



# An active site–tail interaction in the structure of hexahistidine-tagged *Thermoplasma acidophilum* citrate synthase

Jesse R. Murphy, Stefano Donini and T. Joseph Kappock\*

Department of Biochemistry, Purdue University, 175 South University Street, West Lafayette, IN 47907-2063, USA.

\*Correspondence e-mail: kappock@purdue.edu

Received 14 July 2015

Accepted 25 August 2015

Edited by L. J. Beamer, University of Missouri, USA

**Keywords:** thermophile; euryarchaeon; conformation change; carbon–carbon bond formation.

**PDB reference:** citrate synthase, 4ybo

**Supporting information:** this article has supporting information at journals.iucr.org/f

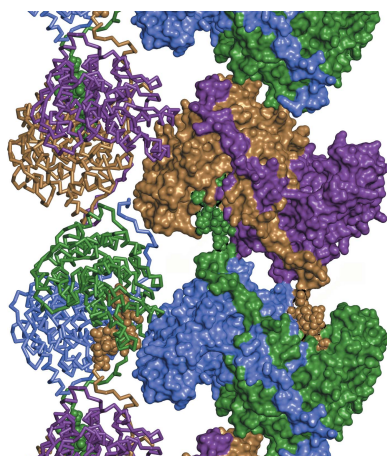
Citrate synthase (CS) plays a central metabolic role in aerobes and many other organisms. The CS reaction comprises two half-reactions: a Claisen aldol condensation of acetyl-CoA (AcCoA) and oxaloacetate (OAA) that forms citryl-CoA (CitCoA), and CitCoA hydrolysis. Protein conformational changes that ‘close’ the active site play an important role in the assembly of a catalytically competent condensation active site. CS from the thermoacidophile *Thermoplasma acidophilum* (TpCS) possesses an endogenous Trp fluorophore that can be used to monitor the condensation reaction. The 2.2 Å resolution crystal structure of TpCS fused to a C-terminal hexahistidine tag (TpCSH6) reported here is an ‘open’ structure that, when compared with several liganded TpCS structures, helps to define a complete path for active-site closure. One active site in each dimer binds a neighboring His tag, the first nonsubstrate ligand known to occupy both the AcCoA and OAA binding sites. Solution data collectively suggest that this fortuitous interaction is stabilized by the crystalline lattice. As a polar but almost neutral ligand, the active site–tail interaction provides a new starting point for the design of bisubstrate-analog inhibitors of CS.

## 1. Introduction

Citrate synthase (CS) performs two sequential reactions: a reversible condensation reaction converts acetyl coenzyme A (AcCoA) and oxaloacetate (OAA) into citryl-CoA (CitCoA), and an irreversible thioester hydrolysis then forms citrate and CoA. This pivotal metabolic reaction is performed by members of at least three enzyme superfamilies (Eggerer, 1965; Gottschalk & Barker, 1966; Kobylarz *et al.*, 2014).

The CS dimer is a classic case of induced-fit substrate binding (Srere, 1966; Bloxham *et al.*, 1980). OAA binding induces domain closure and the formation of an AcCoA binding site between the large and small domains of each subunit (Remington *et al.*, 1982; Wiegand & Remington, 1986; Daidone *et al.*, 2004). Since many conserved active-site residues participate in both the condensation and hydrolysis reactions, the central CS–CitCoA complex is expected to toggle among multiple configurations (Bayer *et al.*, 1981). Strong selective pressure ensures a high degree of substrate specificity: CS does not cleave AcCoA but efficiently hydrolyzes CitCoA (Srere, 1972).

A mechanism proposed for the condensation reaction, based in part on the crystal structures of liganded complexes (Karpusas *et al.*, 1990; Remington, 1992), remains broadly consistent with subsequent experimental findings. Computation-based models for the proton transfers involved



OPEN ACCESS

**Table 1**  
Macromolecule-production information.

|   |   |
|---|---|
| Source organism   | <i>T. acidophilum</i>   |
| DNA source  | Plasmid pJK438  |
| Forward primer†   | ACAGGAGTACATATGCCAGAACTGAAGAA   |
| Reverse primer†   | TTGAGAAAACTCGAGTCACTTTCTTCAGC   |
| Expression vector                                       | pET-24a   |
| Expression host   | <i>E. coli</i> C41(DE3)   |
| Complete amino-acid sequence of the construct produced‡ | PETEEISKGLDENVIKWTRLTTIDGNKGILRYGG-<br>YSVEDIIASGAQDEEIQYLFYGNLPTQEQLR-<br>KYKETVQKGYKIPDFVINAIRQLPRESDAVAM-<br>QMAAVALMAAASETGFKNKDDTRDVAEMIGR-<br>MSAITVNVYRHIMNMPAELPKPSDSYAESFLN-<br>AAFGRKATKEEIDAMNTALILYTDHEVPASTT-<br>AGLVAVSTLSDMYSGITAAALAKGPLHGGAA-<br>EAAIAQFDEIKDPAMVEKWFNDNIINGKKRLM-<br>FGHRYVYKTYDPRAKIFKGAIEKLSKKPEVH-<br>KVVEIATKLEDFGKAFSGKIYPTDYFSGI-<br>VYMSIGFPLRNNIYTALFALSRTVGTGWAHFIE-<br>YVEEQRLIRPRAVYVGAERKYVPIAERKVD-<br>KLAALAEHHHHHH |
| UniProt identifier                                      | P21553  |

† The restriction sites in ODNs 1306 (forward, NdeI) and 1309 (reverse, XhoI) are underlined. Stop codon 385 is replaced by the vector-encoded tag sequence. ‡ The recombinant protein lacks an N-terminal Met. The 15-residue His<sub>6</sub> tag appended to the C-terminus of the native sequence is underlined.

in carbon-carbon bond formation (Donini *et al.*, 2000; Mulholland *et al.*, 2000; Yang & Drueckhammer, 2003; van der Kamp *et al.*, 2008, 2010) have not reached consensus (Alexandrov *et al.*, 2014).

