Impact of cofactor-binding loop mutations on thermotolerance and activity of *E. coli* transketolase

P. Morris[a,1], L. Rios-Solis[a,2], R. García-Arrazola[a], G. J. Lye[a], P. A. Dalby[a*]

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

Present addresses:

[1]Chemical Metrology and Biometry Department, National Institute of Metrology, Klong 5, Klong Luang, Pathumthani, Thailand, 12120

[2]Faculty of Pharmacy, Autonomous University of the State of Morelos. Av. Universidad 1001, Col. Chamilpa, Cuernavaca, Morelos, 62209, México

[*] 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

Tel: +44 20 7679 9566

Abstract

Improvement of thermostability in engineered enzymes can allow biocatalysis on substrates with poor aqueous solubility. Denaturation of the cofactor-binding loops of *E. coli* transketolase (TK) was previously linked to the loss of enzyme activity under conditions of high pH or urea. Incubation at temperatures just below the thermal melting transition, above which the protein aggregates, was also found to anneal the enzyme to give an increased specific activity. The potential role of cofactor-binding loop instability in this process remained unclear. In this work, the two cofactor-binding loops (residues 185-192 and 382-392) were progressively mutated towards the equivalent sequence from the thermostable *Thermus thermophilus* TK and variants assessed for their impact on both thermostability and activity. Cofactor-binding loop 2 variants had detrimental effects on specific activity at elevated temperatures, whereas the H192P mutation in cofactor-binding loop 1 resulted in a two-fold improved stability to inactivation at elevated temperatures, and increased the critical onset temperature for aggregation. The specific activity of H192P was 3-fold and 19-fold higher than that for wild-type at 60 °C and 65 °C respectively, and also remained 2.7-4 fold higher after re-cooling from pre-incubations at either 55 °C or 60 °C for 1 h. Interestingly, H192P was also 2-times more active than wild-type TK at 25 °C. Optimal activity was achieved at 60 °C for H192P compared to 55 °C for wild type. These results show that cofactor-binding loop 1, plays a pivotal role in partial denaturation and aggregation at elevated temperatures. Furthermore, a single rigidifying mutation within this loop can significantly improve the enzyme specific activity, as well as the stability to thermal denaturation and aggregation, to give an increased temperature optimum for activity.

Keywords

Biocatalysis, thermostability, thermo-inactivation, protein engineering, cofactor-binding loops, transketolase

Abbreviations:

GA, glycolaldehyde; HPA, hydroxypyruvic acid; TFA, trifluoroacetic acid; TPP, thiamine diphosphate; TK, transketolase

1. Introduction

Enzymes are extraordinary catalysts that often provide unsurpassed fidelity and selectivity under ambient and near ambient conditions of pH, temperature and in aqueous media [1]. As target compounds in pharmaceuticals become more complex, biocatalysts are increasingly being favoured over non-biological catalysts [2,3]. However, maintaining enzyme stability remains a challenge, particularly where reaction conditions are not physiological or aqueous. For example, improvement of enzyme stability at high temperatures would enable the biotransformation of compounds that have low solubility at ambient temperatures [4]. High temperatures can also increase the rate of enzyme catalysed reactions, typically limited only by the thermal denaturation or aggregation transition temperature [5]. However, the instability of many available enzymes is a major barrier to the adoption of enzymes in organic synthesis [6].

Transketolase (TK) (EC 2.2.1.1) catalyses asymmetric carbon-carbon bond formation with considerable potential to access a wide range of building blocks for further synthesis [7]. Solubilisation of poorly water-soluble aldehydes at elevated temperatures has motivated the cloning and expression of thermostable TK enzymes [8]. Meanwhile, protein engineering has enabled the acceptance of aliphatic and poorly-soluble aromatic aldehydes with *E. coli* TK [9-14], and with aliphatic aldehydes for *Geobacillus stearothermophilus* TK [15]. However, only engineered *E. coli* TK variants have been able to accept a range of benzaldehyde derivatives to date, and so our current focus is to understand and then also improve the thermostability of this enzyme.

The optimal temperature range for wild-type (WT) TK activity was previously observed to be 20-40 °C [16]. The effect of temperature on the structure, stability, aggregation and activity of *E. coli* TK has since been characterized in detail [17], where the enzyme irreversibly aggregated at above 55 °C within 1-3 hours. Surprisingly, heating the enzyme to

40-55 °C for 1 hour, and re-cooling, increased its activity by up to 3-fold [17]. The deactivation and aggregation of *E. coli* TK at extreme pH, high temperature, and in the presence of co-solvents is strongly linked to the binding of cofactors and formation of the two cofactor-binding loops at residues 185-192 and 382-392 (*E. coli* numbering) [17-19].

While the structure of *E. coli* TK has been determined only in the holo form [20], comparative structural studies of yeast apo-TK and holo-TK showed that the cofactor-binding loops were disordered in yeast apo-TK [21]. While several parts of the enzyme active-site also interact with the cofactors, these two loops interact with one another and with TPP in holo-TK to become more ordered, and to form part of the active site [22-23]. Their relative instability suggests that protein engineering of the cofactor-binding loops could be used to minimise the inactivation of TK at high temperatures [24].

Here we aimed to use mutagenesis to further probe the roles of the two cofactor-binding loops in enzyme deactivation, annealing and aggregation. The approach taken was to mutate the loops towards those of a thermostable orthologue, as this had the potential to lead to mutations that both destabilise and stabilise the cofactor loops, and thus probe their roles and impact more fully. The structure of *Thermus thermophilus* (*T. Thermophilus*) TK has been characterised recently which provides an excellent opportunity to identify the structure and sequence differences between the cofactor-binding loops of this enzyme and that of the *E. coli* TK [25]. A mix of single, double, triple and quadruple mutants, were generated, which effectively grafted the whole of each cofactor-binding loop from *T. thermophilus* TK into *E. coli* TK. These variants were then assessed for their impact on the catalytic activity at elevated temperatures, the retention of activity after recooling, and also their thermal denaturation transitions.

2. Materials and Methods

All chemicals were obtained from Sigma-Aldrich UK unless mentioned otherwise.

2.1. Mutant construction

All mutations were introduced into the *tktA* gene in plasmid pQR791 [19], which includes an N-terminal His tag, using the Quikchange (Stratagene, La Jolla, CA) method and primers shown below, as described previously [11]. Plasmids were transformed into XL10-gold cells (Stratagene, La Jolla, CA). The G186R/H192P double-mutant was constructed from the G186R mutant template plasmid with H192P mutagenic primers modified to avoid reverting R186 back to G. The triple and quadruple mutants of loop 2 were constructed directly from wild-type using primers containing all mutations simultaneously.

Loop 1

G186R: GCATTCTACGATGACAAC**CGT**ATTTCTATCGATGGTCAC

H192P: GGTATTTCTATCGATGGT**CCG**GTTGAAGGCTGGTTCACC

Loop 2

L387N: GACCTGGCGCCGTCTAAC**AAC**ACCCTGTGGTCTGGTTCTAAAGC

W390A: CCGTCTAACCTGACCCTG**GCG**TCTGGTTCTAAAGCAATCAAC

L389K/W390A/S391E:
GCGCCGTCTAACCTGACC**AAGGCGGAA**GGTTCTAAAGCAATCAACG

L387N/L389K/W390A/S391E:
GCGCCGTCTAAC**AAC**ACC**AAGGCGGAA**GGTTCTAAAGCAATCAACG

Individual colonies were picked and cultured, and all mutations confirmed by DNA sequencing.

