Structure and function of L-threonine-3-dehydrogenase from the parasitic protozoan *Trypanosoma brucei* revealed by X-ray crystallography and geometric simulations

Eyram Adjogatse*[a], Peter Erskine[a], Stephen Wells[b], John M. Kelly[c], Jonathan D Wilden[d], AW Edith Chan[e], David Selwood[e], Alun Coker[a], Steve Wood[a], Jonathan B Cooper[a]

*Corresponding Author (rmhaead@ucl.ac.uk)

[a]Laboratory for Protein Crystallography, Wolfson Institute for Biomedical Research, University College London, London, WC1E 6BT, United Kingdom

[b]Department of Chemistry, University of Bath, Bath, BA2 7AY, United Kingdom

[c]Department of Pathogen Molecular Biology, London School of Hygiene & Tropical Medicine, London, WC1E 7HT, United Kingdom

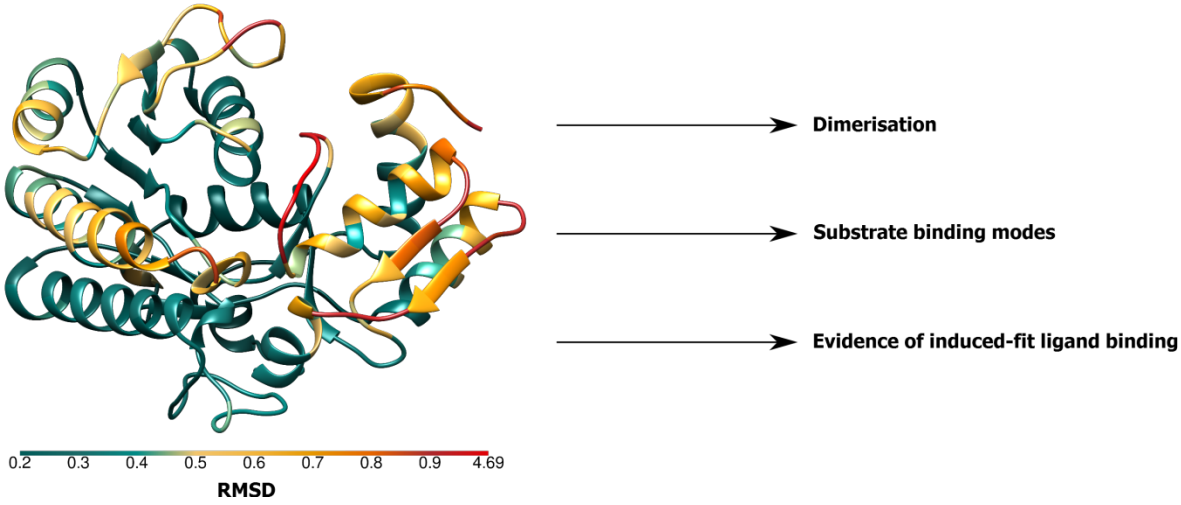
[d]Department of Chemistry, University College London, London, WC1H OAJ, United Kingdom

[e]Drug Discovery, Wolfson Institute for Biomedical Research, University College London, London, WC1E 6BT, United Kingdom

Abstract

Two of the world's most neglected tropical diseases, human African trypanosomiasis (HAT) and Chagas Disease, are caused by protozoan parasites of the genus *Trypanosoma.* These organisms possess specialised metabolic pathways, frequently distinct from those in humans, which have potential to be exploited as novel drug targets. This study elucidates the structure and function of L-threonine-3-dehydrogenase (TDH) from *T. brucei*, the causative pathogen of HAT. TDH is a key enzyme in the metabolism of L-threonine, and a TDH inhibitor has been shown to have trypanocidal activity. TDH in humans is a non-functional pseudogene, suggesting that it may be possible to rationally design safe and specific therapies for trypanosomiasis by targeting this parasite enzyme. As an initial step, the TDH gene from *T. brucei* has been expressed, and the three-dimensional structure of the enzyme has been solved for the first time by X-ray crystallography. In multiple crystallographic structures, *T. brucei* TDH is revealed to be a dimeric short-chain dehydrogenase that displays a considerable degree of conformational variation in its ligand binding regions. Geometric simulations of the structure have provided insight into the dynamic behaviour of this enzyme. Furthermore, the structures of TDH bound to its natural substrates and known-inhibitors have been determined, giving an indication of the mechanism of catalysis of the enzyme. Collectively, these results provide vital details for future drug design to target TDH or related enzymes.

Graphical Abstract



Keywords

TDH; short-chain dehydrogenase; threonine metabolism; trypanosomiasis; geometric simulations

Abbreviations

ADH, alcohol dehydrogenase; AKB, 2-amino-3-ketobutyrate; Ac-CoA, acetyl-coenzyme A; BSF, bloodstream form; DLS, Diamond Light Source; ESRF, European Synchrotron Radiation Facility; GalE, UDP-galactose 4'-epimerase; HAT, human African trypanosomiasis; KBL, 2-amino-3-ketobutyrate; NTD, neglected tropical disease; PDB, protein data bank; PEG, polyethylene glycol; SEC, size-exclusion chromatography; TDH, L-threonine-3-dehydrogenase.

Introduction

Trypanosomiasis is a human and animal disease caused by infection with insect-transmitted parasites of the genus *Trypanosoma*, protozoa distinguished by a single flagellum. In humans, *Trypanosoma brucei* causes human African trypanosomiasis (HAT), which is endemic in sub-Saharan Africa, and *Trypanosoma cruzi* causes Chagas disease, which is widespread in South America. HAT is fatal if left untreated, whilst Chagas disease often results in chronic and life-threatening pathology many years after the initial infection[1,2].

Both HAT and Chagas disease have been designated as Neglected Tropical Diseases (NTDs). These debilitating infections often affect poorer populations within developing countries, and have historically received less attention for pharmaceutical R&D. As a result, there are a limited number of drugs available for both diseases, and these have numerous shortcomings, which include toxicity, limited efficacy, and acquisition of resistance by the pathogens. Developing a more detailed understanding of *Trypanosoma* biochemistry and metabolism should aid the identification of new drug targets.

The main hypothesis of this project is that a genetic difference between trypanosomes and their human hosts is exploitable for drug design. It is known that the insect gut form of *T. brucei* selectively and rapidly metabolises L-threonine by the action of L-threonine 3-dehydrogenase (TDH). Indeed, inhibitors of TDH are lethal to the parasite[3,4]. More recent work has shown that *TDH* is a pseudogene and non-functional in humans, and that an alternative metabolic pathway exists [5]. This suggests that inhibitors of trypanosome TDH may have potential as highly specific anti-parasitic agents. Recent work using reverse genetics has shown that glucose and threonine contribute almost equally to Ac-CoA production, both being crucial in bloodstream form (BSF) and procyclic form parasites, with blockade of the corresponding pathways inducing cell death[6,7]. Hence we

expressed and crystallised *T. brucei* TDH, and derived its crystal structure at high resolution with various physiological ligands bound.

There are three known pathways of threonine catabolism, which are initiated by L-threonine aldolase (EC 4.1.2.5), L-threonine dehydratase (also known as L-threonine deaminase; EC 4.2.1.16) and L-threonine 3-dehydrogenase (TDH; EC 1.1.1.103)., which is the dominant pathway in several prokaryotes and eukaryotes[5,8,9]. The activity of L-threonine aldolase has been shown to be low or insignificant in prokaryotes[10] and eukaryotes[11] and no expression of the enzyme has been detected in humans[12]. In contrast, the pathway initiated by L-threonine dehydratase is a major mechanism for L-threonine catabolism and can become the dominant pathway in certain metabolic states in animals[11]. The other major route for L-threonine catabolism is initiated by the mitochondrial matrix enzyme TDH, and appears to be the dominant pathway in many prokaryotes and eukaryotes[5,8,9]. In this pathway, TDH works in tandem with a second enzyme, 2-amino-3-ketobutyrate ligase (KBL) to catabolise L-threonine to glycine and Ac-CoA (Figure 1).

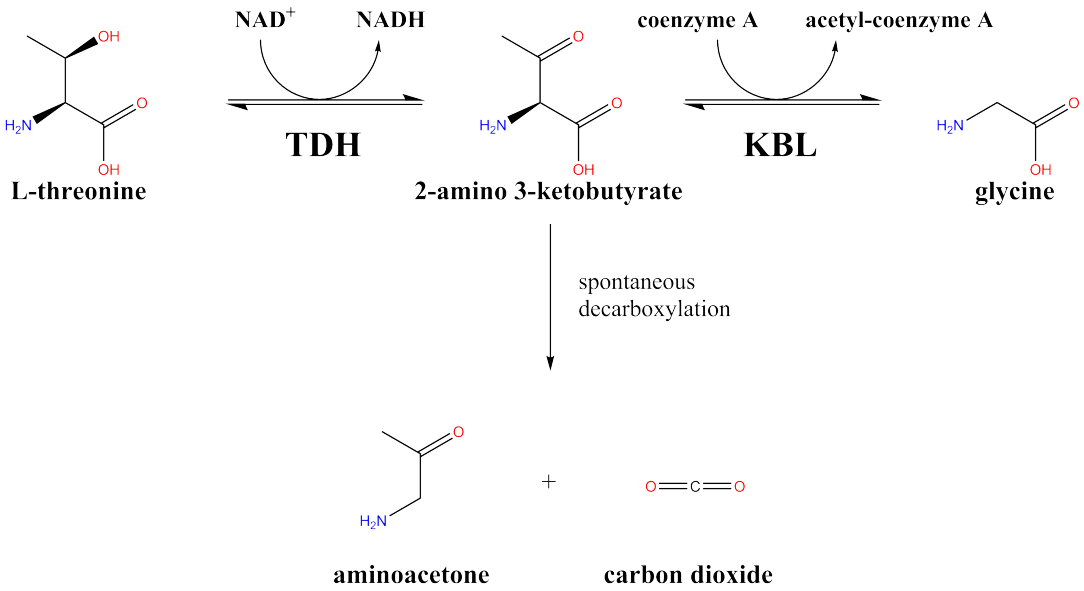


Figure 1.The L-threonine degradation pathway. The enzyme TDH oxidises L-threonine to 2-amino-3-ketobutyrate (AKB). KBL then converts AKB to glycine and Ac-CoA. Alternatively, AKB may undergo spontaneous breakdown to aminoacetone and carbon dioxide.

In the first part of this process, TDH oxidises L-threonine to 2-amino-3-ketobutyrate (AKB) in a reaction dependent on NAD(H). There is evidence that AKB is unstable and spontaneously breaks down to aminoacetone and carbon dioxide. Thus it has been suggested, backed by other evidence, that TDH and KBL form a multi-enzyme complex to facilitate the completion of the second reaction, the breakdown of AKB to Ac-CoA and glycine by KBL[8,13].

TDH belongs to an enzyme family that is sub-divided into short-chain (SDR), medium-chain (MDR) and long-chain (LDR) dehydrogenase/reductases[14]. Several of the most studied TDHs have been MDRs, showing a tetrameric quaternary structure and a requirement for divalent metal cations, such as zinc(8,19–23). More recently, a second group of TDH enzymes that share features with UDP-galactose 4'-epimerase have been characterised. These enzymes belong to the SDR family, and in contrast to the MDR TDH enzymes, they are dimeric and have no requirement for divalent cations(24–26). KBL has not been studied to the same extent, but the crystallographic structure of the enzyme from *Escherichia coli* has been published[23].

The L-threonine degradation pathway catalysed by TDH and KBL plays a variety of roles in both prokaryotes and eukaryotes, including energy production, homeostasis and fatty acid synthesis. The pathway and its constituent enzymes have been studied in *T. brucei*, but crystal structures of the enzymes are yet to be published.

T. brucei was shown by Cross *et al.* to preferentially and exhaustively consume L-threonine when grown in mixed media[3]. Furthermore, it was shown that most of the absorbed threonine was catabolised via the TDH pathway, producing Ac-CoA and glycine. To further underline the importance of this pathway, exposure of cultured *T. brucei* to tetraethyl thiuram disulphide (TETD), an inhibitor of human aldehyde dehydrogenase and a potent inhibitor of TDH, has been demonstrated to lead to trypanosome death in correlation with the degree of TDH inhibition[3,4].

A principal reason for the detrimental effect of TDH inhibition on *T. brucei* is the importance of the pathway for fatty acid synthesis, for which Ac-CoA is a substrate. Millerioux, Mazet and colleagues have demonstrated that Ac-CoA is required for fatty acid synthesis by simultaneously inhibiting TDH and another key Ac-CoA-producing enzyme, pyruvate dehydrogenase, which proved to be lethal[6,7]. The physiological role of L-threonine metabolism through the TDH pathway is not limited to fatty acid production and there may be additional functions not yet elucidated. For instance, glycine produced from L-threonine by TDH in *T. brucei* has been shown to be incorporated into trypanothione[6], an antioxidant analogous to glutathione found in humans. In addition, studies of ^{13}C-labelled glucose metabolism in BSF *T. brucei* showed that glucose-derived acetate was used to acetylate amino acids, and it was suggested that L-threonine-derived acetate plays a role in this process [24].