CS from the thermophilic, acidophilic euryarchaeon *Thermoplasma acidophilum* (TpCS; Danson *et al.*, 1985) has several advantages for mechanistic studies, among them an endogenous fluorophore that reports directly on enzyme chemistry (Kurz *et al.*, 2000, 2005). TpCS Trp348 fluorescence is strongly quenched by OAA binding nearby (Kurz *et al.*, 2005). The subsequent addition of either carboxymethyl-CoA (CMCoA) or carboxymethyl-dethiaCoA (CMX), potent CS inhibitors that resemble the deprotonated AcCoA species produced during the condensation reaction, forms a ‘closed’ ternary complex (Bayer *et al.*, 1981; Karpusas *et al.*, 1990; Kurz *et al.*, 1992) but has little effect on TpCS fluorescence (Kurz *et al.*, 2005). Quenching is, however, alleviated by the subsequent addition of dethiaacetyl-CoA (AcMX), an AcCoA analog that contains a methylene instead of sulfur (Martin *et al.*, 1994; Kurz *et al.*, 1997, 2005). The fluorescence increase is owing to destruction of the Trp348 quencher, the polarized OAA carbonyl, by its conversion to an *sp*<sup>3</sup> carbon in the non-hydrolyzable CitCoA analogue dethiacitryl-CoA (CitMX; Kurz *et al.*, 2009). Reversible, stoichiometric formation of a nonhydrolyzable binary TpCS-CitMX complex was detected by equilibrium fluorescence analysis, pre-steady-state kinetics and a crystal structure (PDB entry 2r9e; C. Lehmann, L. C. Kurz & T. E. Ellenberger, unpublished work). Since the TpCS protein conformation is almost the same in the ternary TpCS-OAA-CMCoA complex (PDB entry 2r26; C. Lehmann, L. C. Kurz & T. E. Ellenberger, unpublished work), the active site arrangement is likely to represent the configuration used for the condensation reaction, not hydrolysis.

This study addresses two barriers to the further study of conformational changes in TpCS. Firstly, an unliganded

structure of TpCS has been reported (Russell *et al.*, 1994) but the coordinates have not been deposited. This hampers the analysis of protein motions accompanying the formation of the TpCS-OAA (PDB entry 2ifc; C. Lehmann, L. C. Kurz & T. E. Ellenberger, unpublished work) and TpCS-OAA-CMCoA complexes. Secondly, TpCS purification involves a dye-linked affinity column (Sutherland *et al.*, 1991) that may displace (acyl-)CoA ligands (Weitzman & Ridley, 1983) and irreversibly binds at least one mutant protein (Constantine, 2009). Here, we report the purification, characterization and crystal structure of TpCS fused to a C-terminal hexahistidine affinity tag (TpCSH6). While the affinity tag has only minor effects on enzyme function in solution, it promotes the formation of a crystal with a His tag bound to an adjacent active site. Liganded and ‘empty’ partner subunits adopt essentially the same ‘open’ conformation.

## 2. Materials and methods

### 2.1. Macromolecule production

*Escherichia coli* C41(DE3) cells transformed with either pJK438 or pJK511 (cloning procedures are given in the Supporting Information) were propagated in LB medium supplemented with 50 mg l<sup>-1</sup> kanamycin. Production cultures (1 l) were grown at 310 K to an optical density at 600 nm of 0.6, at which point isopropyl β-D-1-thiogalactopyranoside (IPTG; Gold Biotechnology, St Louis, Missouri, USA) was added (0.4 mM final concentration) to induce protein production. After a further 16 h, the cells were harvested by centrifugation (8000g for 15 min at 277 K), resuspended in 8 ml buffer *H* (20 mM Tris-HCl pH 8.0, 100 mM KCl) per gram of wet cell weight and disrupted by sonication. Streptomycin was added to a final concentration of 1% (w/v) from a 10% stock and solids were removed by centrifugation (30 000g for 10 min at 277 K). Column-chromatography steps were performed at 295 K under gravity flow. An Ni<sup>2+</sup>-loaded iminodiacetic acid Sepharose column (Sigma-Aldrich, St Louis, Missouri, USA; 2.5 × 3.5 cm, 14 ml column volume) was washed with buffer *H* (140 ml) and the cleared cell lysate was applied. The column was washed with buffer *H* containing 20 mM imidazole (42 ml) and was developed with buffer *H* containing 250 mM imidazole (100 ml). Protein-containing fractions were pooled and concentrated to >5 mg ml<sup>-1</sup> by ultrafiltration using an Amicon Ultra-15 centrifugal filter unit (EMD Millipore, Billerica, Massachusetts, USA). Solid ammonium sulfate was added to 85% saturation at 277 K over 30 min. After stirring for a further 30 min at 277 K, aliquots were taken and stored as a slurry at 277 K. For crystal production, an aliquot was dissolved in a minimal volume of TE (20 mM Tris-HCl pH 8.0, 1 mM EDTA) and applied onto a Sephadex 200 gel-filtration column (2.5 × 28 cm; Pharmacia, Uppsala, Sweden) that was developed in TE. Protein-containing fractions were collected, pooled and concentrated to ~5 mg ml<sup>-1</sup>. Exchanged samples were kept at 277 K and used within 12 h. Proteins were quantitated by the method of Bradford using bovine serum albumin as a standard

**Table 2**  
Crystallization.

|  |  |
|--|--|
| Method                                       | Hanging-drop vapor diffusion                             |
| Plate type                                   | VDX 24-well tray, Hampton Research                       |
| Temperature (K)                              | 294  |
| Protein concentration (mg ml <sup>-1</sup> ) | 5  |
| Buffer composition of protein solution       | 20 mM Tris-HCl pH 8.0, 1 mM EDTA, 1 mM oxaloacetate      |
| Composition of reservoir solution            | 0.1 M CHES pH 9.5, 12% (w/v) PEG 4000, 1 mM oxaloacetate |
| Volume and ratio of drop                     | 2 µl (1:1 ratio)   |
| Volume of reservoir (ml)                     | 0.5  |

(Bradford, 1976). Gel-filtration analysis was used to determine molecular sizes as described by Mullins *et al.* (2013). Macromolecule-production information is summarized in Table 1.

Enzyme activities were determined using a continuous assay that detects CoA release by monitoring the cleavage of 5,5'-dithiobis(2-nitrobenzoic acid) (DTNB) at 412 nm ( $\Delta\epsilon = 14.1 \text{ mM}^{-1} \text{ cm}^{-1}$ ; Srere *et al.*, 1963; Srere, 1969; Riddles *et al.*, 1979). One unit (U) is defined as the amount of enzyme required to produce 1 µmol of product per minute. Additional information on biochemical characterization is given in the Supporting Information.