2.2. Over-expression and purification of His-tagged transketolases

Wild-type and mutant TKs were expressed with an N-terminal His_6-tag from *E. coli* XL10-Gold: pQR791. Cells were grown in 1 L shake flasks with 100 mL LB media (10 g l^{-1} tryptone, 5 g l^{-1} yeast extract, 10 g l^{-1} NaCl and 10 g l^{-1} glycerol) as described previously [26]. Biomass concentration was measured as optical density at 600 nm (OD600) using a spectrophotometer (Thermo Spectronic, Cambridge, UK) and converted to dry cell weight (DCW) where 1 OD600 = 0.4 gDCW l^{-1}. The growth curves and expression profiles for wild type TK and the mutants were virtually identical. TK was purified as described elsewhere [19], dialysed for 24 hours against 25 mM Tris-HCl, pH 7.5, at 4 °C then stored at 4 °C for a maximum of two weeks without loss of activity, and with no precipitation visible. Protein concentration was determined by absorbance at 280 nm, assuming a monomeric molecular weight (MW) of 72260.82 g mol^{-1} and an extinction coefficient (ε) of 93905 L mol^{-1} cm^{-1} [27].

2.3. Temperature dependence of holo-TK activity.

Purified wild-type and mutants of *E. coli* TK were prepared at 0.1 mg mL^{-1} (1.38 μM), with 2.4 mM TPP, 9 mM $MgCl_2$ and 25 mM Tris-HCl, pH 7.0, and pre-incubated at 25 °C for 30 minutes. Enzymes were then incubated for 5 minutes by placing 100 μL samples into a water bath equilibrated at 25, 55, 60, or 65 °C. Sample temperatures were monitored using a digital wired-thermometer (Topac, USA) and shown to equilibrate within 5 minutes. Reactions were initiated by addition of pre-warmed 50 μL of 50 mM Li-hydroxypyruvate (HPA), 50 mM glycolaldehyde (GA) in 25 mM Tris-HCl, pH 7.0. Aliquots were taken and quenched at various times over 180 minutes with 1 vol. 0.2 % (v/v) trifluoroacetic acid (TFA). Sample temperatures were monitored using a digital wired-thermometer (Topac, USA). Triplicate reactions were analysed by HPLC. Li-HPA thermal degradation in the absence of TK in water at temperatures up to 65 ^{o}C was found to be less than 5% after 60 min incubation.

2.4. Thermal inactivation of holo-TK variants.

Purified wild-type and mutants of *E. coli* TK were prepared at 0.1 mg mL^{-1} (1.38 μM), with 2.4 mM TPP, 9 mM $MgCl_2$ and 25 mM Tris-HCl, pH 7.0, and pre-incubated at 25 °C for 30 minutes. Thermal inactivation was initiated by placing the 100 μL samples into a water bath equilibrated at 25, 55, 60, or 65 °C. Sample temperatures were monitored using a digital wired-thermometer (Topac, USA) and shown to equilibrate within 5 minutes. Samples were removed after 1 h, immediately cooled on ice then equilibrated to 25 °C. Reactions were initiated with 50 μL of 50 mM Li-HPA,50 mM GA in 25 mM Tris-HCl, pH 7.0, and aliquots quenched at various times over 180 minutes with 1 vol. 0.2 % (v/v) trifluoroacetic acid (TFA). Triplicate reactions were analysed as above by HPLC.

2.5. Dynamic light scattering (DLS)

The spontaneous aggregation temperature of purified holo-TK variants were measured with a Zetasizer Nano S (Malvern Instruments Ltd., UK). Enzymes were prepared at 0.1 mg mL^{-1} (1.38 μM) in 25 mM Tris-HCl, pH 7.0 with 0.5 mM TPP, 5 mM $MgCl_2$. The temperature was increased step-wise at 1.0 °C per minute, from 4 to 70 °C, between measurements. Measurements of control samples of buffers with containing cofactors were subtracted from each recording. Data were acquired in triplicate with a low volume disposable sizing cuvette with a path length of 1 cm. The hydrodynamic diameters of each sample were calculated from the averaged-measurements using the Zetasizer Nano Series software V.4.20 (Malvern Instruments Ltd., Worcestershire, UK).

2.6. HPLC

A Dionex HPLC system (Camberley, UK) with a Bio-Rad Aminex HPX-87H reverse phase column (300 mm × 7.8 mm, Bio-Rad Labs., Richmond, CA, USA), controlled by Chromeleon client 6.60 software was used for the separation and analysis of erythrulose product as described previously [28].

3. Results

3.1. Design of *E. coli* TK variants

Single, double, triple and quadruple *E. coli* TK variants, culminating in whole cofactor-binding loop grafts, were created based upon sequence and structure alignment to the homologous *T. thermophilus* TK sequence. These variants were then used to assess the relative effects of individual and combined cofactor-binding loop sequence changes, on the stability of TK to high temperatures, as well as the relative contribution of the two loops to thermotolerance.

T. thermophilus TK has 651 residues compared to 680 in *E. coli* TK. This shortening of sequence due to smaller loop regions, is typical in thermostable enzymes as it reduces the flexibility and therefore inherent entropy in the protein structure. The % sequence identity between *T. thermophilus* TK and *E. coli* TK is 49%, and the comparison the cofactor-binding loops 1 and 2 from *E. coli* and *T. thermophilus* TK are shown below, with differences highlighted in bold:

E. coli TK loop 1 (185-192)	N**G**ISIDG**H**
T. thermophilus TK loop 1	N**R**ISIDG**P**

E. coli TK loop 2 (382-392)	L**A**PSN**L**T**LWS**G
T. thermophilus TK loop 2	L**T**PSN**N**T**KAE**G

The two cofactor-binding loops from available *T. thermophilus* and *E. coli* TK structures (2E6K.PDB and 1QGD.PDB) are also overlaid and compared in Fig. 1. The overall backbone structures of the two TK cofactor-loops are virtually identical between the two

organisms, forming the same hairpin in loop 1 and partial helix in loop 2, thus confirming the sequence alignment as accurate. The mutations designed to convert *E. coli* TK loops towards those of *T. thermophilus* were: (loop1) H192P, G186R, G186R/H192P; and (loop 2) L387N, W390A, L389K/W390A/S391E, and L387N/L389K/W390A/S391E. The A383T mutation of loop 2 was not included as this residue is relatively structured in the middle of a single helical turn at the beginning of the loop structure.

3.2. Temperature dependence of holo-TK variant activities.

The optimum temperature range for the wild-type TK enzyme activity has been previously reported as 20-40 °C [16], although an annealing effect has also been observed for wild-type *E. coli* TK that improves activity up to approximately 55 °C [17]. The specific activities of wild-type and variant TKs at 25 °C, 55 °C and 60 °C are compared in Fig. 2, and summarized, also with activity at 65 °C, in Table 1.