Below, we present the first complete description of the tertiary and quaternary structure of *Tb*TDH, as analysed by X-ray crystallography and a range of other biochemical and computational techniques. The results give insight into the relationship between TDH and its natural cofactor and substrate, and also the relationship between TDH and the subsequent enzyme in the pathway, KBL.

Results

Structure of TDH

Crystal growth and Data Collection

Diffracting TDH crystals were successfully grown under a number of different conditions, which are listed in Table 1, along with data collection and refinement statistics for the related data sets. Diffracting TDH crystals were cryo-cooled using glycerol, PEG 400 or ethylene glycol. It was also possible to collect quality data from crystals cooled in the absence of cryoprotectant, as exemplified by the data used to solve two structures (PDB ID: **5K4W** and PDB ID: **5K4Y**).

Table 1. Details of data collected and structures solved by X-ray crystallography

	Model								
RCSB PDB ID	**5L9A**	**5LC1**	**5K4Q**	**5K4T**	**5K4V**	**5K4U**	**5K50**	**5K4W**	**5K4Y**
Crystallisation Conditions									
Conditions	0.1 M HEPES; 20 % w/v PEG 10K; pH 7.5; TDH (5.28 x 10^{-2}mM [2.0 mg/ml])	0.1 M HEPES; 20 % w/v PEG 10K; pH 7.5; TDH (5.28 x 10^{-2}mM [2.0 mg/ml]); NAD+ (1mM); pyruvate (30mM)	0.2M lithium sulphate; 0.1M Tris; 30% w/v PEG 4K; pH 8.5; TDH (6.4 x 10^{-2}mM [2.4mg/ml]); KBL (3.8 x 10^{-2}mM [1.8mg/ml]); NAD^+ (0.09mM); PLP (0.18mM); L-Threonine (7.33mM)	0.1M MES; 0.2M ammonium sulphate; 30% w/v PEG 5K MME; pH 6.5; TDH (9.0 x 10^{-2}mM [3.4mg/ml]); L-serine (50mM)	0.1M tri-sodium citrate; 30% w/v PEG 4K; 0.2M ammonium acetate; pH 8.5; TDH (3.7 x 10^{-2}mM [1.4mg/ml]); NAD^+ (10mM); BPOB (0.5mM)	0.1M tri-sodium citrate; 30% w/v PEG 4K; 0.2M ammonium acetate; pH 5.6; TDH (3.7 x 10^{-2}mM [1.4mg/ml]); NAD^+ (8.5mM); quinine (2.5mM [hydrochloride dehydrate])	0.1M tri-sodium citrate; 30% w/v PEG 4K; 0.2M ammonium acetate; pH 5.6; TDH (10.4 x 10^{-2}mM [3.9mg/ml]); NAD^+ (10mM); L-*allo*-threonine (30mM)	0.2M sodium acetate; 0.1M Tris; 30% w/v PEG 4K; pH 8.5; TDH (9.3 x 10^{-2}mM [3.5mg/ml]); NADH (10mM); L-Threonine (30mM)	0.1M tri-sodium citrate; 25% w/v PEG 4K; 0.2M ammonium acetate; pH 5.6 TDH (3.7 x 10^{-2}mM [1.4mg/ml]); NAD^+ (10mM); methylglyoxal (8mM)
Cryoprotectant	Glycerol	Glycerol	Glycerol	Glycerol	Glycerol	Glycerol	Glycerol	None	None
Data Collection									
Beamline	ESRF ID23-2	ESRF ID29	ESRF	DLS I04-1	DLS I04-1	DLS	DLS I04-1	DLS I02	DLS I04-1
Wavelength (Å)	0.8726	1.07356	0.9334	0.9173	0.9173	0.9173	0.92	0.9795	0.92001
Space group	P1	$P2_12_12$	$P2_12_12$	$P4_32_12$	$P2_12_12$	$P2_12_12$	$P2_12_12$	$P2_12_12$	$P2_12_12$
Unit cell parameters									
a (Å)	46.98	132.03	132.04	91.76	90.40	90.07	133.03	83.45	133.45
b (Å)	57.72	276.49	276.49	91.76	131.53	133.10	273.06	136.12	278.63
c (Å)	70.07	55.74	55.74	93.60	55.02	55.61	55.80	55.69	56.27
α (°)	72.46	90	90	90	90	90	90	90	90
β (°)	70.38	90	90	90	90	90	90	90	90
γ (°)	73.2	90	90	90	90	90	90	90	90

	Model								
RCSB PDB ID	**5L9A**	**5LC1**	**5K4Q**	**5K4T**	**5K4V**	**5K4U**	**5K50**	**5K4W**	**5K4Y**
Max. resolution (Å)	1.45	2.1	2.3	2.1	2.2	1.9	2.26	1.72	1.77
No. of reflections	409937	874324	624537	301645	109277	203173	297998	435135	438691
No. of unique reflections (used)	109054	117204	93173	23982	32475	61881	91289	67909	193067
Solvent content (%)	44.2	45.05	46.34	52.83	42.88	43.94	44.98	41.24	46.7
Matthews coefficient ($Å^3.Da^{-1}$)	2.2	2.24	2.29	2.61	2.15	2.19	2.23	2.09	2.31
Resolution range (Å)	53.82 - 1.45	95.48 - 2.10	48.143 - 2.3	29.073 - 2.1	28.226 - 2.2	28.552 - 1.9	31.09 - 2.26	46.32-1.72	39.60 - 1.77
Multiplicity	3.8	7.5	6.7	12.6	3.4	3.3	1.9	6.4	6.2
Completeness (%)	95	97.3	99.188	99.98	95	97.7	90.4	99.5	99.6
Mean I/σ	10.4	11.4	11.2	33.6	13.4	12.7	5.6	11	12.6
R_{merge}	0.089	0.145	0.12	0.054	0.067	0.074	0.151	0.094	0.121
R_{pim}	0.051	0.055	0.049	0.016	0.04	0.048	0.181	0.04	0.055
R_{meas}	0.103	0.156	0.13	0.056	0.079	0.088	0.214	0.102	0.133
Refinement									
R_{work}	0.1649	0.181	0.1746	0.1741	0.1643	0.1864	0.2073	0.1522	0.16071
R_{free}	0.2218	0.2873	0.2618	0.2400	0.2411	0.2321	0.2809	0.1976	0.20074
RMS bond lengths (Å)	0.011	0.008	0.0178	0.0245	0.0201	0.0241	0.0128	0.0198	0.0201
RMS bond angles (°)	2.198	1.837	1.764	2.059	1.826	1.99	1.6492	2.1037	1.9749
Average B-factor ($Å^2$)	18.68	22.2	21.09	26.769	26.234	22.138	33.76	27.689	19.218
Ramachandran favoured, % of total	696, 97.6%	1872, 97.0%	1830, 96.62%	309, 96.87%	613, 96.69%	618, 97.48%	1801, 95.14%	622, 97.65%	1824, 97.91%
Ramachandran allowed, % of total	711, 99.7%	1923, 99.6%	61, 3.22%	8, 2.51%	18, 2.84%	14, 2.21%	82, 4.33%	13, 2.04%	33, 1.77%
Ramachandran outliers, % of total	2, 0.3%	7, 0.4%	3, 0.16%	2, 0.63%	3, 0.47%	2, 0.32%	10, 0.53%	2, 0.31%	6, 0.32%
Molecules in the asymmetric unit	2	6	6	1	2	2	6	2	6
Quaternary structure	Dimer	Dimer	Dimer	Dimer	Dimer	Dimer	Dimer	Dimer	Dimer
Residue range	His1-Leu321	His1-Leu321	Pro3-Leu321	His1-Leu321	Pro3-Leu321	Pro3-Leu321	Pro3-Leu321	His1-Leu321	Met2-Leu321
No. protein atoms per monomer	2621	2511	2493	2532	2493	2493	2500.5	2507	2533
No. water molecules per monomer	651	328	252.3	152	220.5	100.5	73.7	210	241.1

	Model								
RCSB PDB ID	**5L9A**	**5LC1**	**5K4Q**	**5K4T**	**5K4V**	**5K4U**	**5K50**	**5K4W**	**5K4Y**
NAD^+ bound?	No	Yes	Yes	No	Yes	Yes	Yes	Yes (NADH)	Yes
Ligands bound, Mean B-factor	-	Pyruvate, 33.66 NAD, 18.37	-	-	-	-	L-*allo*-threonine, 39.53 L-*allo*-threonine, 26.88	L-threonine, 26.06 L-threonine, 30.58	-
Other solvent molecules/ions	acetate	acetate	glycerol, sulphate	glycerol	glycerol, acetate, Na^+	glycerol, acetate	glycerol, acetate	glycerol, Na^+	glycerol, acetate
Conformation(s) of flexible Loop 1	Open	Closed	Open	Open	Closed	Open and Closed	Open and Closed	Closed	Open

*Tb*TDH primary and secondary structure

The TDH sequences from *T. brucei* strain TREU 927 and *T. brucei gambiense* are essentially identical, confirming that the TDH studied in this work is relevant to a human pathogen. The animal pathogen *T. congolense* has a protein that shares 78% sequence identity with the TDH studied here. The sequence identity between TDH from *T. brucei* strain TREU 927 and the Chagas disease-causing pathogen, *Trypanosoma cruzi*, is 72%. Thus, TDH from a range of *Trypanosoma* species are closely related (Supplementary Figure 1).

Other TDH enzymes with a high sequence similarity to *Tb*TDH include the UDP-galactose 4'-epimerase-like (GalE) TDH enzymes from *Thermoplasma volcanium, Flavobacterium frigidimaris* and *Cupriavidus necator*. The corresponding enzymes from these organisms have monomeric molecular weights between 35 and 37.2kDa, matching that of TDH from *T. brucei* (37.8 kDa). Another group of related TDH enzymes originate from mammalian species: *Gallus gallus* (red junglefowl), *Rattus norvegicus* (rat), *Sus scrofa* (boar/pig) and *Capra hircus* (goat). Once again, TDH enzymes from these species have been studied previously and shown to have molecular weights in the range of 36 kDa[25] to 37 kDa [26].

Inspection of the sequences of the various TDHs shows that they belong to the SDR superfamily of enzymes. Two highly conserved structural features that can be identified from the sequence are highlighted in Supplementary Figure 2. The first, a glycine-rich region, is represented by the sequence GxxGxxG, and the second consists of a tyrosine and a lysine residue separated by three other residues: YxxxK.

X-ray models of *Tb*TDH revealed that each individual subunit consists of two domains displaying α/β topology. The distribution of secondary structure elements in TDH appears to be approximately 38% α-helix and 12% β-sheet, with the remaining 50% of the protein not forming specific folds.

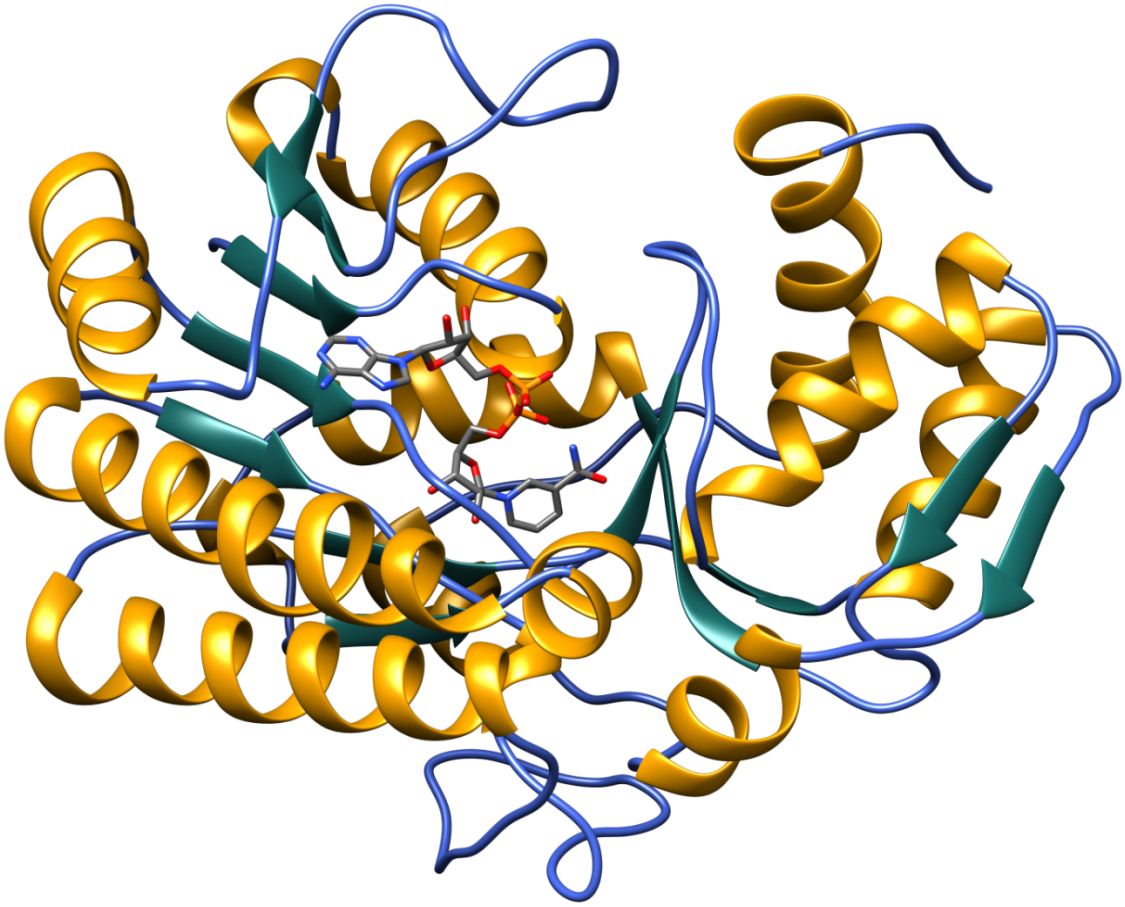


Figure 2.*Tb*TDH coloured by secondary structure (gold = α-helix; dark green = β-strand; light blue = loop). Cartoon representation produced using UCSF Chimera.

