Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

The structure of *Cryptococcus neoformans* thymidylate synthase suggests strategies for using target dynamics for species-specific inhibition

Janet S. Finer-Moore,[a] Amy C. Anderson,[b]* Robert H. O'Neil,[b] M. Paola Costi,[c] Stefania Ferrari,[c] Jolanta Krucinski[a] and Robert M. Stroud[a]

[a]Department of Biochemistry and Biophysics, University of California, San Francisco, 600 16th Street, Room S412B, San Francisco, CA 94143-2240, USA, [b]Department of Chemistry, 6128 Burke Laboratories, Dartmouth College, Hanover, NH 03755, USA, and [c]Dipartimento Di Scienze Farmaceutiche, Universita degli Studi di Modena e di Reggio Emilia, 183 Via Campi, Modena, Italy

Correspondence e-mail: amy.c.anderson@dartmouth.edu

Received 16 February 2005
Accepted 14 July 2005

PDB References: CnTS–dUMP–CB3717 (C2 space group), 2aaz, r2aazsf; EcTS–dUMP–GA9, 2a9w, r2a9wsf.

The ternary complex crystal structures of *Cryptococcus neoformans* and *Escherichia coli* thymidylate synthase (TS) suggest mechanisms of species-specific inhibition of a highly conserved protein. The 2.1 Å structure of *C. neoformans* TS cocrystallized with substrate and the cofactor analog CB3717 shows that the binding sites for substrate and cofactor are highly conserved with respect to human TS, but that the structure of the cofactor-binding site of *C. neoformans* TS is less constrained by surrounding residues. This feature might allow *C. neoformans* TS to form TS–dUMP–inhibitor complexes with a greater range of antifolates than human TS. 3′,3″-Dibromophenol-4-chloro-1,8-naphthalein (GA9) selectively inhibits both *E. coli* TS and *C. neoformans* TS (K_i = 4 μ*M*) over human TS (K_i >> 245 μ*M*). The *E. coli* TS–dUMP–GA9 complex is in an open conformation, similar to that of the apoenzyme crystal structure. The GA9-binding site overlaps the binding site of the pABA-glutamyl moiety of the cofactor. The fact that human apoTS can adopt an unusual fold in which the GA9-binding site is disordered [Phan *et al.* (2001), *J. Biol. Chem.* **276**, 14170–14177] may explain the poor affinity of GA9 for the human enzyme. These observations highlight the critical need to incorporate multiple target conformations in any computational attempt to facilitate drug discovery.

1. Introduction

Cryptococcus neoformans is an opportunistic fungal pathogen that is a particular threat to patients with T-cell deficiencies, such as those suffering from AIDS (Mitchell & Perfect, 1995). The organism enters the host respiratory system and in immunocompromized patients the infection then frequently invades the central nervous system, causing meningo-encephalitis (Mitchell & Perfect, 1995). If untreated, cryptococcal meningitis is fatal and conventional antifungal treatments are only moderately effective and toxic; thus, novel treatments are needed (Brouwer *et al.*, 2004).

Thymidylate synthase (TS) is an essential enzyme required for the sole *de novo* synthesis of 2′-deoxyribosylthymine 5′-monophosphate (dTMP). Since dTMP is required for synthesis of DNA, human TS (hTS) has been a primary target for anti-cancer drugs such as 5-fluorouracil (Heidelberger *et al.*, 1983). TSs are required for dTMP synthesis in almost all eukaryotes and in a wide range of bacteria and archaea; TSs from pathogenic species such as *C. neoformans* are potential drug targets for treating infectious diseases.

Several TS inhibitors have been developed that are analogs of the enzyme substrate or cofactor. TS binds to its substrate, 2′-deoxyuridine 5′-monophosphate (dUMP), using 14 hydrogen bonds that are conserved among both prokaryotic

and eukaryotic TS species (Stroud & Finer-Moore, 2003). The high specificity of the binding site precludes the binding of most substituents of dUMP, thus aside from the mechanism-based inhibitor 5-fluoro-2′-deoxyuridine 5′-monophosphate (FdUMP), most drugs or drug candidates targeting TS are analogs of the cofactor 5,10-methylenetetrahydrofolate (CH_2H_4folate; Santi & Danenberg, 1984; Takemura & Jackman, 1997). Classical antifolates contain components corresponding to the pterin, *p*-amino benzoic acid (pABA) and glutamyl moieties of the cofactor (Fig. 1). They have single glutamyl moieties, but as with the cofactor, additional glutamate residues are added inside cells to form a poly-glutamyl tail (Shane, 1989).