All specific activities, except for G186R, increased between 2- and 5-fold (4.4-fold for wild-type) at 55 °C relative to that at 25 °C. At 60 °C, these then decreased to between 10% and 90% of the 55 °C levels (42% for wild type), for all except H192P which increased by a further 10% to give an activity that was 3-fold higher than for wild type at 60 °C. All specific activities dropped to zero at 65 °C, except for H192P which retained 20% of the initial activity observed at 25 °C. Therefore, the optimum temperature for specific activity was increased only for H192P, and to approximately 60 °C. The activity of H192P at 65 °C was 20-fold higher than that for wild-type at 25 °C. This suggests that H192P could be useful as a biocatalyst at 65 °C, if required for increasing the substrate solubility.

In loop 1, G186R was not active under any condition. By contrast, H192P in loop1 resulted in a two-fold increased specific activity compared to wild-type TK at 25 °C, whereas the double mutant G186R/H192P had similar activity to wild type (Fig. 2A). This suggests that

while G186R is destabilizing and/or inactivating, and that the stability induced by the H192P mutation is also sufficient to rescue the destabilization/inactivation impact of G186R. This synergistic effect is potentially mediated by the direct interaction of these two residues in cofactor-binding loop 1 (Fig. 1).

For loop 2, all of the variants had a lower catalytic activity than the wild type at 25 °C, particularly for the quadruple mutant which was less than 50% as active (Fig. 2A). The single mutant L387N, and the triple mutant L389K/W390A/S391E, were the most active of the loop 2 variants at 25 °C, with 55% and 50% of the wild-type activity, respectively. The quadruple mutant L387N/L389K/W390A/S391E clearly had the lowest activity at only 20% that of wild type at 25 °C. The lower activity of loop 2 variants relative to wild type persisted at all elevated temperatures (Fig. 2B and 2C). The only notable exception was that the activity of the triple mutant L389K/W390A/S391E did not fall as sharply at 60 °C as it did for all other variants and wild-type TK (except H192P), but decreased by approximately only 10%. This indicated a broadened, and perhaps slightly increased optimum temperature, although clearly to a lesser degree than for H192P.

At 60 °C, the wild-type TK and loop 2 variants all failed to reach complete bioconversion, indicating the loss of active enzyme before the reaction was complete. Wild-type TK, achieved 50% bioconversion within 140 minutes, whereas L387N, W390A and L389K/W390A/S391E achieved only 20% bioconversion, and L387N/L389K/W390A/S391E achieved less than 10% bioconversion (Fig. 2C). There was no appreciable catalytic activity of wild-type TK, or for any of the mutants in loop 1 and loop 2, other than for H192P, when tested at 65 °C.

3.3. Activity retention after 1 h of thermal inactivation for holo-TK variants

The specific activities of holo-TK variants were also investigated after a 1 hour incubation at 55-65 °C, followed by re-cooling of the samples and measurement of activity at 25 °C. After incubation at 55 °C (Fig. 3A) for 1 h and re-cooling to 25 °C, the wild-type and the loop 1 variants H192P, and G186R/H192P gained activities that were 2.2-, 2.7- and 3-fold higher than without the 55 °C incubation (Table 1). This annealing effect had been observed previously for wild-type TK [17], and was yet even greater for the cofactor-binding loop 1 variants. By contrast, loop 2 variant L387N simply reverted back to the original activity observed at 25 °C, while the other loop 2 variants lost between 30% and 70% of the original activity.

Enzyme inactivation at 60 °C and 65 °C is already known to be driven by protein aggregation, as evidenced previously by a lag phase in the time dependence of inactivation for wild-type TK at 60 °C [17]. This lag phase was less evident at 65 °C due to faster aggregation kinetics. Increasing the 1 hour pre-incubation temperature to 65 °C was sufficient to knock out all enzyme variant activities, while at 60 °C (Fig. 3B) all except some residual activities in wild-type, G186R/H192P and H192P were lost, where H192P again retained the greatest activity.

3.4. Thermal aggregation temperatures measured by dynamic light scattering (DLS)

The effect of temperature on the Z-average hydrodynamic radius of holo-TK was determined for each variant by DLS (supplementary information Fig. S1). The Z-averages of all variants were approximately 10 nm as expected for the homodimeric enzymes, from 20 °C until a sharp increase at higher temperatures. The increase can be characterised as the temperature (T_{agg}) at which a pre-defined Z-average is reached. The values of T_{agg} for forming large aggregates (>1,000 nm), are summarised for all variants in Table 2.

For wild-type TK, a Z-average of >1,000 nm was reached at 60 °C. All of the variants appeared to aggregate to larger average particle sizes than for wild-type TK. The T_{agg} of 61.5 °C for H192P was 1.5 °C higher than for wild type, whereas that of G186R/H192P was approximately unchanged at 59.4 ^{o}C. These data indicate that H192P is more resistant to aggregation at elevated temperatures than for wild-type TK, and G186R/H192P, consistent with their activities after pre-incubations (Fig. 3) or actually at the elevated temperatures (Figs. 2B and 2C).

All of the loop 2 variants gave T_{agg} values lower than that of wild-type TK. In fact the T_{agg} values of all variants and wild-type TK correlated well to their activities at 25-60 °C (Fig. 4), demonstrating that protein stability to denaturation and aggregation was the primary mechanism of inactivation for the mutants at elevated reaction temperatures, as found previously for wild type. These results also show that cofactor-binding loop stability plays a significant role in the aggregation mechanism for TK, and that engineering them can improve their stability to temperature denaturation, as found for H192P.

The proportion of activity retained after thermal inactivation for 1 hour, was also directly correlated with the T_{agg} values (Fig. 5), such that for the 55 °C incubation, the high T_{agg} variants gave a greater annealing effect. While the mechanism leading to the observed annealing effect is still not known, these results indicate that it can also be influenced by cofactor-binding loop mutations. The correlation to T_{agg} suggests that it is related to conformational stability, and potentially arises from a degree of partial unfolding or misfolding of the cofactor loops, even within the initially purified holo-TK population, that is then re-equilibrated to the correctly folded state by incubation at temperatures elevated to just below the T_{agg}.

4. Discussion

TK contains two identical active-sites at the homodimer interface, in each of which two cofactor-binding loops become ordered upon binding of TPP and Ca^{2+} (or Mg^{2+}). The TPP cofactor forms many, mostly hydrophobic, interactions within both chains A and B, including a conserved hydrophobic contacts with the sidechains of L382 in loop 2, and I189 in loop 1. The Ca^{2+} ion also coordinates to the two phosphates of TPP. Loop 1 in chain A interacts with loop 2 chain B (and *vice versa* for the other active site) across the dimer interface, to form one side of the active-site funnel with TPP at the base. The I189 sidechain in loop 1 also forms a hydrophobic contact with P384 in loop 2, while D190 (loop 1) hydrogen bonds to the mainchain amide of A383 in loop 2. Although these contacts may lead to concerted behaviour between loops 1 and 2, a comparison of loop 1 and loop 2 variants reveals their relative effects on the thermal aggregation, annealing and deactivation of wild-type TK.