*Tb*TDH tertiary structure

*Tb*TDH encompasses an NAD-binding domain, which includes the first 180-200 amino acids, and a catalytic domain at the C-terminal end of the protein (Supplementary Figure 3). In three dimensions, there is some crossover between these domains. There is a Rossmann fold motif (parallel β-strands flanked by α-helices), which is characteristic of NAD-binding proteins, and hence SDRs. The structural motif GxxGxxG, which is common to the 'extended' class of SDRs (eSDR), is located within the NAD binding region and is represented by the sequence Gly9-Ala10-Leu11-Gly12-Gln13-Ile14-Gly15. This motif is involved in binding the diphosphate group of NAD (Figure 6). The YxxxK motif and several other conserved residues, including Met81, Ser82, Thr119, Thr186, Trp280, and Tyr144 and Lys148 of the SDR YxxxK motif are located in the active site (Figure 7).

When comparing the different crystallographic models, which were obtained with different ligands and crystallised in different space groups, no large domain movements were observed. By comparing corresponding residues on different TDH models, it was possible to detect a certain degree of variation in the relative positions of certain atoms (see methods), as indicated by the root-mean-square deviation (RMSD). Furthermore, an analysis of the average *B*-factor for each residue revealed high B-factors for some residues, supporting the notion that those regions of the protein are more disordered than others, and thus more flexible.

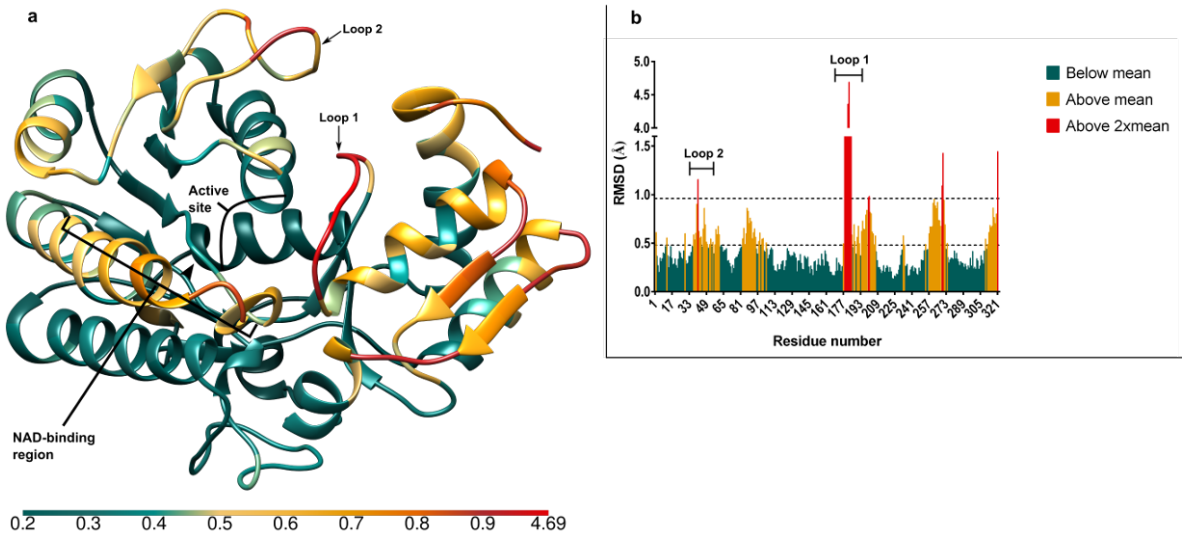


Figure 3. Conformational flexibility as indicated by RMSD of Cα positions of the different *Tb*TDH structures (PDB ID: 5K4Q; PDB ID: 5K4T; PDB ID: 5K4V; PDB ID: 5K4U; PDB ID: 5K50; PDB ID: 5K4W) when superimposed in UCSF Chimera. (a) A ribbon representation of TDH, coloured by the RMSD of α-carbon positions in each residue. The scale at the bottom of the diagram indicates the colour scheme relating to the RMSD measured in Ångstrom. (b) A bar chart showing the RMSD of α-carbons per TDH residue.

A number of regions in *Tb*TDH appear to be flexible (Figure 3). Firstly, there are two disordered loop regions: Loop 1 includes residues Thr179 and Ala185, and Loop 2 lies between residues Asp35 and Asn60. There also appears to be a higher degree of flexibility in the catalytic domain (gold and red-coloured region on the right-hand side of Figure 3a). The most striking evidence of conformational

variability is found in the residues of Loop 1. Indeed, in different crystallographic structures, this loop was found to occupy two distinctive positions: one where the loop was in an 'open' position, lying at a distance of approximately 5 Å from the active site, and another where the loop was 'closed' and positioned adjacent to the active site (Figure 4). Therefore, for each model of a TDH monomer, the structure can be classified as 'open' or 'closed', on the basis of the conformation of Loop 1. Comparing the average *B*-factors in the open and closed structures highlighted changes in flexibility or disorder that accompanied the change in the Loop 1 conformation. For example, Figure 4b shows the average *B*-factors of TDH residues in 'open' structures, whilst Figure 4c shows the average *B*-factors for the 'closed' structures.

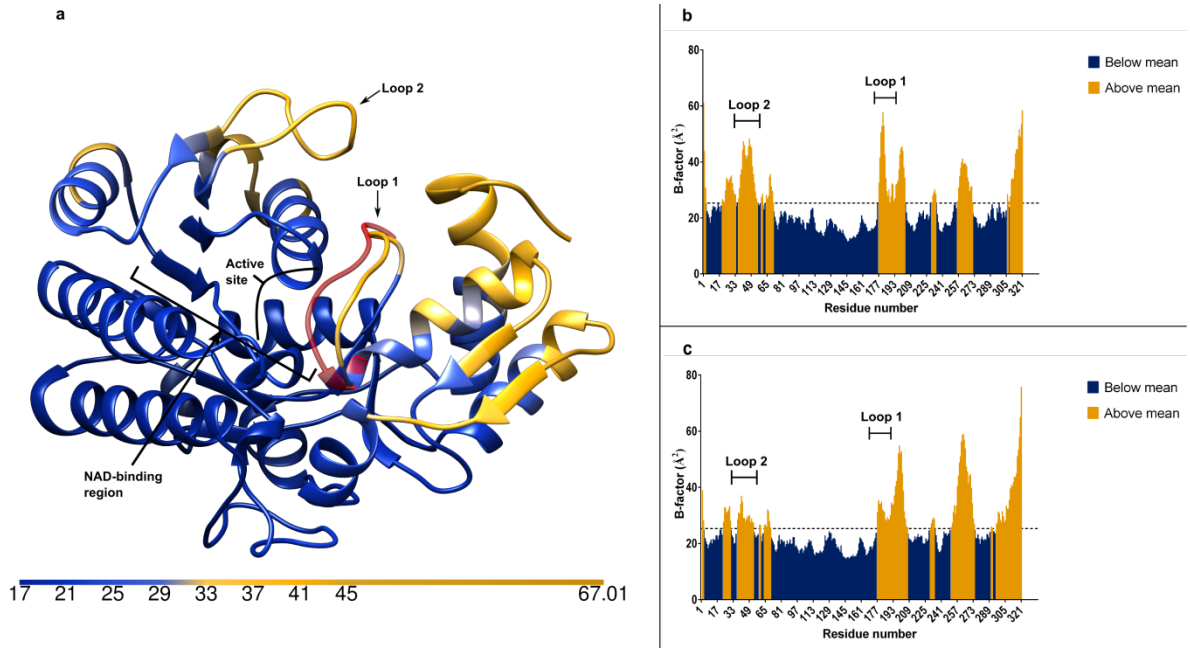


Figure 4. A colour representation of the average *B*-factors of TDH residues, derived from structures PDB ID: 5K4Q, PDB ID: 5K4T, DB ID: 5K4V, PDB ID: 5K4U, PDB ID: 5K50 and PDB ID: 5K4W. (a) A ribbon diagram of TDH coloured by residue-average *B*-factor (scale in horizontal bar). The distinct 'closed' position of Loop 1 is displayed in red. (b) A chart showing the average *B*-factor of residues in 'open' TDH structures. (c) A chart showing average *B*-factors of TDH residues in 'closed' structures

In the 'open' structures, the highest *B*-factors are seen around both loop regions. However, in the closed structures, although these same residues have above average *B*-factors, they are much less

disordered. Inspection of Figure 4c demonstrates that the catalytic domain appears to be more disordered in closed structures. In three dimensions, this region can be seen as one of a pair of adjacent α-helices and β-strands, which are the gold-coloured regions on the right hand sides of the structures in Figure 3a and Figure 4a.

The NAD-binding region, particularly around Loop 2, also appears to be flexible. Although the B-factors are low, the RMSDs around residues 82 to 100 are particularly high (Figure 3b). This region lies at the dimerization interface (discussed below), indicating some conformational variability there.

These findings indicate that there may be a relationship between the conformation of Loop 1, and the flexibility of other regions of the protein, particularly the catalytic domain. The different conformational changes do not appear to correlate with the space-group of the crystal (Table 1). Thus, the specific conformation of Loop 1 and the flexibility of the catalytic region may actually be influenced by ligand binding. Indeed, three of the four 'closed' structures (from PDB ID: **5K4U**, PDB ID: 5K4W and PDB ID: **5K50**) examined were bound to NAD/NADH and another ligand, whilst only one of the four 'open' structures (from PDB ID: **5K50**) examined possessed bound NAD and another ligand.

*Tb*TDH quaternary structure and complex formation

TDH was shown to exist as a dimer following refinement of all the X-ray crystallographic models. The dimerization interface consists of two α-helices (Pro89-Tyr110 and Val143-Tyr162) and a loop (Lys126-Thr142), with the same regions on each monomer making contact in the dimer. As can be seen in Figure 5, the dimer is stabilised by a large number of hydrophobic, polar and ionic interactions. In particular, polar side chains of residues on each monomer form hydrogen bonds with each other and charged residues form salt bridges. Interactions across this interface include the following: Glu86 with Trp157; Asp90 with Arg101; Asp94 with Arg101; Lys126 with Asp135. This

probably gives greater stability to the dimer, whilst the length of some of the side chains allows their interactions to be maintained in the case of any change in the orientation of the two monomers.

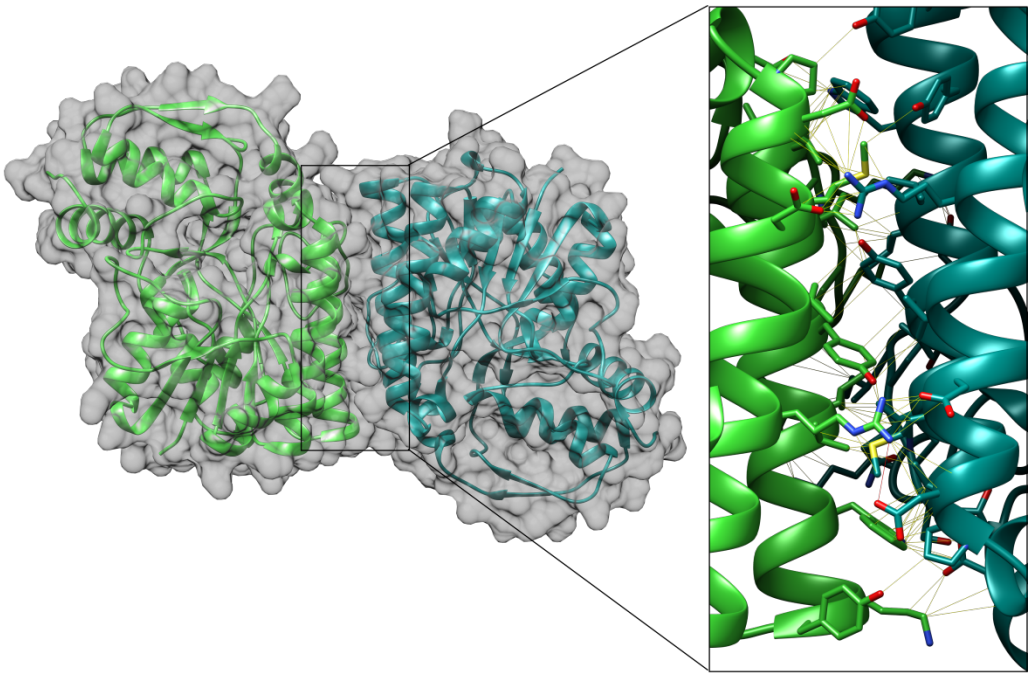


Figure 5. The TDH dimer and its dimerisation interface. Lines between amino acids indicate Van der Waals, polar or ionic interactions. Key residues involved in the interaction include Glu86, Asp90, Asp94, Arg101, Lys126, Asp135 and Trp157

Further evidence of the existence of the dimer and that this is the largest oligomer present in all of the crystallographic structures, was provided using the software PISA[27]. Supplementary Table 1 lists a number of different interfaces between different TDH subunits in the different structural models and it can be seen that the only stable association predicted by PISA is the dimer relationship previously described. There is no clear relationship between the crystal space group and the nature of the dimerisation interface.

