<td>3-HBA</td>
</tr>
<tr>
<td>WT</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>D469T</td>
<td>2%</td>
<td>4%</td>
</tr>
<tr>
<td>D469Y</td>
<td>4%</td>
<td>n.d.</td>
</tr>
<tr>
<td>D469K</td>
<td>2%</td>
<td>n.d.</td>
</tr>
<tr>
<td>D469E</td>
<td>2%</td>
<td>0%</td>
</tr>
<tr>
<td>F434A</td>
<td>10%</td>
<td>6%</td>
</tr>
</tbody>
</table>

aReported, except D469Y where the reaction was carried out for 48 h. Isolated yields were obtained after purification and therefore incur minor losses relative to reaction conversion yields. (n.d. is not determined).

bConversion yields were determined from the consumption of the aldehydes after 24 h at 50 mM 3-FBA 2b and 50 mM HPA, or 30 mM 4-FBA 2c and 30 mM HPA, in 50 mM Tris buffer, pH 7.0, 2.4 mM ThDP and 9 mM MgCl₂ using 10% v/v clarified lysate in a total reaction volume of 300 μL.

to 3-FBA 2b, 4-FBA 2c were investigated, and compared to isolated yields for benzaldehyde 2a and 3-hydroxybenzaldehyde (3-HBA) using the TK variants (Table 2). A 24 hour time-course reaction was performed using 3-FBA 2b and 4-FBA 2c with WT-TK, D469T, D469Y and F434A and monitored by HPLC. Only D469T gave any significant conversion of 3-FBA 2b or 4-FBA 2c, producing 3b and 3c, respectively. While 3-FBA 2b gave 3b in 65% conversion yield, 4-FBA 2c gave rise to a 30% conversion yield (Table 2). These yields were both considerably higher than those obtained previously with any benzaldehyde substrate and mutant combination (0-10%).

The reaction of 3-FBA 2b with TK variant D469T reached a final yield of 65% yield within 2 hours, compared to the 2% yield reached after 17 hours with 2a. This potentially represents an increase in specific activity of up to 250-fold with 3-FBA 2b relative to 2a, although enzyme inactivation by exposure to unreacted aldehyde could also be a factor for the slower reaction with 2a. Since D469T TK appears to accept aromatic aldehydes, it was tested against aromatic aldehydes bearing non-acid groups at the 3- and 4-positions, namely 3- and 4-methylbenzaldehyde, 3- and 4-methoxymethylbenzaldehyde and 3- and 4-isopropylbenzaldehyde (see...
Table S1) in order to confirm the importance of the negatively charged species at 3- and 4- positions. In all cases no conversions were noted which confirms the importance of the acidic functionality and hydrophobic interaction alone was insufficient to obtain high turnover for benzaldehyde derivatives. Although, the bioconversions of two hydrophobic heteroaromatic aldehydes; 2-phenylpropanal and phenylacetaldehyde, by D469T mutant were reported to achieve 50% conversion, the reactivity and factors influencing their reactions with D469T cannot be compared with benzaldehyde derivatives. These aldehydes are significantly more reactive than the benzaldehydes due to electronic and steric effects.

The reaction using D469T and 3-FBA 2b was scaled-up and carried out over 18 hours, and after purification 3b was isolated in 88% yield. Notably the D469T reaction with 2a also produced a double addition by-product (5% isolated yield), whereas with 3-FBA 2b only a single product 3b was observed by HPLC and isolated. The solubility of 3-FBA 2b is much greater than 2a, which may have enhanced the reaction rate. However, no biphasic behaviour was observed with either 2a or 3-FBA 2b at 50 mM concentrations.

**Figure 1.** Computational docking of 3-FBA 2b and 4-FBA 2c into the active sites of D469T-based TK, and comparison with the binding of cyclic ribose-5-phosphate in the wild type active site. (a) Crystal structure of ribose-5-phosphate bound to *E. coli* WT-TK (PDB: 2R5N). Computational docking of 3-FBA 2c in models of D469T (b), and D469T/R520Q (c), each derived from the wild-type structure (PDB: 1QGD). (d) Computational docking of 4-FBA 2e in D469T. Hydrogen bonds to substrates are represented by black dotted lines. ThDP-enamine bound TK structure was obtained from computational docking of energy-minimised ThDP-enamine into modelled mutant structures.
The activity with D469T reflects the generally higher yields (2-6%) previously obtained for this mutant relative to the other D469 mutants on benzaldehyde 2a, 3-hydroxybenzaldehyde and hetero-aromatic aldehydes. Crystal structures of yeast and E. coli TK reveal that D469 interacts with the α-hydroxyl group of substrates, and replacing this negatively charged residue with the more hydrophobic and uncharged threonine was reported to enhance the activity toward non-hydroxylated substrates, due to the more favorable interactions with non-polar groups. For similar reasons, D469T also previously gave a low activity towards benzaldehyde that was not observed in the WT enzyme. A decrease in steric hindrance for benzaldehyde to access the active site may also play a role in its reactivity with D469T.