## 2.2. Crystallization

Sparse-matrix crystallization screens performed using the Wizard I and II kits (Emerald Bio) yielded two hits after 27 d: condition I-1 [20% PEG 8000, 0.1 M 2-(cyclohexylamino)ethanesulfonate (CHES) pH 9.5] and condition I-26 (Table 2), each supplemented with 1 mM OAA. Crystals with a similar appearance were obtained in the same buffer after 10 d in drops (2–4 µl prior to mixing with an equal volume of reservoir solution) containing 8–10% PEG 4000 and 5 mg ml<sup>-1</sup> TpCSH6 with or without OAA. Single crystals were loaded into nylon loops (Teng, 1990), transferred to reservoir solution supplemented with 15% (w/v) sorbitol for 10 min, rapidly immersed in liquid nitrogen and maintained at or below ~100 K until data collection was complete.

## 2.3. Data collection and processing

All screened crystals diffracted X-rays weakly and with high mosaicity. The best diffraction patterns were obtained from a TpCSH6 crystal that adhered to the side of the mounting loop. The unusual sample geometry hampered the collection of a complete X-ray diffraction data set, which was recorded in three passes using different regions of the crystal. Diffraction data from one-pass, two-pass or three-pass sets were processed using the *HKL-2000* suite (Otwinowski & Minor, 1997). Single-pass data sets had similar statistics (not shown), including individual  $R_{\text{sym}}$  values that were comparable to the  $R_{\text{merge}}$  value for the full (three-pass) data set (Table 3).

## 2.4. Structure solution and refinement

*PHENIX* and *Coot* were used for structure solution and refinement (Emsley *et al.*, 2010; Adams *et al.*, 2010). Molecular replacement was performed with *Phaser* (McCoy *et al.*, 2007) using subunit A from PDB entry 2ifc, with all buffer compo-

**Table 3**  
Data collection and processing.

Values in parentheses are for the outer shell.

|   |                          |
|---|--------------------------|
| Diffraction source                                    | 21-ID-G, APS             |
| Wavelength (Å)  | 0.97856                  |
| Temperature (K)                                       | 100                      |
| Detector  | MAR Mosaic 300 mm CCD    |
| Crystal-to-detector distance (mm)                     | 200, 247.8 and 250       |
| Rotation range per image (°)                          | 0.5, 1.0 and 1.0         |
| Total rotation range (°)                              | 93, 122 and 360          |
| Exposure time per image (s)                           | 3.0, 2.8 and 3.5         |
| Space group   | $P2_1$                   |
| $a, b, c$ (Å)   | 56.505, 113.406, 120.066 |
| $\alpha, \beta, \gamma$ (°)                           | 90, 95.08, 90            |
| Resolution range (Å)                                  | 50–2.18 (2.24–2.18)      |
| Total No. of reflections                              | 983117                   |
| No. of unique reflections                             | 76084                    |
| Completeness (%)                                      | 97.2 (97.0)              |
| Multiplicity  | 12.9 (13.5)              |
| $\langle I/\sigma(I) \rangle$                         | 11.8 (2.1)               |
| $R_{\text{meas}}$                                     | 0.249 (1.400)            |
| $R_{\text{p.i.m.}}$                                   | 0.085 (0.422)            |
| $CC_{1/2}$  | 0.993 (0.762)            |
| Overall $B$ factor from Wilson plot (Å <sup>2</sup> ) | 24.9                     |

nents and side-chain atoms beyond  $C^\beta$  removed, as the search model. The sole solution contained four subunits in the asymmetric unit (residues 4–383; the TpCS numbering excludes Met0). One round of *phenix.autobuild* (with default settings) was performed to repair side chains. Iterative cycles of model improvement and refinement, using noncrystallographic symmetry (NCS) restraints and a hybrid TLS-isotropic ADP model, were performed until  $R_{\text{free}}$  appeared to converge. The C-terminal appendage (residues Val386–His399) was added using *Coot*. NCS restraints were dropped in the final refinement cycles. *MolProbity* was used to check the protein geometry (Chen *et al.*, 2010).

Pairwise alignments and root-mean-square deviation (r.m.s.d.) computations were performed using one cycle of the *align* algorithm implemented in *PyMOL* (v.1.7.4.2; DeLano, 2002). Domain motions were computed using the *DynDom* server (v.1.5; default settings; Hayward & Berendsen, 1998; Hayward & Lee, 2002). Protein images were based on *PyMOL* or *LIGPLOT* (Wallace *et al.*, 1995) output.

## 3. Results and discussion

### 3.1. Comparison of TpCSH6 and TpCS

TpCSH6 was overexpressed in *E. coli*, purified and found to be a dimer (Supplementary Fig. S1) with a specific activity of 52 U mg<sup>-1</sup> at 328 K and  $\leq 22$  U mg<sup>-1</sup> at 298 K (8 s<sup>-1</sup>, assuming one active site per subunit). Substrate-saturation analyses (Supplementary Fig. S2) yielded Michaelis constant ( $K_M$ ) values of 3.0 µM for AcCoA and 5.2 µM for OAA. AcMX was a competitive inhibitor versus AcCoA ( $K_i = 20$  µM; Supplementary Fig. S3). These results, and the relatively high batch-to-batch variability in specific activity, are comparable to those for TpCS (Sutherland *et al.*, 1991; Kurz *et al.*, 2000).

**Table 4**  
Structure refinement.

Values in parentheses are for the outer shell.

|                                       |                        |
|---------------------------------------|------------------------|
| Resolution range (Å)                  | 29.15–2.18 (2.24–2.18) |
| Completeness (%)                      | 89.8 (76.0)            |
| $\sigma$ Cutoff                       | 0                      |
| No. of reflections, working set       | 70618 (4415)           |
| No. of reflections, test set          | 1867 (126)             |
| Final $R_{\text{cryst}}$              | 0.1642 (0.2385)        |
| Final $R_{\text{free}}$               | 0.2118 (0.2965)        |
| No. of non-H atoms                    |                        |
| Protein (all atoms)                   | 12251                  |
| Protein (His <sub>6</sub> -tag atoms) | 277                    |
| Ligand                                | 8                      |
| Water                                 | 868                    |
| Total                                 | 13127                  |
| R.m.s. deviations                     |                        |
| Bonds (Å)                             | 0.012                  |
| Angles (°)                            | 1.139                  |
| Average $B$ factors (Å <sup>2</sup> ) |                        |
| Protein (all atoms)                   | 33.8                   |
| Protein (His <sub>6</sub> -tag atoms) | 63.6                   |
| Ligand                                | 45.5                   |
| Water                                 | 36.7                   |
| Ramachandran plot                     |                        |
| Favoured regions (%)                  | 97.93                  |
| Additionally allowed (%)              | 2.01                   |
| Outliers (%)                          | 0.06                   |
| <i>MolProbity</i> clashscore†         | 4.86                   |
| <i>MolProbity</i> overall score       | 1.08                   |

† Clashscore is the number of interatomic overlaps ( $\geq 0.4$  Å) per 1000 atoms (Chen *et al.*, 2010).