Oligomerisation studies

When assayed separately in size-exclusion chromatography experiments (SEC), elution volumes of TDH and KBL indicated that they exist in dimeric forms in solution (Supplementary Table 2). The TDH

results were therefore consistent with the findings gained by X-ray crystallography. Interestingly, when mixed and assayed together, TDH and KBL co-eluted predominantly at a volume suggesting the existence of the proteins as monomers. Miniscule amounts of both proteins eluted at a volume corresponding to a large oligomer of approximately 188 kDa , but the low protein concentrations make this observation inconclusive. Overall, the SEC findings confirm the dimeric structure of TDH. Additionally, an indication of potential interactions between TDH and KBL was obtained, however further study is required to confirm the dissociation of TDH and KBL dimers or the existence of a TDH:KBL complex.

The cross-linking studies carried out using dimethyl suberimidate (DMS) provide more evidence in regards to the oligomeric states of TDH and KBL. SDS-PAGE analysis of the samples revealed protein bands representing monomeric and dimeric forms of TDH and KBL (Supplementary Figures 6-8). A single two-fold dilution of the concentration did not have an effect on the apparent oligomeric states detected by the cross-linking experiments. Experiments conducted with TDH and KBL together do not suggest that the two proteins are complexed in solution, as there are no additional bands present (Supplementary Figure 6, lanes 7 and 8), compared to when samples of the two proteins were tested separately (Supplementary Figure 6, lanes 1, 2, 4 and 5). The pull-down assay using His-tagged KBL and native TDH was in agreement with the data from the cross-linking studies and the majority of the SEC studies, suggesting that TDH and KBL do not appear to form a complex with each other in solution. When His-tagged KBL was applied to the Ni-NTA column, non-tagged TDH washed through with the wash buffer. Meanwhile, KBL was retained on the column, and elution with the 250mM imidazole elution buffer caused KBL to elute alone, demonstrating that the two proteins had not bound to each other or remained bound to each other under the conditions tested (Supplementary Figure 9).

Structural insights into ligand binding and catalysis

To gain a complete picture of TDH structure as it relates to function, various ligands, including the natural substrates, were co-crystallised with TDH prior to obtaining diffraction data. As can be seen in Figure 6 and Figure 7, the TDH substrates, NAD^+ and L-threonine, occupy a deep cleft in the protein. The L-threonine binding pocket and the pocket occupied by the nicotinamide group of NAD^+ are partially separated by the side chain of Met81 (Figure 7).

NAD binds TDH through a combination of hydrogen bonds and hydrophobic interactions with several main chain and side chain atoms of TDH residues. Whilst the nicotinamide and the adjacent ribose groups bind in deep pockets of the protein, the remainder of the NAD molecule faces the exterior. The interactions with important TDH residues are displayed in Figure 6b. In addition, the cofactor also hydrogen-bonds with several water molecules. This is particularly the case for the phosphate groups, the ribose hydroxyl groups and the nicotinamide amide (Figure 6a).

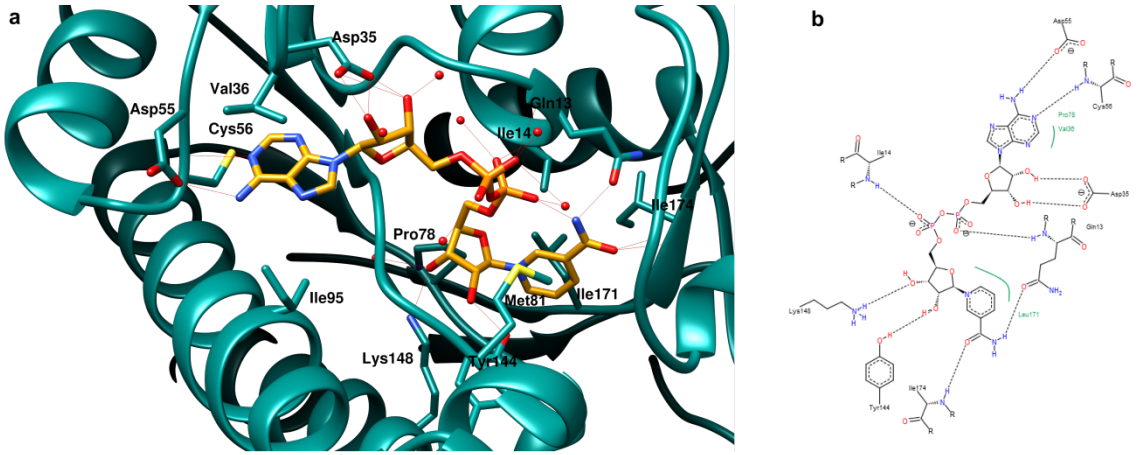


Figure 6. Binding of NAD^+ to TDH in 3D (a) and 2D (b) representations. In (a) hydrogen bonds are indicated by red lines; water molecules are depicted as red spheres. In (b) hydrogen bonds are indicated by dotted lines, whilst Van der Waals interactions are represented by green curved lines, with the corresponding TDH residues in green text.

To gain insight into L-threonine binding, TDH was co-crystallised with the reduced form of the cofactor (NADH) to prevent turnover, which involves oxidation of the amino acid. The L-threonine-bound structure (PDB ID: **5K4W**) at 1.72 Å resolution (Figure 7) reveals that the carboxyl group of L-

threonine hydrogen-bonds with the side chain and main chain atoms of Ser82 and Thr186. The amine group of L-threonine is seen to hydrogen-bond with the side chain hydroxyl group of Thr186, with the nicotinamide carbonyl of NADH and with a nearby water molecule. The side chain hydroxyl group of L-threonine is hydrogen-bonded to Thr119 and Tyr144, whilst the methyl group is directed towards the hydrophobic amino acid Trp280. This aligns the β-carbon with the C4 atom of the nicotinamide ring (discussed in more detail below). Another interesting feature of this structure was that the flexible loop, Loop 1, is closed over the active site, meaning that it is completely shielded by the enzyme (Figure 7c). This was the case in both TDH monomers in the asymmetric unit of the structure, both of which were bound to NADH and L-threonine.

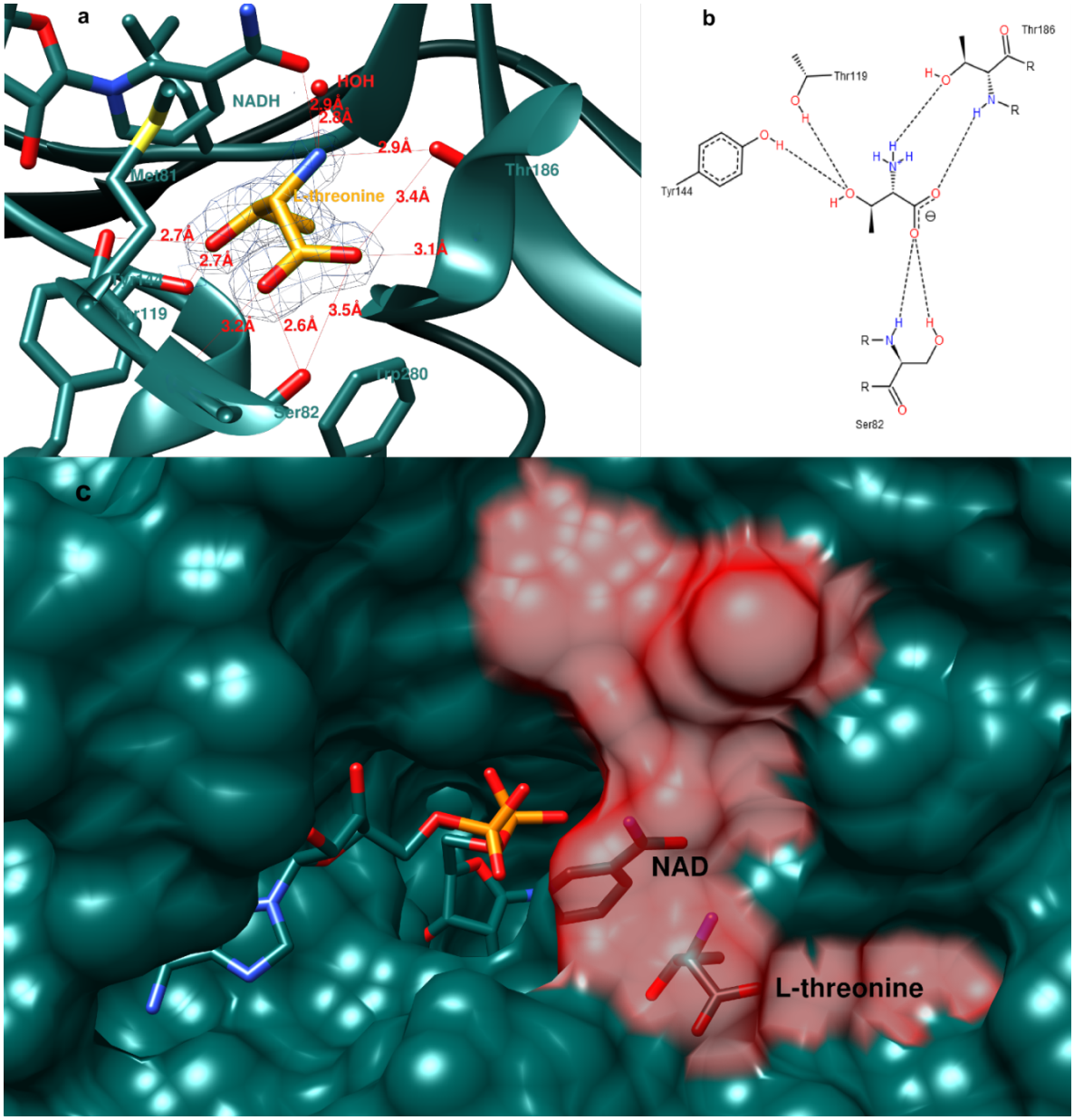


Figure 7. Binding of L-threonine to TDH in 3D (a) and 2D (b) representations, derived from structure 5K5W. In (a) hydrogen bonds are indicated by red lines, electron density of L-threonine is represented by a blue mesh and a water molecule is depicted by a red sphere. In (b) hydrogen bonds are indicated by dotted lines, whilst Van der Waals interactions are represented by green curved lines, with the corresponding TDH residues in green text. (c) shows a surface representation of threonine-bound TDH derived from PDB ID: 5K4W; the surface of loop 1, which is closed over the active site, is coloured red and is semi-transparent.

X-ray crystallographic structures of TDH bound to compounds reported to be inhibitors, L-*allo*-threonine[28] and pyruvate[21,29], were also successfully solved. These compounds were shown to occupy the L-threonine binding site, forming interactions with some of the same active site residues

as the natural TDH substrate, thereby demonstrating their modes of inhibition (Supplementary Figure 10).

The putative mechanism of reaction of TDH involves transfer of a hydride from the β-carbon of L-threonine to the C4 atom of the nicotinamide ring[20]. The data presented here confirm that the atoms involved in this reaction are well aligned and at a distance that would permit such a reaction to happen (Figure 7). X-ray structures described herein confirmed the presence and interaction of key putative catalytic residues, such as Thr119, Tyr144 and Thr186 (Figure 7). Another structural feature that is conserved in SDRs and deemed important for catalysis is a chain of water molecules that is believed to form a proton relay that transfers the proton extracted from L-threonine to the bulk solution, allowing the enzyme to catalyse another reaction. This water chain requires a certain orientation of residues to bind the water molecules. The feature observed in several SDRs is an interaction between an asparagine side chain and the main chain of another residue. In this case, the interaction occurs between Asn96 and Ile80 (Figure 8a), causes a kink in the α-helix spanning residues Pro89-Lys109 and changes the orientation of the Asn96 carbonyl group. This allows the Asn96 carbonyl group to hydrogen-bond one of the water molecules in the water chain (Figure 8b).