Docking of 3-FBA 2b into D469T (Figure 1b) and alignment of the WT-TK structure (not shown) reveals that the carboxyl moiety of D469 would be just 2.9 Å from the aromatic ring of 3-FBA 2b, compared to a distance of 4.1 Å from the D469T side chain hydroxyl group. This supports steric hindrance as a factor, but more important is that the D469T mutation avoids the burial of a charged group when 3-FBA 2b is bound.

Our observations with D469Y are consistent with steric hindrance being a key factor. While D469Y accepts the non-hydroxylated aliphatic substrate propionaldehyde 2d to give 3d in high yield, and even reverses the product ee, it only gave a 4% isolated yield after 48 hours with benzaldehyde 2a, and did not accept the larger 3-FBA 2b and 4-FBA 2c derivatives (Table 2). The aromatic tyrosine side chain has much greater steric bulk than threonine, which may restrict the access of benzaldehydes to the C-2 carbon of the thiazolidine ring resulting in poor or no activity. Interestingly, F434A did not accept 3-FBA 2b or 4-FBA 2c (Table 2), even though activity has been previously reported for 2a. This suggested that while F434A produces a hydrophobic pocket large enough to accept benzaldehyde (10% yield), it may not accept the aromatic ring in an orientation required to accommodate the added steric bulk or charge from the 3-FBA 2b or 4-FBA 2c. This view is consistent with the lower isolated yield observed previously with 3-HBA (6%) which is uncharged and slightly smaller than either 3-FBA 2b or 4-FBA 2c.

Kinetics of mutants with 3-FBA 2b and 4-FBA 2c

To determine the rate enhancing influence of interactions between the phosphate binding residues R358, H461, or R520, and the carboxyl groups of 3-FBA 2b and 4-FBA 2c, a series of mutations of these residues were made in D469T based on previously identified functional mutations. Mutations at these residues in yeast TK were previously shown to lower the affinity towards phosphorylated substrates, and improve the activity towards non-phosphorylated substrate. Due to their influences in the activities towards phosphorylated and non-phosphorylated substrates, they were chosen as targets for studying the enzyme-substrate affinity where the phosphate group was mimicked by a carboxyate. The yields and kinetic parameters for D469T and six further mutants of D469T are shown in Table 3 (see Figure S1-S2 for Michaelis-Menten plots). For 3-FBA 2b, all mutants gave similar conversion yield after 24 hours, except for R358L/D469T which gave <10% yield but was also less active and R358P/D469T/R520Q with 14% yield which also has high affinity towards 3-FBA and was subject to stronger substrate and product inhibitions. The yield was independent of kinetic differences between mutants, while up to 20% Li-HPA (1) also remained, indicating possible product inhibition by the same extent for all mutants. The rates and product yields for the mutants with 4-FBA 2c were found to be much lower than for 3-FBA 2b, and the activities for H461S/D469T and H461S/D469T/R520Q were too low to obtain rate data. Lower yields at 24 hours were due to lower activities, although enzyme inactivation and HPA degradation over the longer timescales may also have had a significant influence. For example, after 24 hours 30% of 4-FBA 2c was converted by D469T, yet only 50% HPA remained, suggesting 20% HPA degradation. The causes of HPA degradation were further investigated. HPA was incubated for 18 hours, in the presence and absence of ThDP, for lysates of overexpressed wild type TK, E. coli XL-10 Gold cells with no overexpressed TK, and for a Tris buffer control. Background degradation of HPA was between 2% and 14%, with or without ThDP, even for the Tris buffer control. 5% of the background appears to be catalysed by overexpressed WT TK. This demonstrates that other host enzymes in the lystate, including background WT-ThDP, do not significantly contribute to this degradation. However, 96% HPA degradation did occur with lysates containing D469T or D469T/R520Q TK and additional ThDP (see Table S2). This suggests that HPA degradation is catalysed by D469T and D469T/R520Q containing mutants over long timescales, which is itself an interesting result.