All loop 2 variants decreased T_{agg} by 5-6 °C with concomitant loss of activity at elevated temperatures, indicating that the loss of loop 2 stability led to an increased aggregation propensity. Residues 386-392 of loop 2 form many intrachain contacts with the loop packed against a core helix and beta-sheet within the Pyr domain of TK. L387 forms hydrophobic interactions with the sidechains of R358, K359 and Q362, and the loop 2 L387N mutation could retain these, while introducing a new hydrogen bond to R358, as found in *T. thermophilus* TK. Meanwhile, the partly solvent exposed L389 packs against the aromatic sidechain of Y404.

A similar hydrophobic interaction is retained in the *T. thermophilus* TK structure between a lysine sidechain and the equivalent tyrosine, suggesting that the same interaction could be formed by the L389K mutation of *E. coli* TK. However, the lysine ε-amine in *T. thermophilus* TK also forms a salt bridge to an aspartate residue in place of A395 in *E. coli* TK. The W390A loop 2 mutation removes hydrophobic interactions with the mainchain atoms

of G369 and P370 in a core α-helix, and also to G392 of loop 2. However, these interactions are not present in *T. thermophilus* TK where the resulting cavity is mainly occupied by water molecules. Finally, the S391E loop 2 mutation removes an intrachain hydrogen bond to E366, which is also not present in *T. thermophilus* TK.

Despite the potential for introducing many of the same interactions found in *T. thermophilus* TK, and a minimal likelihood of steric disruptions in each case, the loop 2 variants were still destabilising. Thermophiles are thought to improve their stability often through better packing, rigidification, deletion or shortening of loops [29]. None of the variants contained deletions to shorten the loops, whose lengths are conserved, but presumably led to the loss of optimal packing and interactions with the Pyr domain. Meanwhile, certain stabilising interactions found in *T. thermophilus* TK were not possible in the *E. coli* TK variants, such as the salt bridge formed by the lysine in *T. thermophilus* TK at the residue equivalent to L389.

In loop 1, the H192P mutation improved the thermal stability of TK to aggregation, retained better the activity at elevated temperatures, increased the optimum temperature for activity to 60 °C, enhanced the degree of activity annealing at 55 °C, and also led to a higher activity at 25 °C than for wild-type TK. Introducing a proline can rigidify loops due to the formation of a covalent bond from the side chain back to the backbone amine, thus removing free rotation about the N-Cα bond [29,30,4]. Many proteins from thermophiles with high genomic GC-content, including *T. thermophilus*, contain more prolines in their loop regions [31]. However, not all thermophiles are so dependent on proline residues for stability, and indeed *G. stearothermophilus* and *Thermatoga Maritima* TKs have Glu and His respectively at the equivalent position.

Loop 1 forms many contacts, including coordination of the Ca^{2+}-ion cofactor via the sidechain carbonyl O-atom of N185, and the backbone carbonyl O-atom of I187, as well as an intrachain hydrogen-bond between the S188 hydroxyl moiety, and the W196 indole N-atom. G186 in loop 1 provides the mainchain bridge to bring N185 and I187 into the correct orientations to coordinate Ca^{2+}.

In *T. thermophilus* TK, an arginine residue provides the same function, while also creating a salt-bridge to an aspartate. This position is also an aspartate (D184) in *E. coli* TK, and hence the G186R mutation was fully expected to produce an equivalent salt bridge. Surprisingly, G186R was found to completely inactivate *E. coli* TK. The intended salt-bridge partner, D184 in *E. coli* TK, is involved in a hydrogen bond network with the Y182 phenol, and the D155 carboxyl-group which also coordinates the Ca^{2+} ion. This critical network would be disrupted by the G186R mutation if it competes for D184. Indeed, in *T. thermophilus* TK, the Y182 position is occupied by a tryptophan. The indole N-atom then hydrogen bonds only to the equivalent of D155, and therefore the hydrogen bond network to the Ca^{2+} ion is isolated from the D184-R186 salt bridge. The impact of G186R on *E. coli* TK therefore appears to be electrostatic, and resulted in total loss of activity.

The H192 sidechain is by comparison to other loop 1 sidechains, relatively solvent exposed, despite a hydrophobic interaction with I187, and a 3.8 Å intra-strand salt-bridge to E194. When compared to the results of wild-type, G186R and the double mutant G186R/H192P TK, the H192P mutation is found to partly accommodate the otherwise disruptive G186R mutation. H192 forms a 3.8 Å intra-strand salt-bridge to E194 in *E. coli* TK. This was removed in H192P, and so the G186R mutation in the presence of H192P, would have the potential to form a new intra-strand salt-bridge to E194. This in turn would

reposition the mutant G186R sidechain away from D184, and thus provides a theoretical explanation for rescuing the disruption caused in the single mutant.

Given that H192P and G186R/H192P in loop 1 were both able to enhance the effect of annealing at 55 °C, whereas all of the loop 2 mutations abolished any annealing effect, this phenomenon is apparently linked to both cofactor loops. However, it could also simply be that aggregation and annealing are competing reactions at 55 °C. In this scenario, loop 2 variants promoted aggregation and thus obscured the annealing behviour, whereas the loop 1 variants that suppressed aggregation, allowed the annealing reaction more time to occur.

5. Conclusions

In this work, mutagenesis was targeted to TK cofactor-binding loops, to probe their role in thermal activity annealing, thermal inactivation, and aggregation, as previously implicated for the wild-type enzyme. Mutants of *E. coli* TK were designed based upon the sequence of a thermostable TK from *T. thermophilus,* to examine their impact upon stability and activity at elevated temperatures. The thermal stability of TK was successfully improved by mutation of cofactor-binding loop 1, but not loop 2. H192P in loop 1 resulted in improved stability at elevated temperatures, which correlated with the onset temperature of aggregation measured by thermal scanning. The specific activity of H192P was three times higher than that for wild-type at 60 °C. None of the loop 2 variants improved the specific activity at elevated temperatures, but their negative impact suggested that this loop was still critical in driving thermal inactivation through aggregation of the wild-type enzyme, and both loops remain good targets for further engineering of the stability of the *E. coli* TK enzyme.

Acknowledgments

The Royal Thai Government is acknowledged for the support of Phattaraporn Morris. The Mexican National Council for Science and Technology (CONACYT) is acknowledged for the support of Leonardo Rios-Solis and Ribia García-Arrazola. The UK Engineering and Physical Sciences Research Council (EPSRC) is thanked for the support of the multidisciplinary Biocatalysis Integrated with Chemistry and Engineering (BiCE) programme (GR/S62505/01) at University College London (London, UK).

References

[1] Narancic T, Davis R, Nikodinovic-Runic J, Connor KEO. Recent developments in biocatalysis beyond the laboratory. Biotechnol Lett 2014;37:943-54.

[2] Bornscheuer UT, Huisman GW, Kazlauskas RJ, Lutz S, Moore JC, Robins K. Engineering the third wave of biocatalysis. Nature 2012;485:185-94.

[3] Torrelo G, Hanefeld U, Hollmann F. Biocatalysis. Catal Letts 2014;145:309-45.

[4] Yu H, Huang H. Engineering proteins for thermostability through rigidifying flexible sites. Biotechnol Adv 2014;32:308-15.