Purified TpCSH6-H222Q, a mutant located in the AcCoA binding site, had a specific activity of  $\leq 0.01$  U mg<sup>-1</sup> at 298 K or  $\leq 0.05\%$  of that of the wild type. Since the mutant enzyme was isolated from a *gltA*<sup>+</sup> *E. coli* strain, the observed activity could be associated with a contaminating host enzyme. However, the ‘activity’ did not increase as AcCoA was varied from 30 to 400  $\mu$ M (at 0.4 mM OAA; data not shown), which suggests that host CS is not present. We conclude that TpCSH6-H222Q is nearly or completely devoid of enzymatic activity, and that the protein-isolation method can be used to study even very low-activity mutants produced in common protein-production strains.

Fluorescence titrations were used to measure OAA dissociation constants ( $K_d$ ) of 0.76 and 19  $\mu$ M for TpCSH6 and TpCSH6-H222Q, respectively (Supplementary Fig. S4). Saturating OAA quenched the fluorescence of both forms to about half of the initial level, which differs from a report that TpCS-H222Q fluorescence is quenched to a relatively greater extent (Kurz *et al.*, 2009). As observed for TpCS, the addition of AcMX to the TpCSH6–OAA complex results in a 124% increase in fluorescence ( $K_d = 1.5$   $\mu$ M) owing to the formation of CitMX (Supplementary Fig. S5). The  $K_d$  values determined for untagged TpCS(-H222Q) at 293 K were comparable: 1.3, 33 and 1.47  $\mu$ M, respectively (Kurz *et al.*, 2005, 2009).

### 3.2. Crystal structure of TpCSH6

Initial attempts to crystallize TpCSH6 focused on HEPES/sodium acetate (pH 7.5–8.5) conditions, identified by robotic screening, that yielded crystals of TpCS bound to one or more

ligands (Christopher Lehmann, unpublished observations). No suitable TpCSH6 crystals were obtained in attempts to manually reproduce these conditions. A sparse-matrix screen identified high-pH conditions that yielded large crystals (Table 2). The inclusion of OAA in the crystallization solution (Table 2) had no obvious effect on crystal morphology. A cryocooled crystal (0.1  $\times$  0.1  $\times$  1.0 mm) grown in the presence of OAA was selected for X-ray data collection. Several attempts to crystallize TpCSH6-H222Q under similar conditions yielded only microcrystals.

The three-pass data set was selected for further analysis because the completeness and electron-density maps (Supplementary Fig. S6) were superior to those computed using the one-pass or two-pass data sets (not shown). Molecular replacement and refinement proceeded smoothly, furnishing a final model with acceptable statistics (Table 4).

Like TpCS (Russell *et al.*, 1994), TpCSH6 crystallized in space group  $P2_1$  with two dimers in the asymmetric unit, albeit with rather different unit-cell parameters (Table 3). TpCSH6 had almost the same monomer topology as unliganded TpCS (Supplementary Fig. S7). Pairwise alignments of TpCS (coordinates provided by Linda C. Kurz) and TpCSH6 subunits gave r.m.s.d. values of 0.40–0.59 Å ( $C^\alpha$  only) and 1.15–1.25 Å (all protein atoms). Pairwise alignments of the TpCSH6 subunits (residues 5–383) gave r.m.s.d. values of 0.26–0.42 Å ( $C^\alpha$  only) and 0.68–0.91 Å (all protein atoms). TpCSH6 subunit *C* and the *CD* dimer were selected to represent the open conformation in all subsequent comparisons.

As expected for a relatively long exposure to synchrotron X-rays, alterations consistent with mild X-ray damage were observed, including Asp/Glu decarboxylation and loss of methanethiol in surface-accessible residues (Garman, 2010) such as Asp301*B* and Met241*C*. Two well ordered bicarbonate ligands were included in the final model. This was justified by the high pH of the mother liquor and the presence of OAA, which can spontaneously decarboxylate. These anions ‘cap’ the N-terminal (positive) end of the long  $\alpha$ Q helix in subunits *B* and *C*. The side chain of the single Ramachandran outlier, Lys237*C*, makes a well defined salt bridge with the side chain of Glu158*D* from a neighboring asymmetric unit.

Positive difference electron density was observed early in refinement at two locations: a helix near the *CD* dimer interface and the nearby subunit *D* active site. These features were linked to form the C-terminus of subunit *A*. Dimers *AB* and *CD* are bridged at the center of the asymmetric unit by two C-terminal appendages extending from subunits *A* and *D* (Fig. 1). The C-terminal native sequence of subunit *A* wraps halfway around the large domain of the *B* subunit. The C-terminal tag sequence continues past helix  $\alpha$ T towards the *CD* interface, making a perpendicular turn to form the extra helix ( $\alpha$ U). A second sharp turn near residue His222*D* directs the tag sequence into the active site of subunit *D*, ending with a bidentate salt bridge formed between the C-terminal carboxylate of the tag residue His399*A* and the Arg271*D* guanidinium. A similar interaction is formed between the C-terminus of subunit *D* and the active site of subunit *A*. NCS