A surprising result is that R358P/D469T increased the relative specific activity with 50 mM 3-FBA 2b by 20% compared to D469T. k_cat/K_M values for R358P/D469T and D469T/R520Q also increased to 550 ± 80 s⁻¹ M⁻¹ and 470 ± 170 s⁻¹ M⁻¹, respectively, from 240 ± 50 s⁻¹ M⁻¹ for D469T while the highest k_cat/K_M value was found to be 630 ± 115 s⁻¹ M⁻¹ in the triple mutant R358P/D469T/R520Q. Remarkably, these mutations therefore bring the k_cat/K_M of 3-FBA 2b to higher values than those observed with WT E. coli TK and glycolaldehyde (283 s⁻¹ M⁻¹) and with D469T and propionaldehyde (91 s⁻¹ M⁻¹). Yet are still much lower than those with yeast TK and the natural phosphorylated substrates erythrose-4-phosphate (1.7x10⁵ s⁻¹ M⁻¹) and ribose-5-phosphate (3.2x10⁵ s⁻¹ M⁻¹). The activity of D469T/R520Q was similar to D469T at low 3-FBA 2b concentrations but slightly lower than D469T at above 30 mM 3-FBA. By contrast, R358L/D469T had a severely impaired activity, and H461S/D469T had the lowest activity at all concentrations but was slightly recovered by introduction of R520Q in the triple mutant H461S/D469T/R520Q. The activity of R358L/D469T and yield of product obtained was too low to obtain kinetic parameters. Substrate inhibition was also observed at above 40 mM for all mutants and so inhibiting concentrations were excluded from the kinetic analysis.

Each single charge-neutralising arginine mutation, R520Q or R358P in D469T, surprisingly increased k_cat/K_M for 3-FBA 2b. While k_cat decreased slightly from 13 s⁻¹ in D469T to 6 s⁻¹, 11 s⁻¹, and 3.4 s⁻¹ for D469T/R520Q, R358P/D469T, and R358P/D469T/R520Q respectively, their K_M values decreased more significantly from 56 mM in D469T to 13 mM for D469T/R520Q, and 20 mM for R358P/D469T.
The H461S mutation in H461S/D469T had a similar though weaker effect by decreasing the $K_M$ to 29 mM, though this mutant also led to a significant drop in $k_{cat}$ to just 0.8 s$^{-1}$. Unexpectedly, the highest affinity was found to be 5.5 mM when both R358 and R520 residues were mutated to neutral amino acids in R358P/D469T/R520Q. This significant improvement in the $K_M$ is the major factor that improves the $k_{cat}/K_M$ of 13 (1.0) 66.7 (0.1) 13 (4) 6 (0.75) 468 (170) to the decreased affinities observed for natural substrates at 50 mM 3-FBA 2b and HPA, or 30 mM 4-FBA 2c and HPA, in 50 mM Tris buffer, pH 7.0, 2.4 mM ThDP and 9 mM MgCl$_2$ using 10% v/v clarified lysate in a total volume of 300 μL.

The conversion yields were determined from the consumption of the aldehydes and the productions of 3b and 3c.

The specific activity of D469T towards 3-FBA 2b and 4-FBA 2c are 4.6 μmol mg$^{-1}$ min$^{-1}$ and 0.45 μmol mg$^{-1}$ min$^{-1}$ respectively.

<table>
<thead>
<tr>
<th>Mutant</th>
<th>Substrate</th>
<th>Specific activity (relative to D469T) $^{a}$</th>
<th>Conversion yield (%)</th>
<th>$K_M$ (mM)</th>
<th>$k_{cat}$ (s$^{-1}$)</th>
<th>$k_{cat}/K_M$ (s$^{-1}$ M$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>3-FBA 2b</td>
<td>0</td>
<td>0</td>
<td>n.d.</td>
<td>n.d.</td>
<td>n.d.</td>
</tr>
<tr>
<td>D469T</td>
<td>3-FBA 2b</td>
<td>1.0</td>
<td>67 (1.0)</td>
<td>56 (10)</td>
<td>13 (1.5)</td>
<td>236 (50)</td>
</tr>
<tr>
<td>H461S/D469T</td>
<td>3-FBA 2b</td>
<td>0.08</td>
<td>65 (0.6)</td>
<td>29 (15)</td>
<td>0.8 (0.2)</td>
<td>28 (16)</td>
</tr>
<tr>
<td>D469T/R520Q</td>
<td>3-FBA 2b</td>
<td>0.75</td>
<td>66.7 (0.1)</td>
<td>13 (4)</td>
<td>6 (0.75)</td>
<td>468 (170)</td>
</tr>
<tr>
<td>H461S/D469T/R520Q</td>
<td>3-FBA 2b</td>
<td>0.21</td>
<td>62.7 (0.7)</td>
<td>11 (3)</td>
<td>1.5 (0.2)</td>
<td>138 (45)</td>
</tr>
<tr>
<td>R358P/D469T</td>
<td>3-FBA 2b</td>
<td>1.2</td>
<td>65 (2.5)</td>
<td>20 (2.5)</td>
<td>11.0 (0.7)</td>
<td>553 (80)</td>
</tr>
<tr>
<td>R358L/D469T</td>
<td>3-FBA 2b</td>
<td>&lt;0.1</td>
<td>&lt;10</td>
<td>n.d.</td>
<td>n.d.</td>
<td>n.d.</td>
</tr>
<tr>
<td>R358P/D469T/R520Q</td>
<td>3-FBA 2b</td>
<td>0.42</td>
<td>14 (1.3)</td>
<td>5.5 (1)</td>
<td>3.4 (0.2)</td>
<td>625 (115)</td>
</tr>
</tbody>
</table>