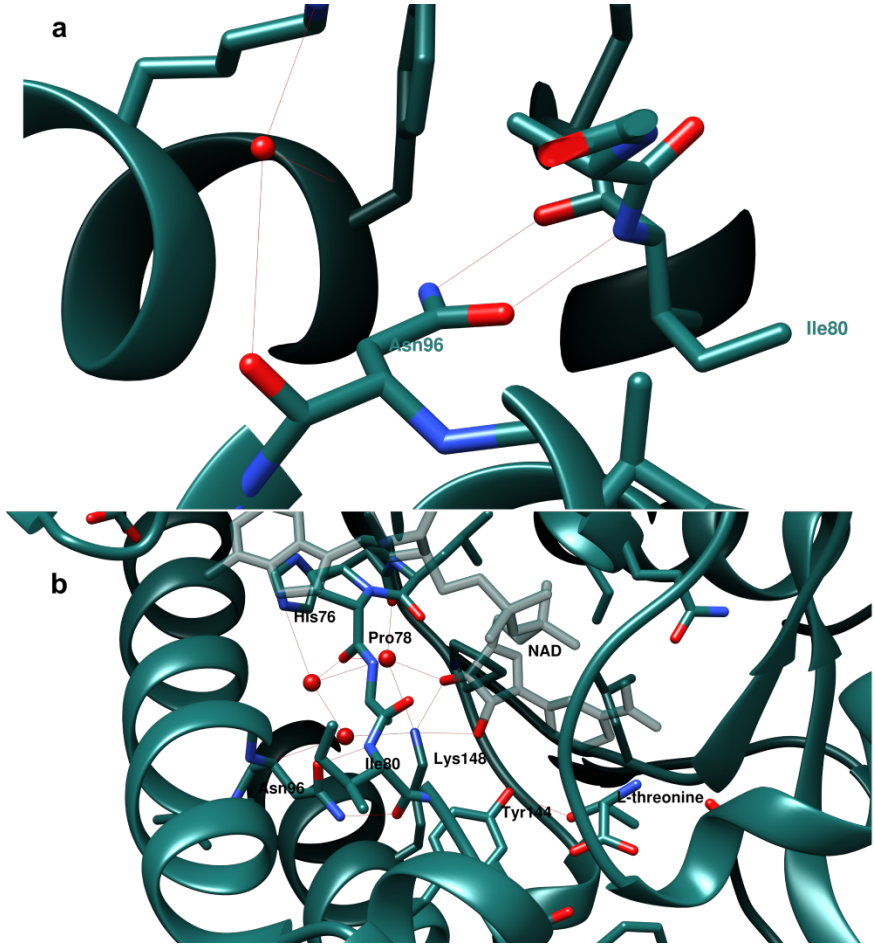


Figure 8. Interaction between Asn96 and Ile80 (a), enabling hydrogen bonding of a water atom in a larger chain of hydrogen-bonded water molecules (b). Hydrogen bonds are indicated by red lines; water molecules are depicted as red spheres.

The sidechain of Lys129 in *Tb*TDH hydrogen-bonds with Asn282 and may have a stabilising role in this region of the enzyme, which is located on the exterior, behind the active site (Supplementary Figure10). Lys129 corresponds to Arg180 in the homologous murine TDH (Supplementary Figure 11). An R180K variant of this enzyme has been shown to have reduced thermostability and impaired recognition of L-threonine , as indicated by reduced catalytic efficiency[30]. Inspection of the structure suggests that replacement of arginine with lysine in *Tb*TDH and bacterial TDH probably results in relatively fewer interactions with surrounding residues, and thus, a weaker stabilising effect.

Other features observed in crystallographic models

A number of molecular species from the storage buffer and from the crystallisation solutions were found to co-crystallise with TDH; glycerol, acetate, sulphate and sodium ions were all observed in TDH structure models. These molecules occupied various positions, and a few positions were occupied in a consistent manner and observed in multiple structures. Most notably, glycerol and acetate bound within the L-threonine binding site (Supplementary Figure 12). Glycerol was also found to bind to the same regions in different TDH models (Supplementary Figure 13).

Typically, different TDH dimers in each of the solved structures are separated from other monomers by an ordered array of water molecules. However, one direct interaction between two monomers in separate dimers was observed in one model (PDB ID: **5K50**). Interestingly, the interaction is formed between residues of the flexible Loop 2 region (Supplementary Figure 14). The interaction does not appear to be strong enough to be an important physiological interaction and is likely to be a result of the crystallisation process.

Geometric Simulations

Noncovalent constraint identification using FIRST

We have examined the dimeric crystal structures using FIRST, which identifies noncovalent polar and hydrophobic interactions and assesses their effect in rigidifying the structure. FIRST identifies a well-formed network of hydrophobic tethers, hydrogen bonds and salt bridges in the dimer interface region, consistent with the PISA analysis. These interactions do not, however, make the dimer into a rigid unit; instead, flexibility remains, consistent with the observed differences between the *apo, holo*, and threonine-bound crystal structures. The Loop 1 region appears not to be constrained by any noncovalent interactions; this is consistent with the crystallographic observations of flexibility in this region.

Low-frequency flexible motion identified using Elnemo and FIRST/FRODA

We have identified low-frequency mode eigenvectors using coarse-grained elastic networking with the Elnemo software, and we have carried out all-atom simulations of flexible motion using the FRODA geometric simulation software (incorporated in FIRST) to explore motion biased along these low-frequency mode directions. The most intriguing large-scale motion we identify is that along the lowest-frequency nontrivial mode of motion (Elnemo's mode 7: modes 1-6 are trivial rigid body motions). This motion is predominantly a twisting motion about an axis running perpendicularly through the dimer interface, i.e. as if the dimer were being twisted about its long axis. The motion is easily permitted by the constraints in the interface region, with large amplitudes of motion (several Å RMSD) being achieved in the geometric simulations; the long flexible sidechains which form the inter-dimer interactions easily accommodate the motion. The character of this motion is essentially identical for the *apo*, *holo* and L-threonine-bound forms.

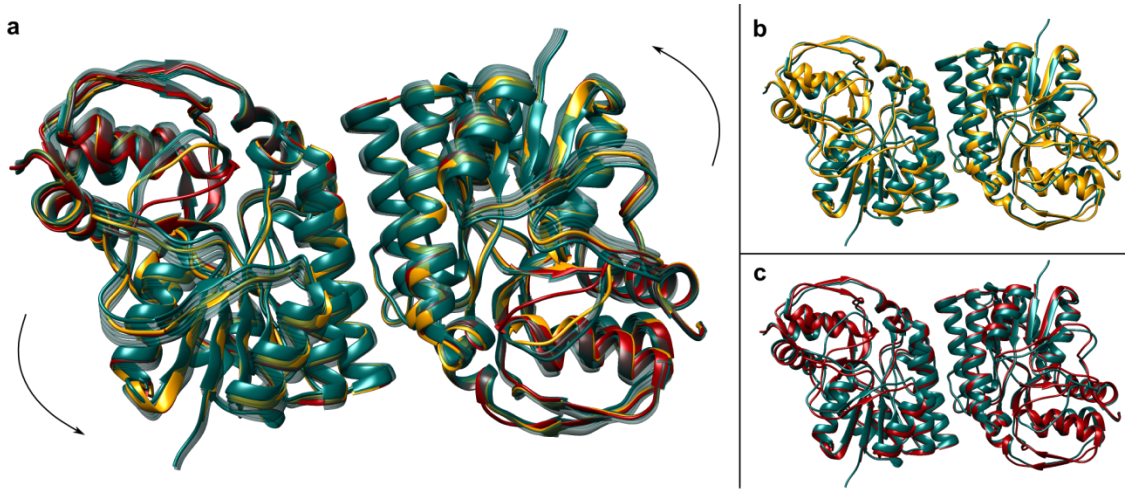


Figure 9. Geometric simulations of *apo-Tb*TDH along Elnemo's mode 7: (a) overlay of multiple projections of *apo-Tb*TDH conformations, superposed on *holo-* (gold) and L-threonine-bound (dark red) *Tb*TDH, with arrows indicating direction of the twisting motion; (b) superposition of *holo-Tb*TDH (gold) over a projection of the *apo* form conformation (fitted backbone RMSD = 0.49 Å); (c) superposition of L-threonine-bound *Tb*TDH (dark red) over a projection of the *apo* form conformation (backbone fitted RMSD = 0.63 Å).

This twisting motion appears to account for most of the global differences observed between the *apo*, *holo*, and L-threonine-bound forms of the dimer. The *holo* and L-threonine-bound forms can be superimposed on the *apo* form with fitted backbone atom RMSDs of 1.13 and 1.582, respectively. However, the geometric simulation of motion of the *apo* form biased along the lowest-frequency nontrivial mode generates a conformation which resembles the *holo* form more closely, with a fitted backbone atom RMSD of 0.49 Å, and, shortly thereafter, a conformation which resembles the L-threonine-bound form, with a fitted backbone atom RMSD of 0.63 Å. These superpositions are shown in Figure 9. Simulations beginning from the *holo* and L-threonine-bound forms (not shown) give consistent results and visit states closely resembling the *apo* form.

Random (diffusive) motion in the flexible Loop 1 region

To assess the freedom of motion in Loop 1, which as noted is not restrained by any identifiable noncovalent interaction, we have carried out unbiased geometric simulations of random diffusive motion in the protein starting from the *apo* state. In one such simulation, the residues making up Loop 1 were observed to drift from their initial "open" state to a much more "closed" state, as shown in in Figure 10. This confirms that the loop has the conformational freedom required to visit both open and closed states regardless of the binding of substrate, and is suggestive of a conformational-selection mechanism rather than a directly induced-fit mechanism of active site closure by Loop 1.

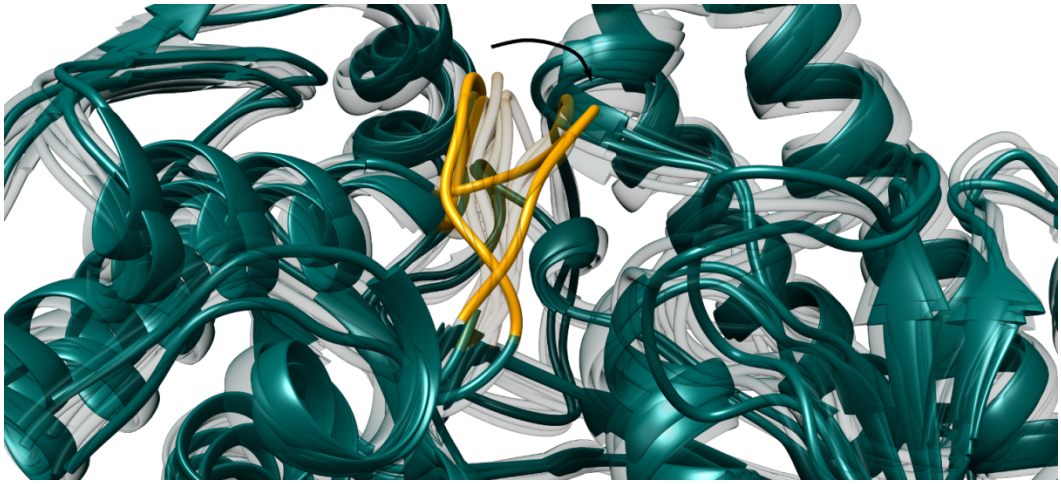


Figure 10. Overlays of multiple states of *apo-Tb*TDH during simulations of random diffusive motion. The arrow indicates the direction of progressive movement of Loop 1 (in gold) during the simulation. States showing extremes of the Loop 1 position are in solid colour, whilst intermediate states are translucent.

Discussion

*Tb*TDH as a GalE-like TDH

*Tb*TDH is a GalE-like TDH, sharing many features with the enzymes from *F. frigidimaris* (*Ff*TDH), *C. necator* (*Cn*TDH) and *T. volcanium* (*Tv*TDH) (20–22,32). We have shown that *Tb*TDH possesses features that are conserved amongst these enzymes, as well as other enzymes within the SDR enzyme superfamily: a monomeric length of around 320 residues, a glycine-rich region (GxxGxxG) within the NAD-binding domain, a structural motif including lysine and tyrosine (YxxxK) near the active site, and several other conserved active site residues (Figure 7). These features contrast with those in TDH in other species, such as *E. coli*, *Pyrococcus horikoshii* and *Thermococcus kodakarensis*, which are homotetramers from a different class of enzyme, rather than homodimers, and which specifically bind divalent metal cations such as zinc[8,17,31].

Another distinguishing feature of GalE-like TDHs is the presence of a loop adjacent to the active site which is directly followed by a motif represented by the sequence GTTDY[21]. In *Tb*TDH, this 'flexible Loop 1' is followed by a similar sequence, GATDY, but it is unclear if this confers any structural or functional difference between these and the other known GalE-like TDHs. Overall, the

structural characteristics observed place TDH in the 'extended' category of SDRs (eSDR), and thus the enzyme is assigned the identifier SDR14E under the SDR nomenclature initiative[32].