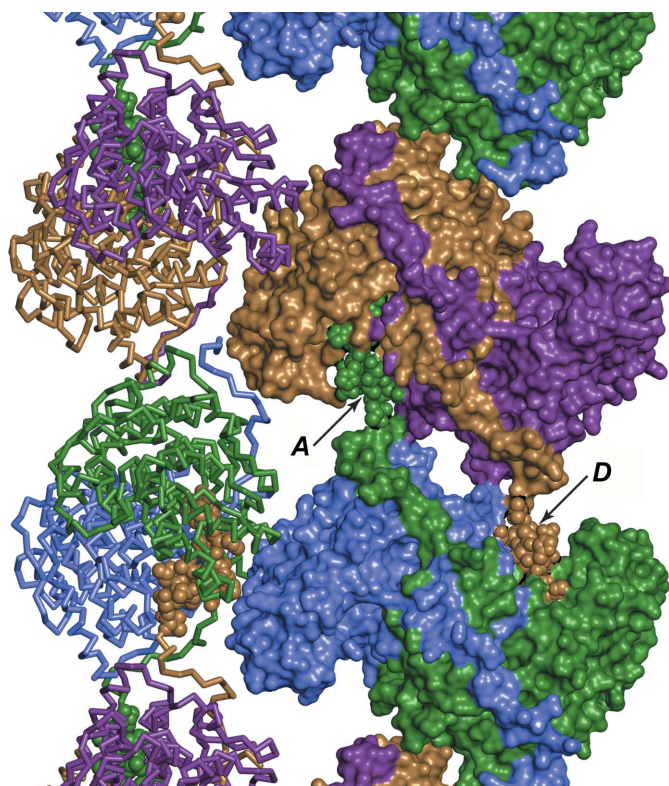


comparisons indicate the backbones diverge after Arg383, with helix  $\alpha$ U displaced along its axis about 2 Å farther from the CD dimer than the AB dimer. No electron density consistent with the C-terminal appendage of subunits B or C was observed. Small affinity-tag appendages occasionally form crystal-packing contacts (Carson *et al.*, 2007). The TpCSH6 His<sub>6</sub> tag is one of a small number of ordered affinity tags that bind to an enzyme active site (Taylor *et al.*, 2005; McDonald *et al.*, 2007; Singh *et al.*, 2011; Wojtkowiak *et al.*, 2012).

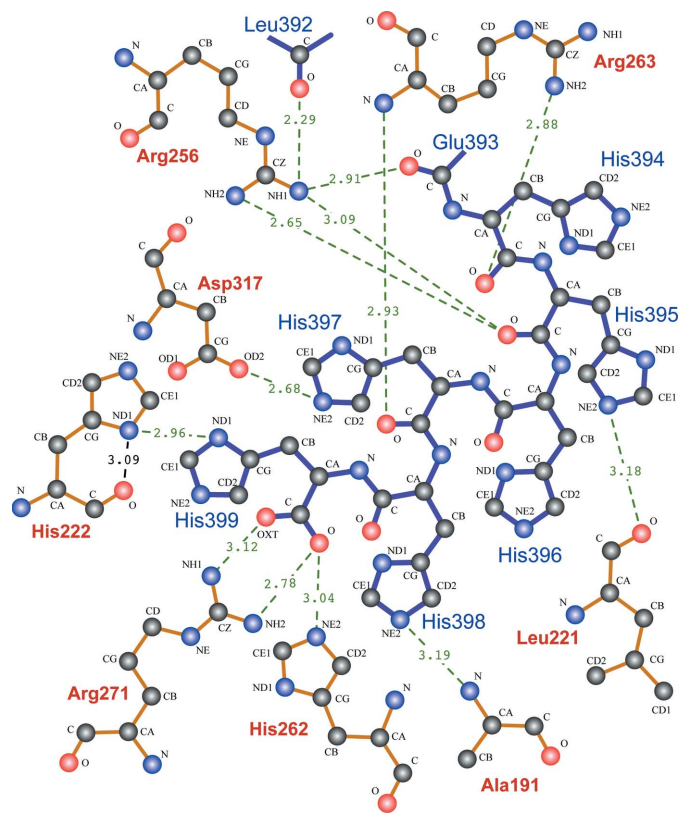
The C-terminal appendage contacts active-site residues involved in binding both OAA and AcCoA (Fig. 2). This was unexpected, as the hexahistidine sequence has no resemblance to either substrate. The C-terminal residue His399 makes a side-chain contact with His222, a residue that provides a hydrogen bond (involving N<sup>δ1</sup>) that polarizes the AcCoA enolate, and the terminal carboxylate contacts both His262 and Arg271, which bind carbonyl and  $\alpha$ -carboxylate O atoms in OAA. Since the His399–His222 contact involves two N<sup>δ1</sup> atoms, one residue adopts a  $\pi$  tautomer (N<sup>δ1</sup>–H), which is uncommon at high pH (Sudmeier *et al.*, 2003; Li & Hong, 2011) but is employed by some active-site residues (Day *et al.*, 2003). The local hydrogen-bonding context and the proposed role for the neutral  $\pi$  tautomer in the condensation reaction (summarized in van der Kamp *et al.*, 2007) suggest that His222

adopts the  $\pi$  tautomer even in the absence of acetyl-CoA. Similarly, flipping the His398 imidazole would worsen the hydrogen-bonding context and introduce clashes, which suggests it may also adopt a  $\pi$  tautomer stabilized by Ala191 NH (Fig. 2). His397 N<sup>ε2</sup> forms a hydrogen bond to one carboxylate O atom in the critical active-site base Asp317, which is not directly contacted by either substrate (Supplementary Fig. S8). Similar contacts are observed between Asp317 equivalents and strong inhibitors such as CMCoA and CMX (Usher *et al.*, 1994; Francois *et al.*, 2006).

Bisubstrate analogues can be potent and selective enzyme inhibitors (Collins & Stark, 1971; Radzicka & Wolfenden, 1995; Schramm, 2013). Clinically useful examples include finasteride, which is converted to a tight-binding bisubstrate analogue by its target enzyme (Bull *et al.*, 1996). The stoichiometric ligand CitMX is a bisubstrate analogue inhibitor of CS that (like finasteride) is formed by its target enzyme (Kurz *et al.*, 2009). Reversible formation of CitMX by TpCS might therefore contribute to inhibition by AcMX (Martin *et al.*, 1994), although its potency could be sapped by harnessing

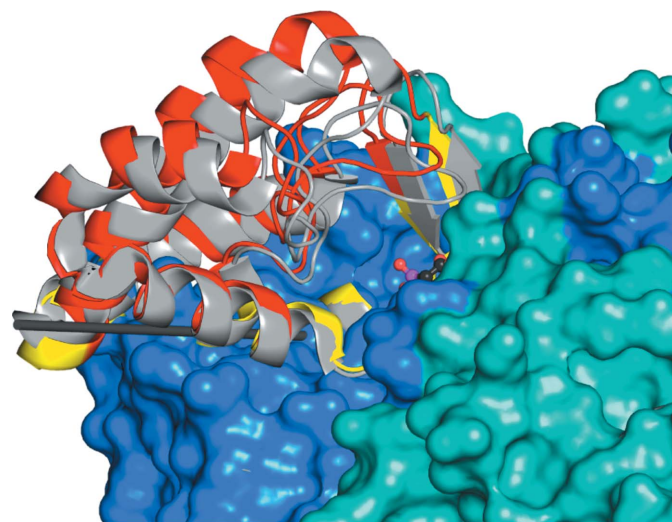


**Figure 1**  
Crystal-packing diagram for TpCSH6. Subunits A–D in the monoclinic asymmetric unit are colored green, blue, purple and gold, respectively. Residues in the C-terminal appendage, shown as spheres in subunits A and D (indicated by arrows), occupy the active sites of subunits D and A, respectively. Left, ribbon rendering shows that the C-termini bind deep within the AB and CD dimers. In this orientation, the twofold screw axis (*b* axis) runs along the Cartesian *x* axis.