The specific activity of D469T towards 3-FBA 2b or 4-FBA 2c is only 42% of 3-FBA 2b, compared to 3-FBA 2b is only 42% of 3-FBA 2b, compared to 3-FBA 2b. Since H461S was found to decrease the specific activity of D469T towards 3-FBA 2b by 70% when a triple mutant is relatively low compared to 3-FBA 2b, which led to the greatest affinity between TK and 3-FBA 2b at 11 mM.

The activity of H461S in H461S/D469T, though this was slightly restored in the triple mutant H461S/D469T/R520Q. It is also possible that the more significant loss of $k_{cat}$ in H461S/D469T/R520Q is related to a decrease in enzyme stability or rearrangement of active site residues. Indeed, H461S was recently found to dramatically decrease the activity of TK towards glycolaldehyde and propanal 2d when combined with other active-site mutations, while the mutation R520Q stabilised certain D469 mutations that otherwise disrupted a coevolved network of residues within the active site.33 Furthermore, the instability of H461S was also indicated by the lower soluble fraction of H461S which was found to be 10% compared with 16% in wild type.33 The activity of each mutant towards 2d was therefore tested as a control. Propanal is uncharged and also too small to interact with the phosphate binding residues when bound in a productive conformation. H461S was found to decrease the specific activity for 2d from 1.39 μmol mg$^{-1}$ min$^{-1}$ with D469T, to just 0.08 μmol mg$^{-1}$ min$^{-1}$.

Standard errors are in parentheses. Specific activities and percentage yields were determined with equimolar substrates at 50 mM 3-FBA 2b and HPA, or 30 mM 4-FBA 2c and HPA, in 50 mM Tris buffer, pH 7.0, 2.4 mM ThDP and 9 mM MgCl$_2$ using 10% v/v clarified lysate in a total volume of 300 μL.


table 3. Kinetic parameters of TK mutants towards 3-FBA 2b, 4-FBA 2c (see Figure S1-S2 for Michaelis-Menten plots).
with H461S/D469T, consistent with a minor active-site rearrangement or partial destabilisation of the enzyme. The specific activity increased slightly to 0.22 µmol mg\(^{-1}\) min\(^{-1}\) in H461S/D469T/R520Q, consistent with stabilisation by R520Q.

For D469T, the specific activity towards 4-FBA 2c was 0.45 µmol mg\(^{-1}\) min\(^{-1}\), approximately 10-fold lower than for 3-FBA 2b, consistent with a 12-fold lower \(k_{cat}/K_M\). The \(K_M\) values for D469T TK indicate that 4-FBA 2c binds with 5-fold lower affinity than 3-FBA 2b even though both substrates carry the same charge, while \(k_{cat}\) was 2.5-fold lower for 4-FBA 2c. This suggests that the relative orientations of the aldehyde and carboxylate moieties are important for both productive binding and catalysis. While the para-carboxylate in 4-FBA 2c is expected to electronically favour nucleophilic attack at the aldehyde relative to the meta-carboxylate in 3-FBA 2b, such an electronic effect is clearly not a major influence on the observed kinetics for 3-FBA 2b and 4-FBA 2c. The reaction using D469T and 4-FBA 2c was scaled-up and carried out over 18 hours to give 3c in a 40% conversion yield. The product 3c was prone to rearrangement to the propan-1-one isomer on purification and was isolated in 8% yield.

Overall, the bioconversion of 4-FBA 2c was significantly impaired when any one of R358, H461 or R520 were mutated to a neutral amino acid, when considering either specific activity or yield. However, at least in the case of R520Q, this was not due to the loss of affinity towards 4-FBA 2c, where the \(K_M\) decreased 10-fold. For all mutants where \(K_M\) was measured, the removal of at least one positive charge in R358, H461 or R520 led to an improved affinity for both 3-FBA 2b and 4-FBA 2c, consistent with an increase in stability achieved by bringing the active-site net charge closer to neutral when bound to these substrates.