*Tb*TDH conformational flexibility

Conformational variability

Figures 3, 4 and 9 indicate the various regions of flexibility and dynamic behaviour that *Tb*TDH may possess in solution. Comparisons of different X-ray crystallographic structures and the geometric simulations show that global motions in the TDH homodimer consist of a combination of hinge and shear motions, where the two subunits appear to twist in opposite directions about the dimer interface. These conformational variations are likely to predominate, as exemplified by the fact that the models arising from various geometric simulations can be superimposed over crystallographic structures of both *apo*- and *holo-TbTDH*. Other important conformational changes may involve alterations in the relative positions of the NAD-binding and catalytic domain.

Recent studies on other dimeric enzymes[33,34] have shown that low-frequency motions can be intimately involved in the opening and closing of enzyme active sites. In the case of *Tb*TDH, however, the lowest-frequency motion gives a good description of the global differences among the *apo/holo*/threonine-bound series of structures, but the motion has no particular effect on the active site geometry. It is therefore plausible that the different degrees of twisting in the different dimer structures represent states from a broad conformational ensemble of flexible variation, selected by crystallisation conditions; the twisting motion may not be directly related to the enzyme function.

The most obvious regions of conformational variability in *Tb*TDH are the flexible Loops 1 and 2. Whereas it is unclear whether Loop 2 has any functional role, it appears that Loop 1 may have an important role in substrate binding. This will be discussed in more detail below.

Substrate binding and catalytic Mechanism of Action

The NADH- and L-threonine-bound structure presented here (Figure 7) is only accompanied in the literature by one other threonine-bound wild-type TDH structure[35]. Both structures provide valuable insights into the proposed mechanism of action of TDH, in which a hydride is transferred from the β-carbon of L-threonine to the C4 atom of the nicotinamide ring of NAD^+; both L-threonine and the co-factor are bound in the necessary orientation for catalysis. Furthermore, the nicotinamide ring was bound in a *syn* conformation, meaning that the hydride is transferred to the pro-S position on the C4 carbon.

The residues of the "catalytic triad"[36,37], which have been described in alcohol dehydrogenases (ADH) and SDRs in general, have been identified as Thr119, Tyr144 and Lys148 in TDH. The residues Thr119 and Tyr144 appear to hydrogen-bond the side chain hydroxyl group of L-threonine. The interatomic distances between the oxygen atoms of the relevant groups on L-threonine and Tyr144 also seem to support the hypothesis that Tyr144 acts as a base and withdraws a proton from the substrate. Molecular dynamics simulations in *Drosophila* ADH have demonstrated that this is made possible by the fact that the protonation states of the active site tyrosine and lysine residues are coupled(36–38). Extrapolating this theory to the TDH mechanism of action, the lysine residue (Lys148 in *Tb*TDH) extracts a proton from a hydroxyl group on a NAD ribose group, which in turn extracts the proton from the tyrosine residue, thus allowing it to extract the proton from L-threonine. In ADH, this proton is then proposed to be transferred to a chain of eight water molecules, which relay the proton until it is removed from the interior of the enzyme. In this way, the tyrosine residue is deprotonated again and is able to participate in the catalysis of another reaction[38,39]. The interatomic distances between the relevant groups observed in TDH structure models all seem to lend support to this theory. The pH dependence of the protonation state of lysine and tyrosine may also explain why the optimum pH in *Tb*TDH(unpublished research) and other GalE-like TDHs is around pH 8-9[4,21,22,29,40–43]. A short chain of water molecules close to Lys148

was identified in several of the *Tb*TDH structures obtained here, but a complete chain of eight water molecules, which has been shown to be essential for ADH activity[39], was not observed. It is possible that a shorter water chain is sufficient for activity in TDH. Alternatively, it may be that the conformational state of TDH in which this water chain is complete has not been captured in any of the crystallographic models to date.

Loops analogous to the flexible Loop 1 identified in *Tb*TDH have been highlighted in TDH from other species[20,35]. Indeed, the amino acid sequence corresponding to the loop appears to be a conserved feature among GalE-like TDHs. The studies here provide crystallographic evidence of distinct conformations adopted by Loop 1 which are characterised by a dramatic shift of approximately 5 Å in distance. In the X-ray crystallographic structures of *Tb*TDH, Loop 1 was observed in the 'open' conformation and the 'closed' conformation in a number of contexts, including in cofactor-bound structures, and structures bound to a cofactor plus L-threonine or another ligand.

Although the data are not conclusive, one can hypothesise that Loop 1 can be in an open or closed conformation, regardless of whether a substrate is bound in the L-threonine binding site or not. It is possible that a change to a closed conformation is induced by binding of a ligand to the L-threonine binding site, as was observed with L-threonine- and pyruvate-bound structures. However, one structure (PDB ID: **5K50**) shows that binding of L-*allo*-threonine did not induce a closed conformation. The observation that different Loop 1 conformations can be adopted in different subunits of the same TDH dimer (as observed in PDB ID: **5K4U** [Supplementary Figure 15]) suggests that the Loop 1 conformation in each monomer is adopted independently of the conformation in the other subunit. The random walk geometric simulations of *Tb*TDH indicate that the loop may have some conformational freedom due to a lack of restraints on this motif. Alternatively, there may be a mechanism whereby the conformation of one subunit affects the conformation in the other subunit.

He *et al.* proposed that the residue Arg180 in *Mm*TDH, which corresponds to Lys129 in *Tb*TDH, may act as a remote "switch" between the open and closed forms of TDH[30]. Although these residues are not located in regions of conformational variability (Figure 3 and Figure 4), Lys 129 does interact with Asn282, which is in close proximity to the flexible region at residues 255-275. Further investigation will be required to understand if Lys129 plays a role in the stability and catalysis of TDH from *T. brucei*.

Considering the effect that it has on the active site of TDH, it is possible that Loop 1 plays an important role in the mechanism of TDH binding to its ligands. The loop may remain open before and after NAD^+ binding to accommodate L-threonine, but then close over the active site once L-threonine is bound. Recently, data from studies of TDH from *C. necator* have also demonstrated that this loop region adopts different conformations in 'open' and 'closed' structures of TDH. Using a combination of crystallographic data from *apo* and *holo* forms of *Cn*TDH, molecular dynamics simulations and enzyme kinetics were used to build a hypothesis of ligand binding and structural changes in this enzyme. Nakano *et al.* proposed a model whereby NAD^+ binding caused a rigidification of flexible regions in the NAD-binding domain, but the flexible loop remained open. Then, in the same model L-threonine binding causes Loop 1 to close over the active site, and this accounts for the selectivity of the enzyme for L-threonine[35].

These findings are in some agreement with those presented herein, which suggest that changes in the conformation of Loop 1 and the catalytic domain of *Tb*TDH happen in response to L-threonine binding. The findings of Nakano *et al.* also support the existence of an induced fit mechanism of ligand binding for TDH, as Loop 1 seems to form a "lid" over the L-threonine active site, something that would prevent L-threonine binding if it happened prior to its encounter with the enzyme. Some researchers assert that all enzymes that form a lid over their active sites must act by induced fit mechanisms[44]; the findings on TDH structure would seem to illustrate this well. However, some structures presented herein show that *Tb*TDH can adopt a 'closed' conformation in the absence of L-

threonine (e.g. PDB ID: **5K4V** or PDB ID: **5K4U**; see Table 1). Furthermore, rigidity analysis and geometric simulations of *apo-Tb*TDH suggest that Loop 1 has the conformational freedom to enter the closed state in the absence of ligands. Therefore, in the case of *Tb*TDH, an induced fit mechanism may occur as part of the wider process of conformational selection, which has been proposed as a way in which proteins with variable conformations can adopt the necessary state required to carry out a function[45]. Although the flexibility of Loop 1 may be linked to the selectivity of TDH for binding and catalysis of L-threonine, the possibility that the enzyme may sample several different conformations in solution may allow a greater variety of ligands to bind, beyond analogues of the natural substrates. This could have significant implications for drug discovery efforts involving TDH as a target.

Interaction between TDH and KBL

The evidence pertaining to the existence of a TDH:KBL heterodimer is inconclusive. Although, under one condition of the SEC experiments, there appeared to be a high molecular weight species that may be attributed to TDH and KBL, the low concentrations of the enzymes in the corresponding fraction make this result uncertain. In addition, X-ray crystallography, cross-linking studies and the pull-down assay all failed to find evidence to confirm a stable association between TDH and KBL. Although it has been suggested that TDH:KBL complex formation would facilitate transfer of the unstable AKB between the two enzymes[8,13], Marcus and Dekker have demonstrated the AKB decarboxylation is pH-dependent and that its half-life could be extended to 11 minutes at pH 7.5, or up to 140 minutes at pH 11.1[46]. This would negate the requirement for TDH and KBL to form a complex. Subjects for future investigation include the influence of ligand binding and catalysis on oligomerisation of TDH and KBL individually, and together in the same environment.

In this study, the structure of *Tb*TDH has been described in detail, providing insights into the dynamic behaviour of the enzyme and how this may relate to its function. The relationship between the enzyme and its natural substrates has been highlighted, and the possible interaction between TDH and KBL has been explored. This new information will contribute to the interpretation of further functional data on TDH, which could in turn help to guide drug discovery efforts. Furthermore, these findings add to the knowledge-base of the related enzymes of the SDR family.

Methods

Materials and Methods

Expression and Purification

The genes for TDH and KBL were amplified by PCR from *T. brucei* genomic DNA and cloned into the NdeI and BamHI sites of the *E. coli* expression vector pET15b using standard methods and then expressed in *E. coli*. The protein was then purified from the supernatant solutions extracted from cell lysates using nickel affinity chromatography, employing a HisTrap HP (GE Healthcare) to isolate His-tagged protein. Purity of protein solutions was determined by SDS-PAGE and by size-exclusion chromatography (SEC). SEC was performed on a Superdex 200 10/300 GL column (GE Healthcare), testing 200-400µl samples using Fast Performance Liquid Chromatography (FPLC) (Pharmacia Biotech) at a flow rate of 0.5ml/min and a maximum pressure of 1.5 MPa. Before any experimental samples were run, the column was calibrated according to the manufacturer's instructions with the standards. Protein elution volumes (V_e) were determined by following UV absorbance at 280 nm (A_{280}). The void volume (V_o) of the column was determined by running a sample of Blue dextran (MW 2000 kDa), which is not retained by the resin, due to its size. TDH and KBL were identified by using a standard calibration curve to predict the molecular weights of proteins eluting at a particular elution volume, with confirmation by SDS-PAGE.

Where it was desired to improve the purity of the protein solutions after affinity chromatography, ion-exchange chromatography (IEC) was employed. Appropriate conditions for IEC were determined by using the ProtParam tool on the ExPASy server to predict the isoelectric points of TDH and KBL from their amino acid sequences. IEC was performed on a 250ml Q-sepharose column using an ÄKTA Prime (GE Healthcare) FPLC machine. Protein was purified by eluting with a gradient of 1000ml solvent progressing from low salt buffer (100 mM NaCl) to high salt buffer (1 M NaCl).

TDH and KBL solutions were stored in ice baths and refrigerated at 2-8°C. The proteins appeared to be able to be stored in this way for at least one month without an appreciable loss in function. Alternatively, solutions were stored for longer periods, frozen at -20°C, and loss of function was not observed after several months of freezing the proteins.

X-ray crystallography

Crystallisation

Proteins were crystallised by use of the hanging drop method, using commercial crystallisation screens or manually prepared solutions (Table 1), as previously described.

Cocrystals of the proteins with substrates, inhibitors and metal ions were prepared using the same method, or crystals were soaked in solutions containing additional substances of interest for a period of 5 minutes to 1 hour.

As described previously, prior to data collection, crystals were harvested using a fabric or plastic loop and cryo-cooled by one of three methods: the loop and crystal were plunged directly into a bath of liquid nitrogen; the loop and crystal were suddenly exposed to a stream of cryo-cooled nitrogen gas;

the loop and crystal were plunged into liquid ethane for one second and then into liquid nitrogen. As cryo-cooling presents the risk of causing damage to crystals, cryoprotectants such as glycerol, ethylene glycol or PEG 400 were sometimes added. In these cases, crystals were transferred to a 10 µl drop of crystallisation well solution and four 1 µl drops of cryoprotectant were sequentially mixed into this solution before the crystal was cryo-cooled as described above.