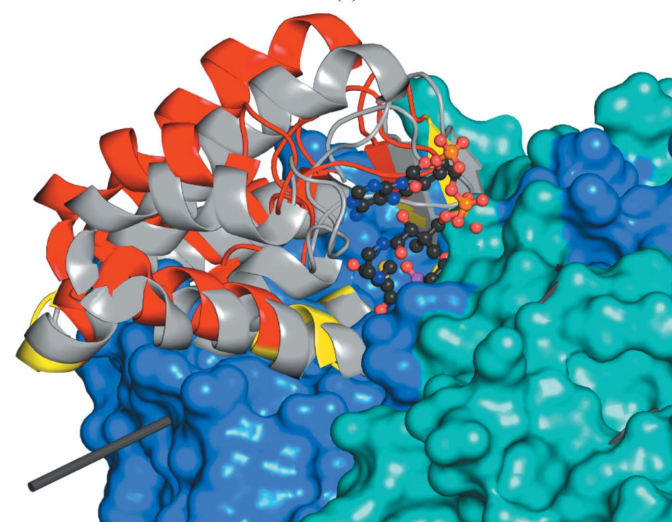


**Figure 2**  
Active site of TpCSH6 subunit A bound to the C-terminus of subunit D (blue bonds). All illustrated residues also contact OAA (His262 and Arg271) or AcCoA (the remainder). His222 and Asp317 are responsible for deprotonating the AcCoA acetyl group, stabilizing the resulting enolate/carbanion and promoting the condensation reaction that produces CitCoA. An intra-residue hydrogen bond that stabilizes the neutral  $\pi$  tautomer of His222 is shown in black. A similar analysis of the TpCS active site bound to the CitCoA analogue CitMX (Supplementary Fig. S8) shows that many of the same residues are involved in substrate contacts. All polar contacts presumed to result in hydrogen-bonding interactions are shown, with distances in Å.

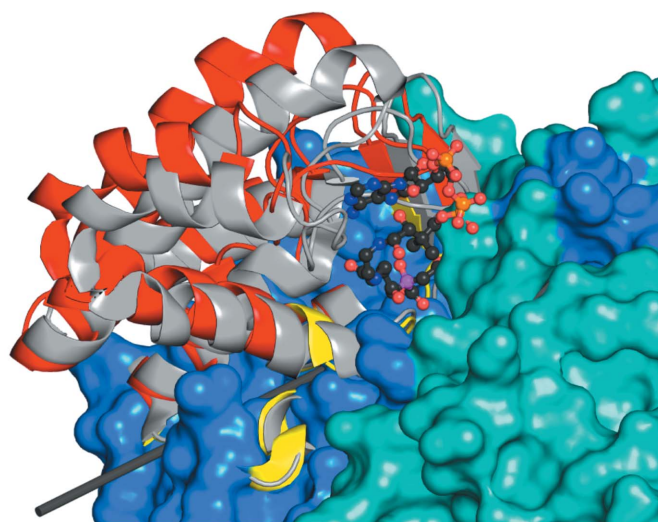
'binding energy' to drive the condensation reaction (Page & Jencks, 1971). CS is a highly specific enzyme that possesses a highly basic active site that binds two polyanionic substrates. All known potent CS inhibitors are unpromising starting points for drug design: they closely resemble the substrates and would encounter substantial barriers to cellular uptake. In contrast, imidazoles should be mostly uncharged at the basic pH, which may allow crystal growth by alleviating electrostatic repulsion between the active site and the His<sub>6</sub> tail. This near-neutral peptide ligand could represent a useful alternative starting point for the identification of CS inhibitors.



(a)



(b)



(c)

**Figure 3**

Domain motions in TpCS(H6). Each panel shows the open TpCSH6 structure superimposed on a closed TpCS structure in pairwise *DynDom* analyses. The static portion, mainly the large domains that form the majority of the dimer interface, is shown as a surface rendering. The mobile domain extends from the surface and is shown in cartoon rendering, with the TpCSH6 mobile region colored red and the closed structure colored grey. Hinging residues in each TpCSH6 domain are colored yellow. Residues involved in each pairwise comparison are shown in Supplementary Fig. S7. The rotation axis is shown as a black bar; closure results in clockwise rotations about this axis. (a) Comparison with the TpCS–OAA complex (PDB entry 2ifc subunit A) shows 52% closure and a 8.9° rotation (translation of 0.1 Å). Mobile domain: 99 residues, r.m.s.d. 0.41 Å. Static domain: 277 residues, r.m.s.d. 0.52 Å. (b) Comparison with the TpCS–OAA–CMCoA complex (PDB entry 2r26 subunit A) shows 86% closure and a 10.6° rotation (translation of 0.4 Å). Mobile domain: 115 residues, r.m.s.d. 0.68 Å. Static domain: 258 residues, r.m.s.d. 0.69 Å. CMCoA is a tight-binding analogue of the deprotonated AcCoA formed during the condensation reaction. (c) Comparison with the TpCS–CitMX complex (PDB entry 2r9e subunit B) shows 82% closure and a 10.3° rotation (translation 0.4 Å). Mobile domain: 123 residues, r.m.s.d. 0.83 Å. Static domain: 252 residues, r.m.s.d. 0.7 Å. CitMX is formed from AcMX and OAA, presumably within the crystal.

To test the hypothesis that the CS active site has detectable affinity for a poly-His ligand, TpCSH6 was incubated with a dansyl-His<sub>3</sub> synthetic peptide (Genscript, Piscataway, New Jersey, USA). Fluorescence changes consistent with Förster resonance energy transfer between the tripeptide and TpCSH6, presumably the nearby Trp115, were not observed (data not shown).