While the R520Q mutation led to a 10-fold improved affinity for 4-FBA 2c in D469T/R520Q, \(k_{cat}\) decreased almost 25-fold. This contrasts to 3-FBA 2b where the affinity improved 4-fold in D469T/R520Q but \(k_{cat}\) only decreased to 50% that of D469T. Moreover, the \(K_M\) of D469T/R520Q towards 4-FBA 2c is only 2-fold higher than for 3-FBA 2b, whereas the \(k_{cat}\) for 4-FBA 2c is 30-fold lower than for 3-FBA 2b. Again this suggests that the affinity gained by creating a specific interaction between the substrate carboxylate-moieties and the mutated phosphate binding pocket in TK serves to differentiate the two substrates by placing the aldehyde group of 3-FBA 2b in a more productive orientation for catalysis than with 4-FBA 2c. The conversion yield of 2c with R358P/D469T was 60-fold lower than that for D469T at 30 mM of HPA and 4-FBA 2c, consistent with the much lower specific activity. This contrasts with the activity found for R358P/D469T with 3-FBA 2b, which was higher even than the activities of D469T and D469T/R520Q (Table 3). Therefore the R358P mutation improves the affinity to 3-FBA 2b but severely impairs the activity towards 4-FBA 2c. The R358L mutation results in a partial loss of activity towards 4-FBA 2c and a slightly greater loss for 3-FBA 2b.

The removal of these positively charged amino acids has illustrated the change in charge distribution within the active site and it could provide a better microenvironment for the binding of benzaldehyde. However, the bioconversion of benzaldehyde by D469T/R520Q was less than 3% which is not significantly improved from D469T. This still implied that hydrophobic alone is insufficient for high bioconversion of aromatic aldehydes. The triple mutants have low performance, so they were not further investigated.

**Computational Docking of 3-FBA 2b and 4-FBA 2c**

The computational docking of 3-FBA 2b into D469T showed that the carboxylate group of 3-FBA 2b is in close proximity to R358, S385, H461, and R520 in D469T (Figure 1b). 3-FBA 2b was found to have two major binding clusters in D469T and the free binding energy of both clusters were similar. In one cluster the aldehyde group of 3-FBA 2b forms a hydrogen bond with H26 to bring the aldehyde into close proximity to the enamine (4.89 Å). However, in the second cluster the aldehyde was at a non-productive distance from the enamine (7.21 Å) due to the hydrogen bond formed between the aldehyde O-atom and H473, and could potentially even inhibit enzyme function. It should be noted that benzaldehyde did not dock into the D469T active site with any distinct cluster of structural poses.

The interactions between the carboxylate group of 3-FBA 2b and 4-FBA 2c were slightly different to those in D469T, whereby a small shift towards R358 removed a hydrogen bond with S385 in D469T/R520Q (Figure 1c). This also moved the 3-FBA 2b molecule to bring the aldehyde group much further away (6.59 Å) from the enamine than in D469T, consistent with the 2-fold lower \(k_{cat}\) observed experimentally for D469T/R520Q.

Only one binding cluster was predicted for 4-FBA 2c in the D469T active site (Figure 1d). While 4-FBA 2c formed hydrogen bonds between one of the oxygen atoms of the carboxylate group in 4-FBA 2c and residues R358, H461, and R520, it lacked several others found with 3-FBA 2b, consistent with the 5-fold higher \(K_M\) for 4-FBA 2c. In particular, 4-FBA 2c lacked the hydrogen bond to S385 and also all interactions with the second carboxylate O-atom. The aldehyde was placed 5.05 Å away from the enamine, which is only 0.16 Å further than for 3-FBA 2b in D469T which may play some part in the 2.5-fold lower \(k_{cat}\) for 4-FBA 2c relative to 3-FBA 2b. However, this was much closer than the distance between the aldehyde of 3-FBA 2b and the enamine in D469T/R520Q, which gave a similar \(k_{cat}\) to 4-FBA 2c in D469T, indicating that factors additional to proximity, such as orientation of the aldehyde carbonyl relative to the enamine intermediate, are also a factor. Finally, D469T/R520Q was found to have four highly populated binding clusters with similar binding energies when docking with 4-FBA 2c. These gave aldehyde to enamine distances of 4.6 Å, 6.5 Å, 8.3 Å and 8.9 Å, with 35%, 10%, 16% and 26% occupancies respectively, consistent with the poor \(k_{cat}\) observed.