[5] Elias M, Wieczorek G, Rosenne S, Tawfik DS. The universality of enzymatic rate-temperature dependency. Trends Biochem Sci 2014;39:1-7.

[6] Wijma HJ, Floor RJ, Janssen DB. Structure- and sequence-analysis inspired engineering of proteins for enhanced thermostability. Curr Opin Struct Biol 2013;23:588-94.

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

[8] Abdoul-Zabar J, Sorel I, Helaine V, Charmantray F, Devamani T, Yi D, de Berardinis V, Louis D, Marliere P, Fessner WD, Hecquet L. Thermostable transketolase from Geobacillus stearothermophilus: Characterization and catalytic properties. Advanced Synthesis and Catalysis. Adv Syn Catal 2013;355:116-128.

[9] Payongsri P, Steadman D, Hailes HC, Dalby PA. Second generation engineering of transketolase for polar aromatic aldehyde substrates. Enzyme Microb Technol 2015;71:45-52.

[10] Payongsri P, Steadman D, Strafford J, MacMurray A, Hailes HC, Dalby PA. Rational substrate and enzyme engineering of transketolase for aromatics. Org Biomol Chem 2012;10:9021-29.

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

[12] Hibbert EG, Senussi T, Smith MEB, Costelloe SJ, Ward JM, Hailes HC, Dalby PA. Directed evolution of transketolase substrate specificity towards an aliphatic aldehyde. J Biotechnol 2008;134:240-45.

[13] Cázares A, Galman JL, Crago LG, Smith MEB, Strafford J, Ríos-Solís L, Lye GJ, Dalby PA, Hailes HC. Non-alpha-hydroxylated aldehydes with evolved transketolase enzymes. Org Biomol Chem 2010;8:1301-9.

[14] Ranoux A, Karmee SK, Jin J, Bhaduri A, Caiazzo A, Arends IWCE, Hanefeld U. Enhancement of the substrate scope of transketolase. ChemBioChem 2012;13:1921-31.

[15] Yi D, Saravanan T, Devamani T, Charmantray F, Hecquet L, Fessner WD. A thermostable transketolase evolved for aliphatic aldehyde acceptors. Chem Commun 2015;51:480-3.

[16] Sprenger GA, Pohl M. Synthetic potential of thiamin diphosphate-dependent enzymes. J Mol Catal B: Enzym 1999;6:145-59.

[17] Jahromi RRF, Morris P, Martinez-Torres RJ, Dalby PA. Structural stability of *E. coli* transketolase to temperature and pH denaturation. J Biotechnol 2011;155:209-16.

[18] Dalby PA, Aucamp JP, George R, Martinez-Torres RJ. Structural stability of an enzyme biocatalyst. Biochem Soc Trans 2007;35:1606-9.

[19] Martinez-Torres RJ, Aucamp J, George R, Dalby PA. Structural stability of *E. coli* transketolase to urea denaturation. Enzyme Microb Technol 2007;41:653-62.

[20] Littlechild J, Turner N, Hobbs G, Lilly M, Rawas A, Watson H. Crystallization and preliminary X-ray crystallographic data with *Escherichia coli* transketolase. Acta Crystallog Sec D: Biol Crystallog 1995;51:1074-76.

[21] Sundström M, Lindqvist Y, Schneider G. Three-dimensional structure of apotransketolase. flexible loops at the active site enable cofactor binding. FEBS Letts 1992;313:229-31.

[22] Nikkola M, Lindqvist Y, Schneider G. Refined structure of transketolase from *Saccharomyces cerevisiae* at 2.0 A resolution. *J Mol Biol* 1994;238:387-404.

[23] Meshalkina LE, Kochetov GA, Brauer J, Hübner G, Tittmann K, Golbik R. New evidence for cofactor's amino group function in thiamin catalysis by transketolase. *Biochem Biophys Res Comm* 2008;366:692-97.

[24] Kochetov GA, Solovjeva ON. Structure and functioning mechanism of transketolase. Biochim Biophys Acta - Proteins and Proteomics 2014;1844:1608-18.

[25] Iino H, Naitow H, Nakamura Y, Nakagawa N, Agari Y, Kanagawa M, Ebihara A, Shinkai A, Sugahara M, Miyano M, Kamiya N, Yokoyama S, Hirotsu K, Kuramitsu S. Crystallization screening test for the whole-cell project on *Thermus thermophilus* HB8. Acta Crystallog Sec F: Struct Biol Crystal Comm 2008;64:487-91.

[26] Rios-Solis L, Morris P, Grant C, Odeleye AOO, Hailes HC, Ward JM, Dalby PA, Baganz F, Lye GJ. Modelling and optimisation of the one-pot, multi-enzymatic synthesis of chiral amino-alcohols based on microscale kinetic parameter determination. Chem Eng 2015;122:360-72.

[27] Pace CN, Vajdos F, Fee L, Grimsley G, Gray T. How to measure and predict the molar absorption coefficient of a protein. Prot Sci 1995;4:2411-23.

[28] Rios-Solis L, Halim M, Cázares A, Morris P, Ward JM, Hailes HC, Dalby PA, Baganz F, Lye GJ. A toolbox approach for the rapid evaluation of multi-step enzymatic syntheses comprising a 'mix and match' *E. coli* expression system with microscale experimentation. Biocatal Biotrans 2011;29:192-203.

[29] Russell RJ, Ferguson JM, Hough DW, Danson MJ, Taylor GL. The crystal structure of citrate synthase from the hyperthermophilic archaeon P*yrococcus furiosus* at 1.9 Å resolution. Biochemistry 1997;36:9983-94.

[30] Tian J, Wang P, Gao S, Chu X, Wu N, Fan Y. Enhanced thermostability of methyl parathion hydrolase from Ochrobactrum sp. M231 by rational engineering of a glycine to proline mutation. FEBS J 2010;277:4901-8.

[31] Suzuki Y, Oishi K, Nakano H, Nagayama T. A strong correlation between the increase in number of proline residues and the rise in thermostability of five *Bacillus* oligo-1,6-glucosidase Appl Microbiol Biotechnol 1987;26:546-51.

[32] DeLano WL. The PyMOL Molecular Graphics System. 2002. DeLano Scientific, Palo Alto, CA, USA. http://www.pymol.org.

List of Legends

Fig 1. Structural comparison of the cofactor loops. Loop 1 (left) and loop 2 (right), from wild-type *E. coli* TK, 1qgd.pdb (green) and wild-type *T. thermophilus* TK, 2e6k.pdb (cyan). The TPP cofactors and Mg^{2+} ion are shown as spheres. Residues that differ in amino acid, between the two structures are represented as sticks. Fig. generated in PyMol [32].

Fig 2. Catalytic activity of TK variants at different temperatures. The catalytic activity of purified wild-type TK compared to the variants was measured in triplicate at **A)** 25 °C, **B)**

55 °C and **C)** 60 °C in 50 mM Tris-HCl, pH 7.0. (a) loop 1 (▲) H192P (Δ) G186R/H192P (b) loop 2 (∇) L387N, (○) W390A, (▼) L389K/W390A/S391E, (♦) L387N/L389K/W390A/S391E, each compared to (●) wild-type TK. The catalytic activity was measured using 50 mM GA and 50 mM HPA as substrates and L-erythrulose determined as product.