Data collection and structure solution

X-ray data were collected at synchrotron radiation sources, the European Synchrotron Radiation Facility (ESRF) and Diamond Light Source (DLS). Early data processing was either carried out manually or by use of the integrated data reduction pipeline, xia2[47], at DLS. When early data processing was carried out manually, data integration was carried out using iMosflm[48]. When automated processing was carried out by xia2, the program XDS was used. Space group determination was initially attempted by selecting from iMosflm-provided options, or it was carried out automatically using POINTLESS[49,50]. Scaling and merging were carried out using Scala[49] or Aimless[51],together with ctruncate[52] to convert the intensities into structure factors. As an analogous TDH structure was already available in the PDB, phase determination for all TDH structures described herein was carried out by means of molecular replacement, using Molrep[53,54] or Phaser(174). In order to know the appropriate number of molecules per asymmetric unit, Matthews_coef[55] was used. Refinement was performed using Refmac5[56,57]. To incorporate ligands into the model, their structures were produced as .pdb files, along with restraints (stored in .cif files, used by Refmac), using the online ProDRG server[58]. Once positioned correctly, in accordance with its associated electron density, ligand files were merged with protein .pdb files.

The structure visualisation and manipulation software, Coot[59,60], was used on Windows and Linux systems to carry out real space refinement and validation of structures. Manipulation of structures was guided by 2Fo-Fc maps at σ levels of 1.5 and by Fo-Fc maps at σ levels of 2.00.

To validate models, bond distances and geometries, including those in hydrogen bonding networks, were observed visually and modified accordingly. The Ramachandron plot function in Coot was used to find and correct errors in the phi (ϕ) and psi (ψ) angles of amino acid residues. Coot's rotamer analysis tool was also used to correct side chain orientations in TDH structures.

Structural Analysis

Sequence Analysis

Nucleotide and amino acid sequences from *Tb*TDH (Tb927.6.2560) were compared with those of TDH from other organisms using the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST)[61]. Certain evolutionarily conserved sequences were also identified in the 2D sequences and visually using Coot and other visualisation software, such as UCSF Chimera[62].

3D structure

Protein structure models and their interactions with ligands were observed visually using PyMOL (Schrödinger LLC), Coot and UCSF Chimera(182). To analyse conformational variability evident in various *Tb*TDH structures, eight representative monomers from solved structures were compared in a method similar to Huntington [63]. The CCP4 program baverage was used to calculate average *B*-factors per residue for all main chain atoms. UCSF Chimera was used to visualise the mean *B*-factors and the RMSDs calculated between alpha carbons of the eight structures when superimposed.

The program PISA (Protein Interfaces, Surfaces and Assemblies)[27] in the CCP4 package was used to output the quaternary structures of TDH models for further analysis.

Geometric simulations of flexible motion

Amplitudes of motion in models of TDH were simulated using a combination of rigidity analysis and coarse-grained elastic network normal mode analysis, as described previously[33,64]. Elnemo software is used[65] to obtain normal mode eigenvectors from coarse-grained elastic network modelling, and FIRST/FRODA software [66,67] to carry out rigidity analysis (FIRST)[68], identifying the noncovalent interaction network and labelling dihedral angles as locked or variable, and template-based geometric simulations of flexible motion (FRODA)[67] which project the all-atom structure over large amplitudes of motion while maintaining local bonding and steric geometry.

Normal mode eigenvectors were generated in Elnemo in a one-site-per-residue coarse-graining using the Cα geometry of the input structure, placing springs of equal spring constant between all sites lying within an interaction distance cut-off of 12 Å. A rigidity analysis of the all-atom input structure is carried out in FIRST using the "pebble game" algorithm [66,69] which matches degrees of freedom against bonding constraints in the molecular framework of the protein. Bonding constraints include covalent, hydrophobic and polar (hydrogen bond and salt bridge) interactions. As the strength of the polar interactions can be gauged from their geometry, the results of the analysis depend on an "energy cut-off" which selects the set of polar interactions to include in the constraint network[68]. A cut-off of -2.0 kcal/mol is used in this study.

Template-based geometric simulation of flexible motion, carried out using FRODA, explores the mobility of the all-atom structure by iterative perturbation and relaxation of atomic positions in parallel and antiparallel to the direction of normal-mode eigenvectors. Several thousand iteration steps are carried out to generate large motion amplitudes. The simulation generates an initial phase of "easy" motion, where the bonding geometry is easily maintained, followed by the onset of "jamming" as the motion encounters steric and bonding constraints which naturally limit its

amplitude. The conformational changes of geometric simulations of TDH projected using this method were observed and compared with superpositions of crystallographic models in PyMOL and UCSF Chimera.

Other biochemical techniques

Size-exclusion chromatography

SEC was used as described above to predict the oligomeric states of TDH and KBL. Before any experimental samples were run, the column was calibrated with the standards: β-amylase (MW 200 kDa), alcohol dehydrogenase (MW 150 kDa) bovine serum albumin (BSA) (dimer MW 132 kDa; monomer MW 66 kDa), ovalbumin (MW 45 kDa), carbonic anhydrase (MW 24 kDa) and cytochrome C (12.4 kDa). To estimate the molecular weights of proteins eluting at a particular elution volume (V_e), a plot of $\log_{10}$ molecular weight of standard protein against the elution volume: void volume ratio (V_e/V_0) was made using Microsoft Excel. A number of experiments testing TDH and KBL in isolation, together and in the presence of substrates were carried out to measure changes in the observed elution volumes and oligomeric states predicted by linear regression analysis.

Cross-linking

Cross-linking studies of TDH and KBL were carried out using a method similar to that of Davies and Stark[70], which employed dimethyl suberimidate (DMS) as the cross-linking agent. As cross-linking reactions predominate within oligomers, this was used to provide an indication of the oligomeric states of TDH and KBL. Protein concentrations and DMS concentrations were carefully chosen to avoid precipitation.

All solutions were made up to 50µl or 100µl in 200mM Tris.HCl buffer at pH 8.5. Final protein concentrations were approximately 1mg/ml. Reactions were carried out in duplicate, at room temperature and at 2-8°C. The samples were incubated for 3 hours or overnight and the reactions

were then halted by the addition of Laemmli buffer in preparation for analysis by SDS-PAGE. Gels of 9% polyacrylamide, rather than the usual 12%, were used to increase the migration of higher molecular weight complexes.

Accession numbers

Coordinates and structure factors for all structures cited herein have been deposited in the Protein Data Bank with the following accession numbers: **5L9A**; **5LC1**; **5K4Q**; **5K4T**; **5K4V**; **5K4U**; **5K50**; **5K4W**; **5K4Y**.

Acknowledgements

We are grateful to UCL and UCL Division of Medicine for an Impact Studentship to EA and to the ESRF (Grenoble, France) and DLS (Didcot, UK) for synchrotron beam time and user-support (awards MX12342 and MX-1372). We are also grateful to UCL Business for a Proof of Concept award (POC-12-023) which supported PE.

References

[1] WHO | Trypanosomiasis, human African (sleeping sickness), WHO. (n.d.). http://www.who.int/mediacentre/factsheets/fs259/en/ (accessed December 1, 2016).
[2] WHO | Chagas disease (American trypanosomiasis), (n.d.). http://www.who.int/mediacentre/factsheets/fs340/en/index.html (accessed December 2, 2010).
[3] G.A. Cross, R.A. Klein, D.J. Linstead, Utilization of amino acids by Trypanosoma brucei in culture: L-threonine as a precursor for acetate, Parasitology. 71 (1975) 311–326.
[4] D.J. Linstead, R.A. Klein, G.A.M. Cross, Threonine Catabolism in Trypanosoma brucei, J Gen Microbiol. 101 (1977) 243–251. doi:10.1099/00221287-101-2-243.
[5] A.J. Edgar, The human L-threonine 3-dehydrogenase gene is an expressed pseudogene, BMC Genet. 3 (2002) 18.
[6] Y. Millerioux, C. Ebikeme, M. Biran, P. Morand, G. Bouyssou, I.M. Vincent, M. Mazet, L. Riviere, J.-M. Franconi, R.J.S. Burchmore, P. Moreau, M.P. Barrett, F. Bringaud, The threonine degradation pathway of the Trypanosoma brucei procyclic form: the main carbon source for lipid biosynthesis is under metabolic control, Mol. Microbiol. 90 (2013) 114–129. doi:10.1111/mmi.12351.
[7] M. Mazet, P. Morand, M. Biran, G. Bouyssou, P. Courtois, S. Daulouède, Y. Millerioux, J.-M. Franconi, P. Vincendeau, P. Moreau, F. Bringaud, Revisiting the Central Metabolism of the Bloodstream Forms of Trypanosoma brucei: Production of Acetate in the Mitochondrion Is

Essential for Parasite Viability, PLoS Negl Trop Dis. 7 (2013) e2587. doi:10.1371/journal.pntd.0002587.
[8] A. Bowyer, H. Mikolajek, J.W. Stuart, S.P. Wood, F. Jamil, N. Rashid, M. Akhtar, J.B. Cooper, Structure and function of the l-threonine dehydrogenase (TkTDH) from the hyperthermophilic archaeon Thermococcus kodakaraensis, J. Struct. Biol. 168 (2009) 294–304. doi:10.1016/j.jsb.2009.07.011.
[9] J.P. Marcus, E.E. Dekker, Threonine formation via the coupled activity of 2-amino-3-ketobutyrate coenzyme A lyase and threonine dehydrogenase., J. Bacteriol. 175 (1993) 6505–6511.
[10] V.M.S. Lam, I.P.R. Chan, Y.G. Yeung, Role of l-Threonine Deaminase and l-Threonine 3-Dehydrogenase in the Utilization of l-Threonine by Pseudomonas aeruginosa, J. Gen. Microbiol. 117 (1980) 539–542. doi:10.1099/00221287-117-2-539.
[11] M.I. Bird, P.B. Nunn, Metabolic homoeostasis of L-threonine in the normally-fed rat. Importance of liver threonine dehydrogenase activity., Biochem. J. 214 (1983) 687–694.
[12] A.J. Edgar, Mice have a transcribed L-threonine aldolase/GLY1 gene, but the human GLY1 gene is a non-processed pseudogene, BMC Genomics. 6 (2005) 32. doi:10.1186/1471-2164-6-32.
[13] R.A. Dale, Catabolism of threonine in mammals by coupling of l-threonine 3-dehydrogenase with 2-amino-3-oxobutyrate-CoA ligase, Biochim. Biophys. Acta BBA - Gen. Subj. 544 (1978) 496–503. doi:10.1016/0304-4165(78)90324-0.
[14] K.L. Kavanagh, H. Jörnvall, B. Persson, U. Oppermann, Medium- and short-chain dehydrogenase/reductase gene and protein families, Cell. Mol. Life Sci. 65 (2008) 3895–3906. doi:10.1007/s00018-008-8588-y.
[15] S.A. Boylan, E.E. Dekker, L-threonine dehydrogenase of Escherichia coli K-12, Biochem. Biophys. Res. Commun. 85 (1978) 190–197. doi:10.1016/S0006-291X(78)80028-X.
[16] S.A. Boylan, E.E. Dekker, L-threonine dehydrogenase. Purification and properties of the homogeneous enzyme from Escherichia coli K-12., J. Biol. Chem. 256 (1981) 1809–1815.
[17] N. Higashi, T. Matsuura, A. Nakagawa, K. Ishikawa, Crystallization and preliminary X-ray analysis of hyperthermophilic L-threonine dehydrogenase from the archaeon Pyrococcus horikoshii, Acta Crystallograph. Sect. F Struct. Biol. Cryst. Commun. 61 (2005) 432–434. doi:10.1107/S174430910500881X.
[18] K. Ishikawa, N. Higashi, T. Nakamura, T. Matsuura, A. Nakagawa, The First Crystal Structure of l-Threonine Dehydrogenase, J. Mol. Biol. 366 (2007) 857–867. doi:10.1016/j.jmb.2006.11.060.
[19] R. Machielsen, J. van der Oost, Production and characterization of a thermostable L-threonine dehydrogenase from the hyperthermophilic archaeon Pyrococcus furiosus, FEBS J. 273 (2006) 2722–2729. doi:10.1111/j.1742-4658.2006.05290.x.
[20] K. Yoneda, H. Sakuraba, I. Muraoka, T. Oikawa, T. Ohshima, Crystal structure of UDP-galactose 4-epimerase-like l-threonine dehydrogenase belonging to the intermediate short-chain dehydrogenase-reductase superfamily, FEBS J. 277 (2010) 5124–5132. doi:10.1111/j.1742-4658.2010.07916.x.
[21] K. Yoneda, H. Sakuraba, T. Araki, T. Ohshima, Crystal Structure of Binary and Ternary Complexes of Archaeal UDP-galactose 4-Epimerase-like l-Threonine Dehydrogenase from Thermoplasma volcanium, J. Biol. Chem. 287 (2012) 12966–12974. doi:10.1074/jbc.M111.336958.
[22] T. Ueatrongchit, Y. Asano, Highly selective L-threonine 3-dehydrogenase from Cupriavidus necator and its use in determination of L-threonine, Anal. Biochem. 410 (2011) 44–56. doi:10.1016/j.ab.2010.11.003.
[23] A. Schmidt, J. Sivaraman, Y. Li, R. Larocque, J.A.R.G. Barbosa, C. Smith, A. Matte, J.D. Schrag, M. Cygler, Three-Dimensional Structure of 2-Amino-3-ketobutyrate CoA Ligase from Escherichia coli Complexed with a PLP−Substrate Intermediate: Inferred Reaction Mechanism, Biochemistry (Mosc.). 40 (2001) 5151–5160. doi:10.1021/bi002204y.