Although the prediction of binding sites and the positions of 3-FBA 2b and 4-FBA 2c by AutoDock agreed with the kinetic studies, the free energies of binding (Table 4) were consistent with experimental data only for wild-type and D469T TK, whereas D469T/R520Q calculations deviated considerably from the best line of fit to the plot of experimental versus calculated \(\log(K_i)\) values (supplementary information Figure S3 and Table S3). The deviations for D469T/R520Q calculations were likely to be due to small active-site structure changes during the energy minimisation of the D469T/R520Q double mutant using CharmM. Small structure shifts may have resulted in a systematic shift in the calculations of binding free energy in AutoDock relative to those in wild-type of D469T.
Table 4. The free energy of binding, predicted $K_a$ and $K_M$ values of 3-FBA 2b and 4-FBA 2c for D469T and D469T/R520Q.

<table>
<thead>
<tr>
<th>Mutant</th>
<th>$\Delta G$ predicted (kcal/mol)</th>
<th>$K_a$ predicted (mM)</th>
<th>$K_M$ experiment (mM)</th>
<th>Distance of aldehyde to enamine (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3-FBA 2b</td>
<td>4-FBA 2c</td>
<td>3-FBA 2b</td>
<td>4-FBA 2c</td>
</tr>
<tr>
<td>D469T</td>
<td>-3.35</td>
<td>-3.15</td>
<td>3.30</td>
<td>4.64</td>
</tr>
<tr>
<td>D469T/R520Q</td>
<td>-1.96</td>
<td>-1.91*</td>
<td>35.3</td>
<td>38.7</td>
</tr>
</tbody>
</table>

The $K_a$ values were calculated from $\Delta G = -RT \ln(K_a)$ where $\Delta G$ is free energy of binding.

*Population weighted average from four docking clusters (ranged -1.87 to -1.96 kcal/mol).

**Range observed for four docking clusters.

Conclusion

We have illustrated that the naturally existing phosphate binding pocket in the E. coli transketolase active site can be used to improve the bioconversion of benzaldehyde derivatives by appropriate positioning of a carboxylate moiety on the aromatic ring. Reliance on only hydrophobic interactions between the enzyme and benzaldehyde leads to inefficient catalysis, but aromatic substituents can provide the basis for hydrogen bonding and electrostatic interactions for high catalytic conversion. In particular, the meta-carboxylate group significantly improves the activity observed in 3-FBA 2b relative to benzaldehyde. Decreasing the total net positive charge in the phosphate binding site by mutation of arginine residues led to an improved binding affinity to carboxylated benzaldehyde derivatives.

Materials and Methods

Chemicals and reagents

All chemical reagents were obtained from Sigma-Aldrich, unless otherwise stated. Lithium hydroxypropionate was prepared as described previously. Propanal 2d was purchased from AcroSeal (Acros Organics, Fisher Scientific, UK). 1,3-Dihydroxypyrazine-2-one 3d was synthesised as described previously, as a product standard for HPLC calibration.

Transketolase library

All transketolases were expressed under the control of tktA gene promoter in the plasmid pQR791 in XL-10 Gold (Stratagene, La Jolla, California, USA). WT-TK, D469T, D469T/R520Q, D469Y, and F434A were available from our previous published works. Additional mutants were constructed using the Quikchange™ site directed mutagenesis kit (Stratagene) transformed into XL-10 gold competent cells (Stratagene), and confirmed by DNA sequencing. Primers used for additional mutant construction were as follows, where the target codons are in bold, and the changed bases underlined:

R358P: GAAATCGCACAGCCAAGGGTCTCACAGATT
R358L: GAAATCTGCACAGCAATTTAACCGTCTCAAGATG
H461S: GGTGATGTTTACACCAGCGACTCCATCGGTTCGG
D469Y: TCATCCTGCGGAATAAAGGGCAGATCTCACAG
R520Q: GATCATCCTCCACACAGAATCTGGCGACAG
F434A: CTGACACCCCTACCAGGGCTAGTGTTCG
D469T: TCGGTCGCGCGAAACCGGGCGACCTCAGCAG

Enzyme preparation and quantification

Glycerol stocks of all mutants were streaked on LB plates containing 150 µg/mL ampicillin and cultured overnight at 37 °C. Single colonies were picked and cultured in 250 mL flask containing 20 mL of LB with 150 µg/mL of ampicillin for 16 hours at 37 °C, shaking at 250 rpm. Cells were harvested by centrifugation, the pellet resuspended in 50 mM Tris buffer, pH 7.0, then sonicated (MSE Soniprep 150 probe, Sanyo) on ice with 10 s on, 15 s off for 10 cycles. The lysate was clarified by centrifugation at 17,700 g for 10 min at 4 °C. The supernatant was aliquoted and stored at -80 °C. Total protein concentrations were determined by Bradford assay using BSA as a standard. The concentration of TK was determined by SDS-PAGE and densitometry as previously described.