Fig 3. Catalytic activity of TK after pre-incubation at different temperatures for 1 hour. The catalytic activity of pure wild-type TK compared to variants was measured in triplicate after incubation at **A)** 55 °C and **B)** 60 °C, in 50 mM Tris-HCl, pH 7.0 for 1 h and re-cooling for reactions at 25 °C. (a) loop 1 mutants (▲) H192P (Δ) G186R/H192P; and (b) loop 2 mutants (∇) L387N, (○) W390A, (▼) L389K/W390A/S391E, (♦) L387N/L389K/W390A/S391E are each compared to (●) wild-type TK. The catalytic activity was measured using 50 mM GA and 50 mM HPA as substrates and L-erythrulose determined as product.

Fig 4. Correlation between T_{agg} values and the specific activities of TK variants at (■) 25 °C, (♦) 55 °C, (▲) 60 °C and (×) 65 °C.

Fig 5. Correlation between T_{agg} values and the retention of activities after pre-incubating TK variants for 1 hour at (♦) 55 °C and (▲) 60 °C.

Enzyme	Specific Activity (μmol min^{-1} mg^{-1})						
	Temperature of reaction (^oC)				Temperature of pre-incubation (^oC) for 1 hr, prior to reaction at 25 ^oC		
	25	55	60	65	55	60	65
Wild-type	8.5	37.5	15.8	0.2	18.7	0.4	0
Loop 1 variants							
H192P	18.5	43.3	48.2	3.8	50.6	1.7	0
G186R	0	0	0	0	0	0	0
G186R/H192P	7.2	18.4	1.4	0.94	21.3	0.3	0
Loop 2 variants							
L387N	4.7	18.5	7.4	0	4.9	0	0
W390A	3.1	15.1	7.4	0	2.3	0	0
L389K/W390A/S391E	4.2	9.4	8.3	0	2.8	0	0
L387N/L389K/W390A/S391E	1.9	4.8	1.1	0	0.6	0	0

Table 1. Specific activity of TK variants at different temperatures, and after pre-incubation at elevated temperatures. Specific activity of wild-type and variant TKs in loop 1 and 2 were determined at 25, 55, 60 and 65 °C directly, and also 25 °C after pre-incubation for 1 h at 55, 60 and 65 ^oC. All standard deviations were determined to be $\leq \pm 0.1$ μmol min^{-1} mg^{-1} (n=3).

Enzyme variant	Temperature (T_{agg}) (^{o}C)
Wild-type	60.0
Loop 1 variants	
H192P	61.5
G186R/H192P	59.4
Loop 2 variants	
L387N	55.2
W390A	55.0
L387N/W390A/S391E	54.0
L387N/L389K/W390A/S391E	53.7

Table 2. Temperature (T_{agg}) at which TK variant samples reach an average particle size of >1,000 nm. Temperature at which the average particle diameter exceeds 1,000 nm for wild-type and variant TK, as determined by dynamic light scattering. TK at 0.1 mg mL^{-1} (1.38 μM) was incubated at 25 °C for 1 h in 25 mM Tris–HCl, pH 7.0, 5 mM $MgCl_2$, 0.5 mM TPP prior to measurements. Mean diameter (Z-ave) was calculated from the % intensity distribution at each temperature. All temperatures are determined to the nearest ± 0.1 ^{o}C.

Fig. 1.

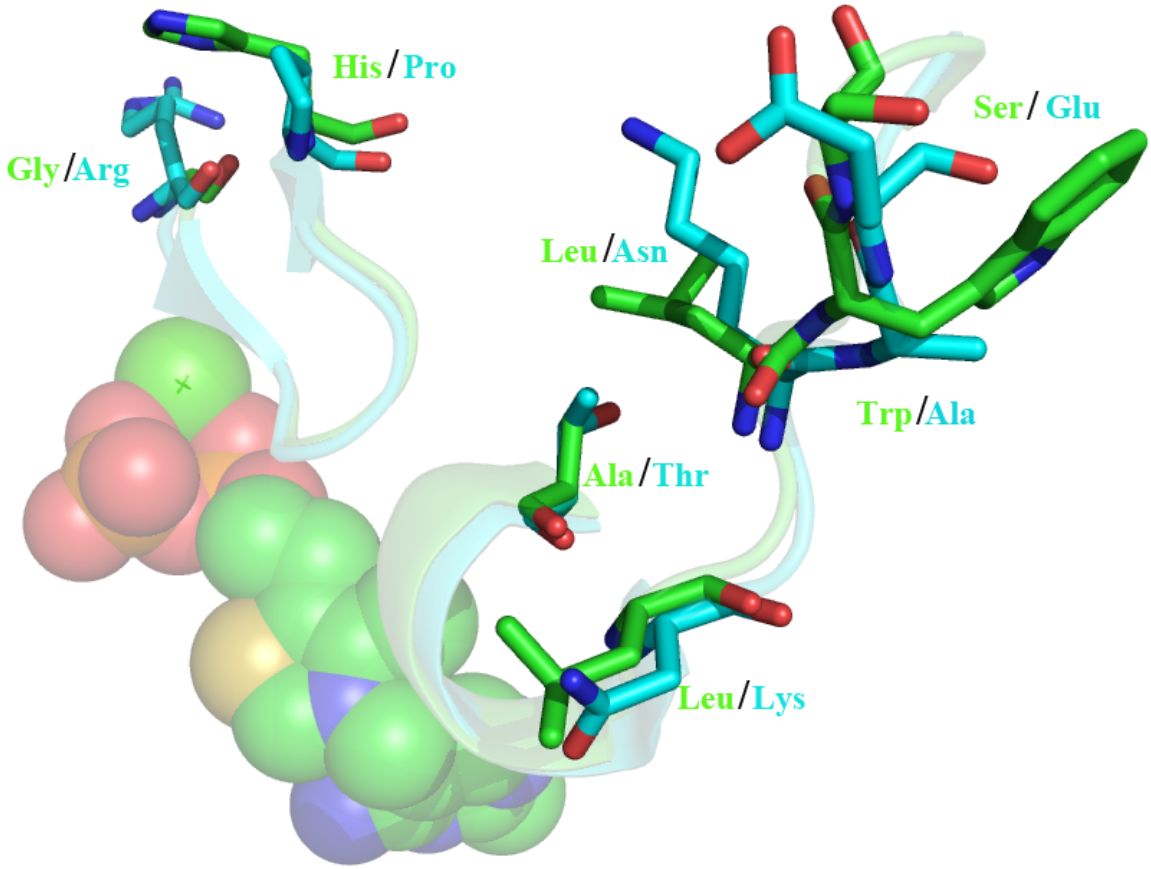


Fig. 2.