Gel-filtration studies indicated that unliganded TpCS, TpCSH6 and TpCSH6-H222Q were each exclusively dimeric in solution (Supplementary Fig. S1). Deletion of a hydrogen bond between the active-site residue His222 and the appendage residue His399 appears to thwart TpCSH6-H222Q crystal growth. This sensitivity suggests a weak peptide–protein interaction that is stabilized by cooperative crystal-packing interactions.

### 3.3. TpCS conformational changes associated with ligand binding

Substrate exclusion by the C-terminal appendage could explain why the crystal morphology was unaffected by the addition of saturating OAA but only half of the active sites are occupied by peptide in the crystal (ordered solvent is observed in the other two active sites). The striking finding that all four subunits adopt open conformations suggests that the TpCS dimer cannot close just one active site. Previous CS crystal structures have a high degree of NCS and contain the same ligand(s) in both active sites. Since to our knowledge there is



no evidence for cooperative binding of OAA by any CS, it is possible that the open conformation is trapped within the crystalline lattice.

The preferred rotamer at Trp348, or the residue occupying this position, has been reported to change upon ligand binding and concomitant enzyme closure (Kurz *et al.*, 2005). The TpCS structure possesses *m*95 (Trp348*A*) and *p*90 (Trp348*B*) rotamers (Russell *et al.*, 1994), each modeled with a  $C^\alpha-C^\beta-C^\gamma$  bond-angle outlier ( $\sigma > 5.5$ ). In TpCSH6 and the three deposited TpCS structures, however, only *m*0 rotamers were observed in unambiguous electron density in all subunits (Supplementary Fig. S6). It is possible that binding of the C-terminal appendage to the TpCSH6 active site, which does not cause domain closure, triggers a Trp348 ring flip. A simpler model, however, is that the Trp348 indole does not ring flip, consistent with the fluorescence properties, which indicate a rigidly immobilizing protein environment (Kurz *et al.*, 2005).

Using the new TpCSH6 structure as a reference, we examined the series of protein conformational changes associated with ligand binding, domain closure and the condensation reaction (Supplementary Fig. S7). As anticipated, the binding of OAA is associated with a conspicuous closure of the active site: a rotation of the small domain relative to the large domains that form the major dimer interface (Fig. 3*a*). Each of the domains is relatively monolithic, with pairwise r.m.s.d. values of  $<1$  Å, and similar sets of hinge residues are involved. This motion is similar to the larger ( $19.4^\circ$ ) rotation observed in the canonical vertebrate CS open/closed pair (PDB entries 1cts/1csh; Remington *et al.*, 1982; Usher *et al.*, 1994; Hayward & Berendsen, 1998). Domain closure increases when a CoA analogue is also present (Figs. 3*b* and 3*c*). As noted above, the TpCS–OAA–CMCoA and TpCS–CitMX complexes adopt similar protein structures, suggesting that they both represent the configuration associated with the condensation reaction.

#### 4. Conclusions

TpCSH6 is an active enzyme useful for solution studies of the CS mechanism. It unexpectedly produced a new crystal form that appears to prevent ligand cocrystallization. The serendipitous peptide ligand, with a total charge of near zero, binds both parts of the CS active site. Using this complex as an inspiration, it may be possible to identify an alternative drug scaffold to target this key step in primary metabolism.

#### Acknowledgements

We thank Joshua A. Mosberg for constructing pJK438, Jason Markovich for initial protein-expression trials, Elwood A. Mullins for help with X-ray data collection, Kelly L. Sullivan for performing gel-filtration analysis, Barbara L. Golden and Jeremy R. Lohman for processing X-ray data and Linda C. Kurz for materials, advice related to CS and discussions of CS conformational dynamics. This research used the resources of the Advanced Photon Source, a US Department of Energy (DOE) Office of Science User Facility operated for the DOE

Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. Use of LS-CAT Sector 21 was supported by the Michigan Economic Development Corporation and the Michigan Technology Tri-Corridor (Grant 085P1000817).

#### References

- Adams, P. D. *et al.* (2010). *Acta Cryst.* **D66**, 213–221.  
 Aleksandrov, A., Zvereva, E. & Field, M. (2014). *J. Phys. Chem. B*, **118**, 4505–4513.  
 Bayer, E., Bauer, B. & Eggerer, H. (1981). *Eur. J. Biochem.* **120**, 155–160.  
 Bloxham, D. P., Ericsson, L. H., Titani, K., Walsh, K. A. & Neurath, H. (1980). *Biochemistry*, **19**, 3979–3985.  
 Bradford, M. M. (1976). *Anal. Biochem.* **72**, 248–254.  
 Bull, H. G., Garcia-Calvo, M., Andersson, S., Baginsky, W. F., Chan, H. K., Ellsworth, D. E., Miller, R. R., Stearns, R. A., Bakshi, R. K., Rasmuson, G. H., Tolman, R. L., Myers, R. W., Kozarich, J. W. & Harris, G. S. (1996). *J. Am. Chem. Soc.* **118**, 2359–2365.  
 Carson, M., Johnson, D. H., McDonald, H., Brouillette, C. & DeLucas, L. J. (2007). *Acta Cryst.* **D63**, 295–301.  
 Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* **D66**, 12–21.  
 Collins, K. D. & Stark, G. R. (1971). *J. Biol. Chem.* **246**, 6599–6605.  
 Constantine, C. Z. (2009). PhD thesis. Washington University in St Louis, St Louis, Missouri, USA.  
 Daidone, I., Roccatano, D. & Hayward, S. (2004). *J. Mol. Biol.* **339**, 515–525.  
 Danson, M. J., Black, S. C., Woodland, D. L. & Wood, P. A. (1985). *FEBS Lett.* **179**, 120–124.  
 Day, R. M., Thalhauser, C. J., Sudmeier, J. L., Vincent, M. P., Torchilin, E. V., Sanford, D. G., Bachovchin, C. W. & Bachovchin, W. W. (2003). *Protein Sci.* **12**, 794–810.  
 DeLano, W. L. (2002). *PyMOL*. <http://www.pymol.org>.  
 Donini, O., Darden, T. & Kollman, P. A. (2000). *J. Am. Chem. Soc.* **122**, 12270–12280.  
 Eggerer, H. (1965). *Biochem. Z.* **343**, 111–138.  
 Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* **D66**, 486–501.  
 Francois, J. A., Starks, C. M., Sivanuntakorn, S., Jiang, H., Ransome, A. E., Nam, J.-W., Constantine, C. Z. & Kappock, T. J. (2006). *Biochemistry*, **45**, 13487–13499.  
 Garman, E. F. (2010). *Acta Cryst.* **D66**, 339–351.  
 Gottschalk, G. & Barker, H. A. (1966). *Biochemistry*, **5**, 1125–1133.  
 Hayward, S. & Berendsen, H. J. C. (1998). *Proteins*, **30**, 144–154.  
 Hayward, S. & Lee, R. A. (2002). *J. Mol. Graph. Model.* **21**, 181–183.  
 Kamp, M. W. van der, Perruccio, F. & Mulholland, A. J. (2007). *Proteins*, **69**, 521–535.  
 Kamp, M. W. van der, Perruccio, F. & Mulholland, A. J. (2008). *Chem. Commun.*, pp. 1874.  
 Kamp, M. W. van der, Żurek, J., Manby, F. R., Harvey, J. N. & Mulholland, A. J. (2010). *J. Phys. Chem. B*, **114**, 11303–11314.  
 Karpusas, M., Branchaud, B. & Remington, S. J. (1990). *Biochemistry*, **29**, 2213–2219.  
 Kobylarz, M. J., Grigg, J. C., Sheldon, J. R., Heinrichs, D. E. & Murphy, M. E. P. (2014). *J. Biol. Chem.* **289**, 33797–33807.  
 Kurz, L. C., Constantine, C. Z., Jiang, H. & Kappock, T. J. (2009). *Biochemistry*, **48**, 7878–7891.  
 Kurz, L. C., Drysdale, G., Riley, M., Tomar, M. A., Chen, J., Russell, R. J. M. & Danson, M. J. (2000). *Biochemistry*, **39**, 2283–2296.  
 Kurz, L. C., Fite, B., Jean, J., Park, J., Erpelding, T. & Callis, P. (2005). *Biochemistry*, **44**, 1394–1413.  
 Kurz, L. C., Roble, J. H., Nakra, T., Drysdale, G. R., Buzan, J. M., Schwartz, B. & Drucekhammer, D. G. (1997). *Biochemistry*, **36**, 3981–3990.