Screening assay when using 3-FBA 2b and 4-FBA 2c

WT TK, mutants D469Y and D469T/R520Q previously reported to enhance the activity toward non-hydroxylated aldehydes, and F434A which was previously found to have some activity with benzaldehyde and other aromatic substrates, were initially screened for activity with 3-FBA 2b and 4-FBA 2c. Clarified lysate (30 µL) was incubated for 20 minutes with 140 µL of 50 mM Tris buffer, pH 7.0, and 30 µL of 10x cofactor solution (24 mM ThDP and 90 mM MgCl2). The reaction was started by addition of 100 µL of 3x substrate solution (150 mM HPA and 150 mM 3-FBA 2b or 90 mM 4-FBA 2c and 90 mM HPA, in 50 mM Tris buffer, pH 7.0, prepared in glass vials). After 24h at 25 °C, 20 µL samples were taken and added to 180 µL of 0.1% TFA, centrifuged at 13,000 rpm for 3 minutes, and the supernatants analysed by HPLC with UV detection (210 nm). The Aminex HPx-87H, 300x7.8mm HPLC column (Bio-Rad), was maintained at 60 °C, and analysis performed by isocratic flow of 0.1% TFA in water at 0.6 mL/min. The consumptions of the starting materials were monitored and the retention times for both 3-FBA 2b and 4-FBA 2c were 64 minutes.

Enzyme kinetics with 3-FBA 2b and 4-FBA 2c

Due to previously observed instabilities with purified TK mutants, clarified lysate was used throughout this study. Enzyme reactions were prepared as for the screening method, except 3-FBA 2b concentrations were varied from 10 mM to 40 mM, with HPA kept constant at 50 mM. Reactions were carried out in triplicate in glass vials at 22 °C. Samples of 20 µL were taken every 4 minutes for at least 20 minutes and added to 380
µL of 0.1% TFA, then analysed by HPLC on an ACE5 C18 reverse phase column (150 x 4.6 mm), using 0.1% TFA at 1.0 mL/min with a gradient of acetonitrile from 15% to 72% over 9 minutes. The gradient was followed by a 2 min equilibration. The retention times of 3-FBA 2b and the product 3b were 5.47 and 2.97 minutes, respectively. The K_M and k_cat of each mutant were determined by non-linear regression fitting of the Michaelis-Menten plot. Repeat kinetics were obtained with enzyme concentrations at between 0.04 – 0.07 mg/mL with negligible effect on the final kinetic parameters. Kinetics for 4-FBA 2c were obtained in the same way but limited to 6-30 mM 4-FBA 2c due to precipitation upon quenching in TFA at above 30 mM. The retention times of the 4-FBA 2c and the product 3c were 5.41 and 2.45 minutes, respectively. For 4-FBA 2c, the TK concentration in the reaction was 0.1 mg/mL for D469T and 0.17 mg/mL for D469T/R520Q. Substrate and product standards of 3-FBA 2b and 4-FBA 2c were used to calibrate the conversion yield.

Propanal activity

Clarified lysate was incubated with 10x cofactor as above for 3-FBA. The reaction was started by addition of 100 µL of 3x substrate solution (150 mM HPA, 150 mM propanal 2d in 50 mM Tris buffer, pH 7.0). Samples of 20 µL were taken every 20 minutes for 2 hours, quenched in 180 µL of 0.1% TFA, centrifuged at 13,300 rpm for 3 minutes, then analysed by HPLC as for 3-FBA 2b screening. A standard of chemically synthesised 3d was used to calibrate the product concentrations obtained.

General methods

Flash chromatography was carried out using silica gel (particle size 40-63 µm). 1H NMR and 13C NMR spectra were recorded at the field indicated using Bruker AMX300 MHz and Avance-600 MHz machine. Coupling constants are measured in Hertz (Hz) and NMR spectra were recorded at 298 K. Mass spectra were recorded on a Thermo Scientific LTQ Orbitrap XL spectrometer. Infrared spectra were recorded on a Perkin Elmer Spectrum 100 FTIR spectrometer.