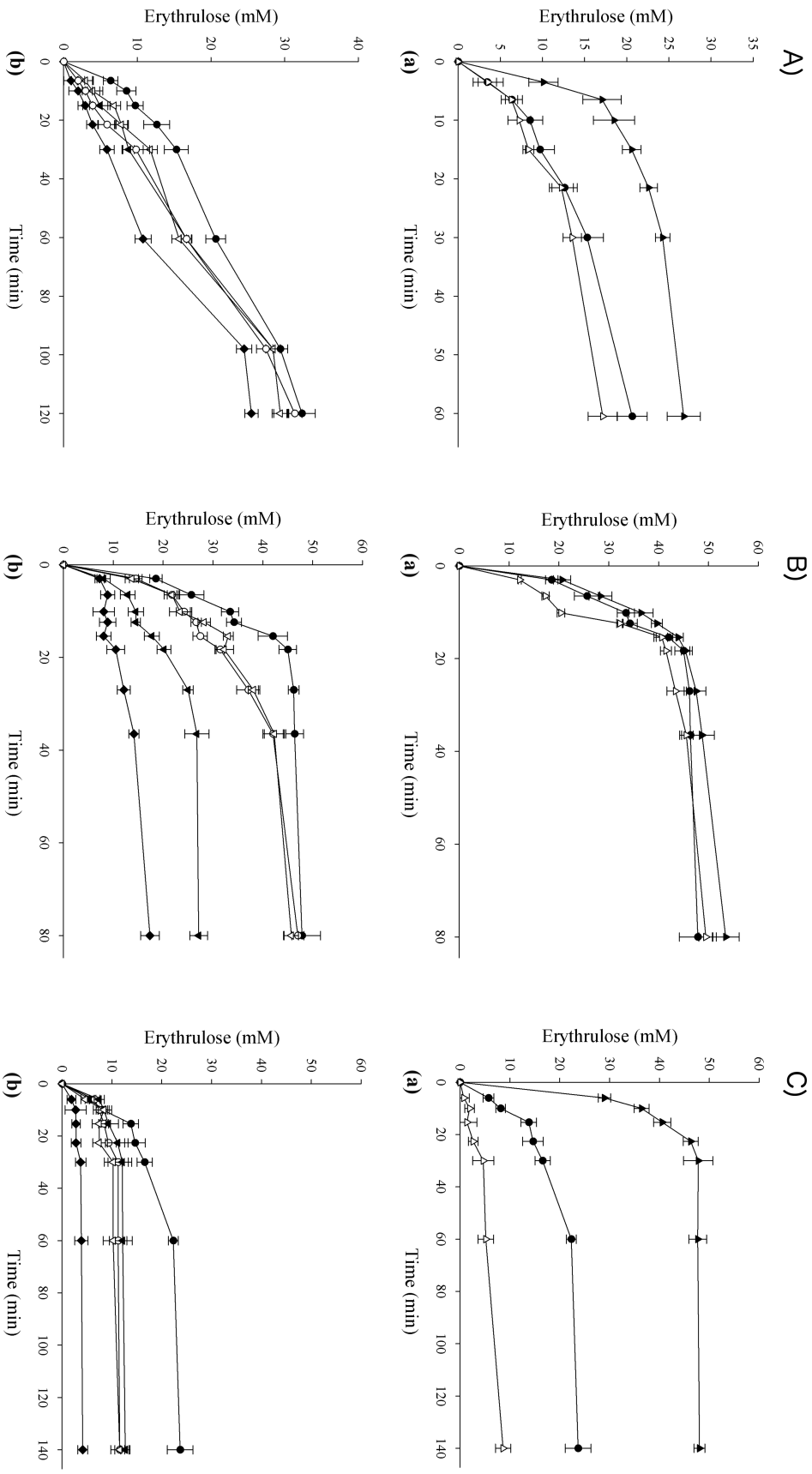


Fig. 3.

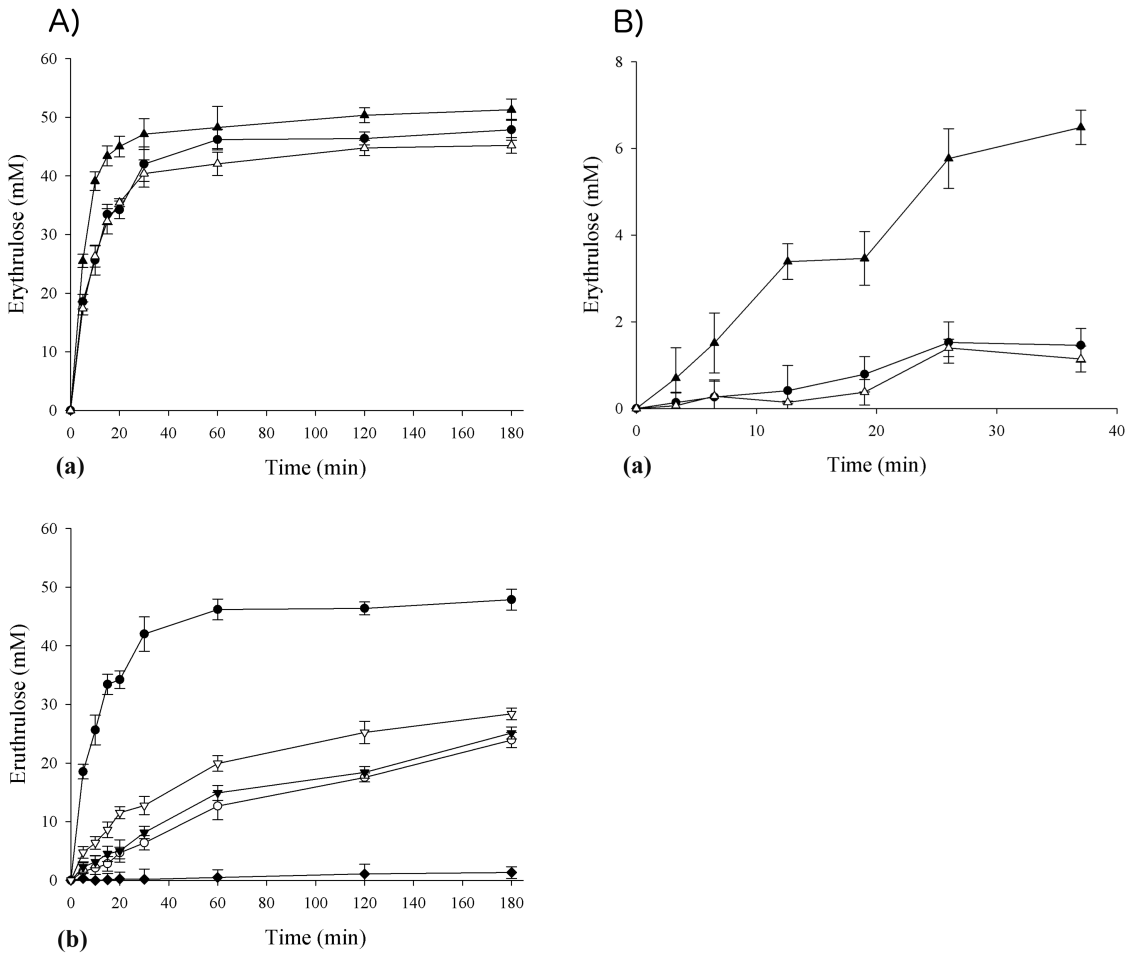

A)
Erythrulose (mM)
Time (min)
(a)
B)
Erythrulose (mM)
Time (min)
(a)
Eruthrulose (mM)
Time (min)
(b)

Fig. 4.