- Kurz, L. C., Shah, S., Crane, B. R., Donald, L. J., Duckworth, H. W. & Drysdale, G. R. (1992). *Biochemistry*, **31**, 7899–7907.
- Li, S. & Hong, M. (2011). *J. Am. Chem. Soc.* **133**, 1534–1544.
- Martin, D. P., Bibart, R. T. & Drueckhammer, D. G. (1994). *J. Am. Chem. Soc.* **116**, 4660–4668.
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
- McDonald, H. M., Pruet, P. S., Deivanayagam, C., Protasevich, I. I., Carson, W. M., DeLucas, L. J., Brouillette, W. J. & Brouillette, C. G. (2007). *Acta Cryst.* **D63**, 891–905.
- Mulholland, A. J., Lyne, P. D. & Karplus, M. (2000). *J. Am. Chem. Soc.* **122**, 534–535.
- Mullins, E. A., Sullivan, K. L. & Kappock, T. J. (2013). *PLoS One*, **8**, e67901.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Page, M. I. & Jencks, W. P. (1971). *Proc. Natl Acad. Sci. USA*, **68**, 1678–1683.
- Radzicka, A. & Wolfenden, R. (1995). *Methods Enzymol.* **249**, 284–312.
- Remington, S. J. (1992). *Curr. Opin. Struct. Biol.* **2**, 730–735.
- Remington, S., Wiegand, G. & Huber, R. (1982). *J. Mol. Biol.* **158**, 111–152.
- Riddles, P. W., Blakeley, R. L. & Zerner, B. (1979). *Anal. Biochem.* **94**, 75–81.
- Russell, R. J. M., Hough, D. W., Danson, M. J. & Taylor, G. L. (1994). *Structure*, **2**, 1157–1167.
- Schramm, V. L. (2013). *ACS Chem. Biol.* **8**, 71–81.
- Singh, H., Malinski, T. J., Reilly, T. J., Henzl, M. T. & Tanner, J. J. (2011). *Arch. Biochem. Biophys.* **509**, 76–81.
- Srere, P. A. (1966). *J. Biol. Chem.* **241**, 2157–2165.
- Srere, P. A. (1969). *Methods Enzymol.* **13**, 3–11.
- Srere, P. A. (1972). *Curr. Top. Cell. Regul.* **5**, 229–283.
- Srere, P. A., Brazil, H., Gonen, L. & Takahashi, M. (1963). *Acta Chem. Scand.* **17**, S129–S134.
- Sudmeier, J. L., Bradshaw, E. M., Haddad, K. E. C., Day, R. M., Thalhauser, C. J., Bullock, P. A. & Bachovchin, W. W. (2003). *J. Am. Chem. Soc.* **125**, 8430–8431.
- Sutherland, K. J., Danson, M. J., Hough, D. W. & Towner, P. (1991). *FEBS Lett.* **282**, 132–134.
- Taylor, E. J., Goyal, A., Guerreiro, C. I. P. D., Prates, J. A. M., Money, V. A., Ferry, N., Morland, C., Planas, A., Macdonald, J. A., Stick, R. V., Gilbert, H. J., Fontes, C. M. G. A. & Davies, G. J. (2005). *J. Biol. Chem.* **280**, 32761–32767.
- Teng, T.-Y. (1990). *J. Appl. Cryst.* **23**, 387–391.
- Usher, K. C., Remington, S. J., Martin, D. P. & Drueckhammer, D. G. (1994). *Biochemistry*, **33**, 7753–7759.
- Wallace, A. C., Laskowski, R. A. & Thornton, J. M. (1995). *Protein Eng. Des. Sel.* **8**, 127–134.
- Weitzman, P. D. J. & Ridley, J. (1983). *Biochem. Biophys. Res. Commun.* **112**, 1021–1026.
- Wiegand, G. & Remington, S. J. (1986). *Annu. Rev. Biophys. Biophys. Chem.* **15**, 97–117.
- Wojtkowiak, A., Witek, K., Hennig, J. & Jaskolski, M. (2012). *Acta Cryst.* **D68**, 713–723.
- Yang, W. & Drueckhammer, D. G. (2003). *J. Phys. Chem. B*, **107**, 5986–5994.