[24] D.J. Creek, M. Mazet, F. Achcar, J. Anderson, D.-H. Kim, R. Kamour, P. Morand, Y. Millerioux, M. Biran, E.J. Kerkhoven, A. Chokkathukalam, S.K. Weidt, K.E.V. Burgess, R. Breitling, D.G. Watson, F. Bringaud, M.P. Barrett, Probing the Metabolic Network in Bloodstream-Form Trypanosoma brucei Using Untargeted Metabolomics with Stable Isotope Labelled Glucose, PLOS Pathog. 11 (2015) e1004689. doi:10.1371/journal.ppat.1004689.
[25] J.H. Yuan, R.E. Austic, Characterization of hepatic L-threonine dehydrogenase of chicken, Comp. Biochem. Physiol. B Biochem. Mol. Biol. 130 (2001) 65–73.
[26] Y.C. Kao, L. Davis, Purification and Structural Characterization of Porcine L-Threonine Dehydrogenase, Protein Expr. Purif. 5 (1994) 423–431. doi:10.1006/prep.1994.1061.
[27] E. Krissinel, K. Henrick, Inference of Macromolecular Assemblies from Crystalline State, J. Mol. Biol. 372 (2007) 774–797. doi:10.1016/j.jmb.2007.05.022.
[28] R.A. Klein, J.M. Angus, A.E. Amadife, L. Smith, Stereospecificity of the threonine dehydrogenase from bloodstream Trypanosoma brucei, Comp. Biochem. Physiol. Part B Comp. Biochem. 66 (1980) 143–146. doi:10.1016/0305-0491(80)90097-8.
[29] T. Kazuoka, S. Takigawa, N. Arakawa, Y. Hizukuri, I. Muraoka, T. Oikawa, K. Soda, Novel psychrophilic and thermolabile L-threonine dehydrogenase from psychrophilic Cytophaga sp. strain KUC-1, J. Bacteriol. 185 (2003) 4483–4489.
[30] C. He, X. Huang, Y. Liu, F. Li, Y. Yang, H. Tao, C. Han, C. Zhao, Y. Xiao, Y. Shi, Structural insights on mouse l-threonine dehydrogenase: A regulatory role of Arg180 in catalysis, J. Struct. Biol. 192 (2015) 510–518. doi:10.1016/j.jsb.2015.10.014.
[31] A.R. Johnson, Y.-W. Chen, E.E. Dekker, Investigation of a Catalytic Zinc Binding Site in Escherichia coli -Threonine Dehydrogenase by Site-Directed Mutagenesis of Cysteine-38, Arch. Biochem. Biophys. 358 (1998) 211–221. doi:10.1006/abbi.1998.0845.
[32] B. Persson, Y. Kallberg, J.E. Bray, E. Bruford, S.L. Dellaporta, A.D. Favia, R.G. Duarte, H. Jörnvall, K.L. Kavanagh, N. Kedishvili, M. Kisiela, E. Maser, R. Mindnich, S. Orchard, T.M. Penning, J.M. Thornton, J. Adamski, U. Oppermann, The SDR (short-chain dehydrogenase/reductase and related enzymes) nomenclature initiative, Chem. Biol. Interact. 178 (2009) 94–98. doi:10.1016/j.cbi.2008.10.040.
[33] S.A. Wells, M.W. van der Kamp, J.D. McGeagh, A.J. Mulholland, Structure and Function in Homodimeric Enzymes: Simulations of Cooperative and Independent Functional Motions, PLOS ONE. 10 (2015) e0133372. doi:10.1371/journal.pone.0133372.
[34] S.A. Wells, S.J. Crennell, M.J. Danson, Structures of mesophilic and extremophilic citrate synthases reveal rigidity and flexibility for function, Proteins Struct. Funct. Bioinforma. 82 (2014) 2657–2670. doi:10.1002/prot.24630.
[35] S. Nakano, S. Okazaki, H. Tokiwa, Y. Asano, Binding of NAD+ and L-Threonine induces Stepwise Structural and Flexibility changes in Cupriavidus Necator L-Threonine Dehydrogenase, J. Biol. Chem. (2014) jbc.M113.540773. doi:10.1074/jbc.M113.540773.
[36] J.O. Winberg, M.K. Brendskag, I. Sylte, R.I. Lindstad, J.S. McKinley-McKee, The catalytic triad in Drosophila alcohol dehydrogenase: pH, temperature and molecular modelling studies, J. Mol. Biol. 294 (1999) 601–616. doi:10.1006/jmbi.1999.3235.
[37] O.A.B.S.M. Gani, O.A. Adekoya, L. Giurato, F. Spyrakis, P. Cozzini, S. Guccione, J.-O. Winberg, I. Sylte, Theoretical Calculations of the Catalytic Triad in Short-Chain Alcohol Dehydrogenases/Reductases, Biophys. J. 94 (2008) 1412–1427. doi:10.1529/biophysj.107.111096.
[38] A. Koumanov, J. Benach, S. Atrian, R. Gonzàlez-Duarte, A. Karshikoff, R. Ladenstein, The catalytic mechanism of Drosophila alcohol dehydrogenase: Evidence for a proton relay modulated by the coupled ionization of the active site Lysine/Tyrosine pair and a NAD+ ribose OH switch, Proteins Struct. Funct. Bioinforma. 51 (2003) 289–298. doi:10.1002/prot.10354.
[39] Y. Wuxiuer, E. Morgunova, N. Cols, A. Popov, A. Karshikoff, I. Sylte, R. Gonzàlez-Duarte, R. Ladenstein, J.-O. Winberg, An intact eight-membered water chain in drosophilid alcohol

dehydrogenases is essential for optimal enzyme activity, FEBS J. 279 (2012) 2940–2956. doi:10.1111/j.1742-4658.2012.08675.x.
[40] M.L. Green, W.H. Elliott, The enzymic formation of aminoacetone from threonine and its further metabolism, Biochem. J. 92 (1964) 537–549.
[41] Y. Aoyama, Y. Motokawa, L-Threonine dehydrogenase of chicken liver. Purification, characterization, and physiological significance., J. Biol. Chem. 256 (1981) 12367–12373.
[42] M. Ray, S. Ray, L-Threonine dehydrogenase from goat liver. Feedback inhibition by methylglyoxal., J. Biol. Chem. 260 (1985) 5913–5918.
[43] M. Wagner, J.R. Andreesen, Purification and characterization of threonine dehydrogenase from Clostridium sticklandii, Arch. Microbiol. 163 (1995) 286–290. doi:10.1007/BF00393382.
[44] S.M. Sullivan, T. Holyoak, Enzymes with lid-gated active sites must operate by an induced fit mechanism instead of conformational selection, Proc. Natl. Acad. Sci. 105 (2008) 13829–13834. doi:10.1073/pnas.0805364105.
[45] G.G. Hammes, Y.-C. Chang, T.G. Oas, Conformational selection or induced fit: A flux description of reaction mechanism, Proc. Natl. Acad. Sci. 106 (2009) 13737–13741. doi:10.1073/pnas.0907195106.
[46] J.P. Marcus, E.E. Dekker, pH-Dependent Decarboxylation of 2-Amino-3-ketobutyrate, the Unstable Intermediate in the Threonine Dehydrogenase-Initiated Pathway for Threonine Utilization, Biochem. Biophys. Res. Commun. 190 (1993) 1066–1072. doi:10.1006/bbrc.1993.1157.
[47] G. Winter, xia2: an expert system for macromolecular crystallography data reduction, J. Appl. Crystallogr. 43 (2010) 186–190. doi:10.1107/S0021889809045701.
[48] T.G.G. Battye, L. Kontogiannis, O. Johnson, H.R. Powell, A.G.W. Leslie, iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM, Acta Crystallogr. D Biol. Crystallogr. 67 (2011) 271–281. doi:10.1107/S0907444910048675.
[49] P. Evans, Scaling and assessment of data quality, Acta Crystallogr. D Biol. Crystallogr. 62 (2005) 72–82. doi:10.1107/S0907444905036693.
[50] P.R. Evans, An introduction to data reduction: space-group determination, scaling and intensity statistics, Acta Crystallogr. D Biol. Crystallogr. 67 (2011) 282–292. doi:10.1107/S090744491003982X.
[51] P.R. Evans, G.N. Murshudov, How good are my data and what is the resolution?, Acta Crystallogr. D Biol. Crystallogr. 69 (2013) 1204–1214. doi:10.1107/S0907444913000061.
[52] N. Stein, C. Ballard, Intensity to amplitude conversion using CTRUNCATE, Acta Crystallogr. A. (2009). http://scripts.iucr.org/cgi-bin/paper?a45329 (accessed June 2, 2014).
[53] A. Vagin, A. Teplyakov, MOLREP: an Automated Program for Molecular Replacement, J. Appl. Crystallogr. 30 (1997) 1022–1025. doi:10.1107/S0021889897006766.
[54] A. Vagin, A. Teplyakov, Molecular replacement with MOLREP, Acta Crystallogr. D Biol. Crystallogr. 66 (2009) 22–25. doi:10.1107/S0907444909042589.
[55] K.A. Kantardjieff, B. Rupp, Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein–nucleic acid complex crystals, Protein Sci. 12 (2003) 1865–1871. doi:10.1110/ps.0350503.
[56] G.N. Murshudov, A.A. Vagin, E.J. Dodson, Refinement of Macromolecular Structures by the Maximum-Likelihood Method, Acta Crystallogr. D Biol. Crystallogr. 53 (1997) 240–255. doi:10.1107/S0907444996012255.
[57] A.A. Vagin, R.A. Steiner, A.A. Lebedev, L. Potterton, S. McNicholas, F. Long, G.N. Murshudov, REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use, Acta Crystallogr. D Biol. Crystallogr. 60 (2004) 2184–2195. doi:10.1107/S0907444904023510.
[58] A.W. Schüttelkopf, D.M.F. van Aalten, PRODRG: a tool for high-throughput crystallography of protein-ligand complexes, Acta Crystallogr. D Biol. Crystallogr. 60 (2004) 1355–1363. doi:10.1107/S0907444904011679.

[59] P. Emsley, K. Cowtan, Coot: model-building tools for molecular graphics, Acta Crystallogr. D Biol. Crystallogr. 60 (2004) 2126–2132. doi:10.1107/S0907444904019158.
[60] P. Emsley, B. Lohkamp, W.G. Scott, K. Cowtan, Features and development of Coot, Acta Crystallogr. D Biol. Crystallogr. 66 (2010) 486–501. doi:10.1107/S0907444910007493.
[61] BLAST: Basic Local Alignment Search Tool, (n.d.). http://blast.ncbi.nlm.nih.gov/Blast.cgi (accessed April 12, 2012).
[62] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, UCSF Chimera--a visualization system for exploratory research and analysis, J. Comput. Chem. 25 (2004) 1605–1612. doi:10.1002/jcc.20084.
[63] J.A. Huntington, How Na+ Activates Thrombin - A Review of the Functional and Structural Data, Biol. Chem. 389 (2008) 1025–1035.
[64] J.E. Jimenez-Roldan, R.B. Freedman, R.A. Römer, S.A. Wells, Rapid simulation of protein motion: merging flexibility, rigidity and normal mode analyses, Phys. Biol. 9 (2012) 16008. doi:10.1088/1478-3975/9/1/016008.
[65] K. Suhre, Y.-H. Sanejouand, ElNémo: a normal mode web server for protein movement analysis and the generation of templates for molecular replacement, Nucleic Acids Res. 32 (2004) W610–W614. doi:10.1093/nar/gkh368.
[66] D.J. Jacobs, A. j. Rader, L.A. Kuhn, M. f. Thorpe, Protein flexibility predictions using graph theory, Proteins Struct. Funct. Bioinforma. 44 (2001) 150–165. doi:10.1002/prot.1081.
[67] S. Wells, S. Menor, B. Hespenheide, M.F. Thorpe, Constrained geometric simulation of diffusive motion in proteins, Phys. Biol. 2 (2005) S127. doi:10.1088/1478-3975/2/4/S07.
[68] B.M. Hespenheide, A.J. Rader, M.F. Thorpe, L.A. Kuhn, Identifying protein folding cores from the evolution of flexible regions during unfolding, J. Mol. Graph. Model. 21 (2002) 195–207. doi:10.1016/S1093-3263(02)00146-8.
[69] D.J. Jacobs, M.F. Thorpe, Generic Rigidity Percolation: The Pebble Game, Phys. Rev. Lett. 75 (1995) 4051–4054. doi:10.1103/PhysRevLett.75.4051.
[70] G.E. Davies, G.R. Stark, Use of Dimethyl Suberimidate, a Cross-Linking Reagent, in Studying the Subunit Structure of Oligomeric Proteins, Proc. Natl. Acad. Sci. 66 (1970) 651–656.