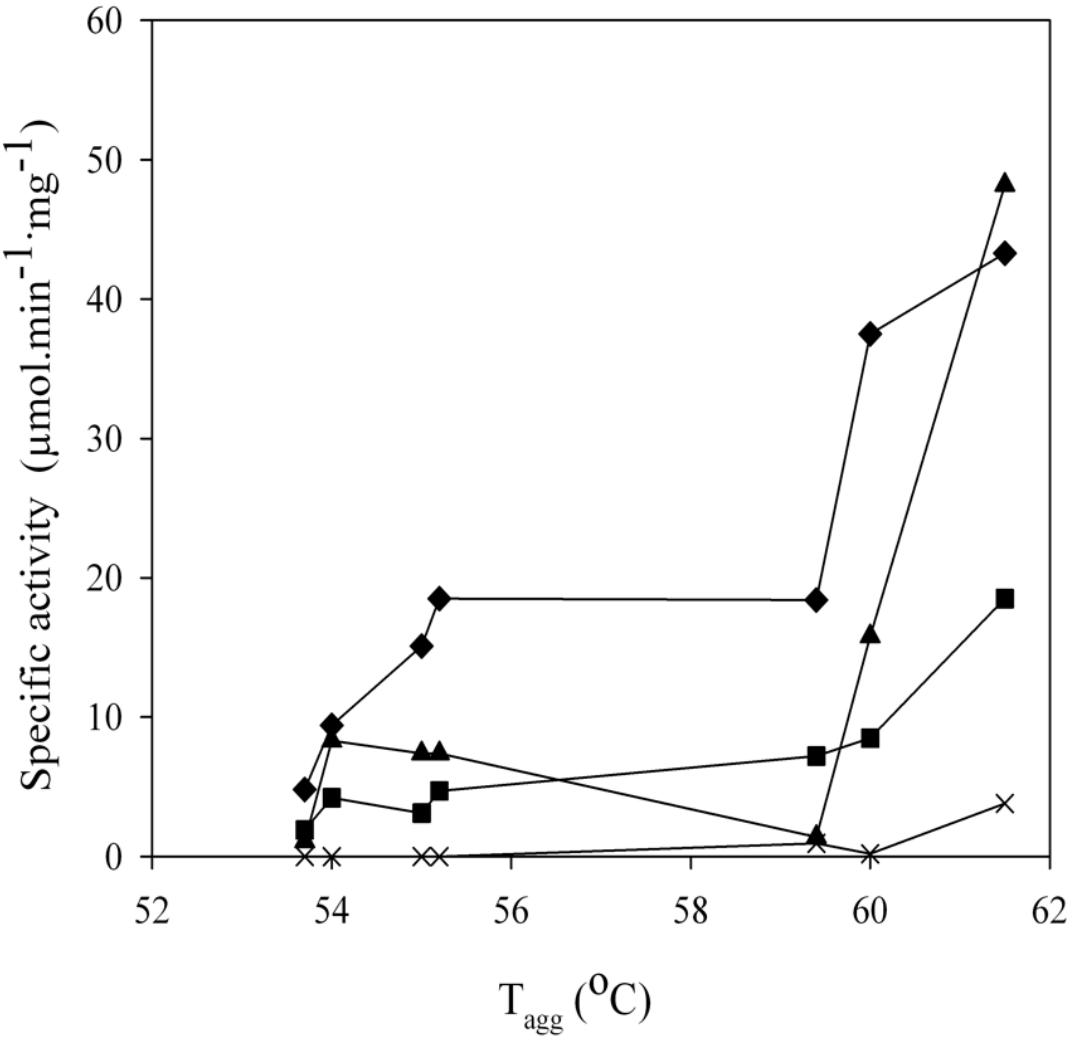
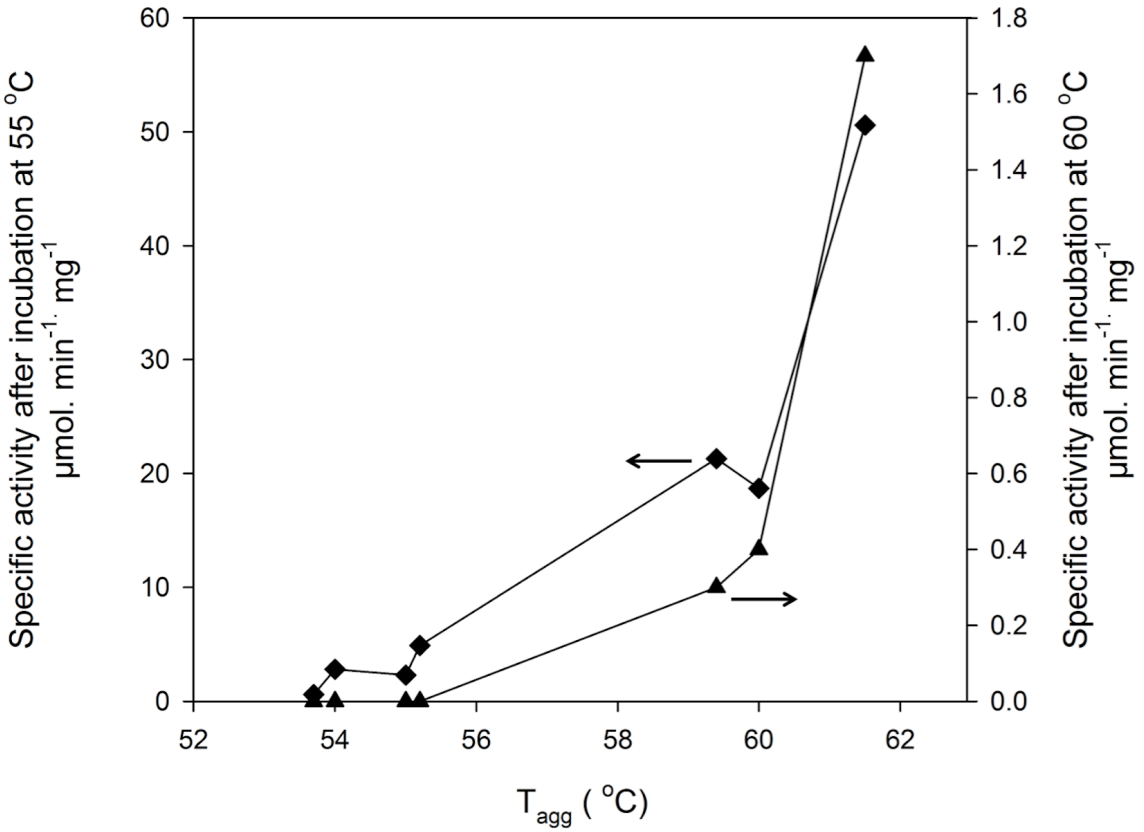

Specific activity (µmol.min-1.mg-1)
60
50
40
30
20
10
0
52
54
56
58
60
62
Tagg (oC)

Fig. 5.


Specific activity after incubation at 55 °C
µmol. min-1. mg-1
Specific activity after incubation at 60 °C
µmol. min-1. mg-1
Tagg (°C)
0
10
20
30
40
50
60
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
1.8
52
54
56
58
60
62