3-(1,3-Dihydroxy-2-oxopropyl)benzoic acid 3b

MgCl₂ (390 µL of 100 mg/mL solution, 0.41 mmol) was added to ThDP (22 mg, 48 mmol) in water (10 mL) at pH 7. Variant D469T TK cell free lysate (2 mL; approx. 1 mg/mL) was added to the enzyme suspension and the reaction was stirred for 18 h. The pH was maintained at 7.0 by the addition of 1 M HCl using a pH stat (Stat Titirino, Metrohm). Silica was added to the reaction mixture and the solvent was removed in vacuo. The product was purified by flash silica chromatography (MeOH:CH₂Cl₂, 1:4) to afford 3-(1,3-dihydroxy-2-oxopropyl)benzoic acid as an amorphous solid (186 mg, 88%). M.p. decom.; Rf 0.57 (MeOH:CH₂Cl₂, 1:4); [α] D = +37.8 (c 0.8, MeOH); ν_max(neat)/cm⁻¹ 3356, 2946, 1690, 1634, 1412; 1H NMR (600 MHz; CDCl₃) δ 8.02 (2H, d, J 7.7 Hz, 3-H), 7.53 (2H, d, J 8.3 Hz, 2-H), 5.35 (1H, s, CHO), 4.42 (2H, s, CH₂OH); 13C NMR (150 MHz; CDCl₃) δ 211.6 (C=O), 175.0 (CO₂H), 139.8, 139.5, 130.4, 130.1, 129.2, 128.9, 78.9, 66.0; m/z (NSI-) 209 ([M-H] -, 100%), 179 (27), 149 (27); m/z (NSI-) found [M-H] - 209.0453. C₁₀H₄O₅ requires 209.0455.

4-(1,3-Dihydroxy-2-oxopropyl)benzoic acid 3c

MgCl₂ (390 µL of 100 mg/mL solution, 0.41 mmol) was added to ThDP (22 mg, 48 mmol) in water (10 mL) at pH 7. D469T TK cell free lysate (2 mL; approx. 1 mg/mL) was added and the mixture stirred for 20 minutes. In a separate flask, Li-HPA (110 mg, 1.00 mmol) was dissolved in water (10 mL) with 4-formylbenzoic acid (0.150 g, 1.0 mmol) and the pH adjusted to 7 using 1 M NaOH. This was then added to the enzyme suspension and the reaction was stirred for 18 hours. The pH was maintained at 7.0 by the addition of 1 M HCl using a pH stat (Stat Titirino, Metrohm). Silica was added to the reaction mixture and the solvent was removed in vacuo. The product was purified by flash silica chromatography (MeOH:CH₂Cl₂, 1:4) to afford 4-(1,3-dihydroxy-2-oxopropyl)benzoic acid as an amorphous solid (16 mg, 8%). M.p. decom.; Rf 0.57 (MeOH:CH₂Cl₂, 1:4); [α] D = +25.0 (c 0.8, MeOH); ν_max(neat)/cm⁻¹ 3356, 2946, 1690, 1634, 1412; 1H NMR (600 MHz; CDCl₃) δ 8.02 (2H, d, J 7.7 Hz, 3-H), 7.53 (2H, d, J 8.3 Hz, 2-H), 5.35 (1H, s, CHO), 4.42 (2H, s, CH₂OH); 13C NMR (150 MHz; CDCl₃) δ 211.6 (C=O), 175.0 (CO₂H), 139.8, 139.5, 130.4, 130.1, 129.2, 128.9, 78.9, 66.0; m/z (NSI-) 209 ([M-H] -, 100%), 179 (27), 149 (27); m/z (NSI-) found [M-H] - 209.0453. C₁₀H₄O₅ requires 209.0455.

Computational docking of 3-FBA 2b and 4-FBA 2c into TK mutant structures

A ThDP-enzyme initial structure, 3-FBA 2b, and 4-FBA 2c structures were drawn in chem3D ultra v.10 and energy minimized by MM2 calculation. The energy-minimised ThDP-enzyme was docked into the active site and compared with yeast-ThDP-enzyme crystal structure (1GPU). The structures of D469T, H461S/D469T, and D469T/R520Q were obtained from WT E. coli TK crystal structure (PDB: 1QGD) by creating the mutations and then energy minimising the structures around the ThDP-enzyme, using the CharmM forcefield, Adopted Basic NR, Implicit Generalised Born solvent model, True SHAKE constant, and 1000 steps in Discovery Studio 2.0 (Accelrys, Inc. San Diego, California, USA). Docking of 3-FBA 2b and 4-FBA 2c was carried out using Autodock 4.2 and AutoDock tools 1.5.4, with the grid centred at 11.777, 27.078, 37.195 and a grid size of 30 A x 30 A x 30 A.

Acknowledgement

The authors would like to gratefully thank the support from the Royal Thai government for Panwajee Payongrsri, UCL Chemistry Department for David Steadman, the EPSRC National Mass Spectrometry Service Centre, Swansea University, for the provision of high resolution MS data.

Notes and references

*Advanced Centre for Biochemical Engineering, Department of Biochemical Engineering, University College London, Torrington Place, London, WC1E 7JE, UK; Email p.dalby@ucl.ac.uk

Department of Chemistry, University College London, 20 Gordon Street, London WC1H 0AJ, UK.
Table of contents entry

The presence of carboxylate substituents on the benzaldehyde ring improves transketolase activity towards aromatic substrates by reconstituting interactions similar to those with cyclic ribose 5-phosphate.