Kinetic studies suggest that the cofactor and its analogs employ similar binding mechanisms (Pogolotti *et al.*, 1986; Santi *et al.*, 1987; Stroud & Finer-Moore, 2003): after forming a rapidly reversible complex with TS–dUMP (they do not bind appreciably to the apoenzyme since dUMP forms a large part of their binding sites), they drive the enzyme into a closed conformation through interactions with several conserved hydrophobic residues on flexible protein segments surrounding the active site. Concomitantly, the protein C-terminus moves into the active site to hydrogen bond to the cofactor or inhibitor and the catalytic thiol forms a covalent bond to the 6-position of dUMP. Because of the flexibility of the binding site, a surprisingly broad range of folate analogs (Fig. 1) can form such closed complexes with a covalently bound dUMP. Once formed, the complexes are stable and well ordered. While the details of their structures depend on the inhibitor (Fig. 2*a*), the ternary complexes of a given inhibitor are structurally conserved between different species of TS (Fig. 2*b*; Anderson *et al.*, 2001; Stroud & Finer-Moore, 2003); therefore, the antifolates are not generally species-specific (Ferrari *et al.*, 2003). However, the antifolate *N*,*O*-didansyl-L-tyrosine (Fig. 1) specifically inhibits bacterial TS species by harnessing the flexibility of the folate-binding site to access nonconserved residues outside the active site (Fritz *et al.*, 2001). Conceivably, the flexibility of the cofactor site may provide the basis for other species-specific antifolates.

One successful strategy for discovering novel classes of species-specific inhibitors uses biological screening against a panel of different TS species to guide synthetic elaboration of a molecular scaffold (Costi *et al.*, 1999). Using this method, a series of 1,8-naphthalein inhibitors that selectively inhibit TSs

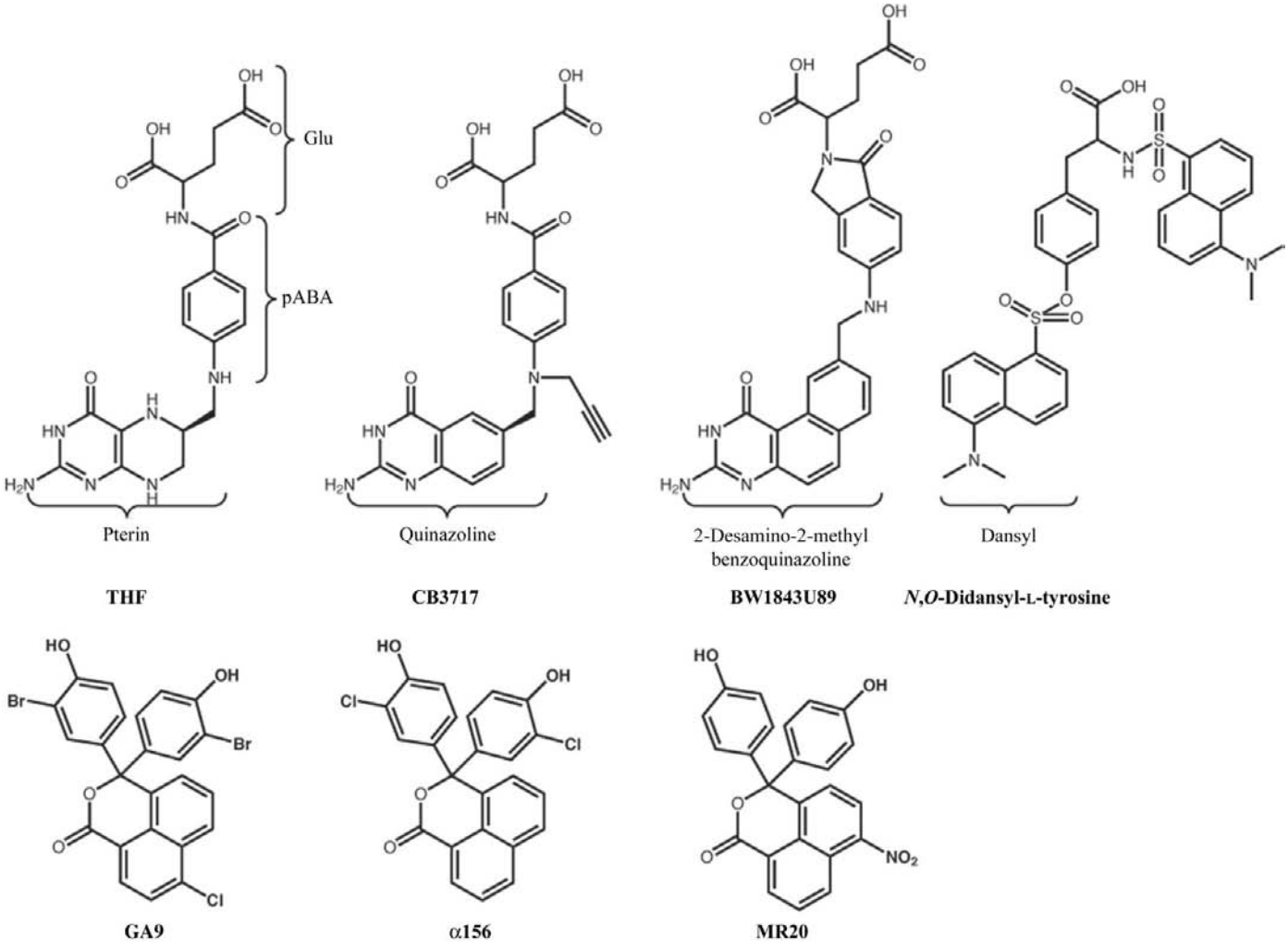

Figure 1
Chemical structures of the TS cofactor, three antifolates and three of a series of 1,8-naphthalein derivatives that inhibit TS.

from several pathogens, including *C. neoformans*, over human TS has been developed (Fig. 1). These inhibitors do not resemble antifolates in structure or binding mechanism. They often bind within the spacious active-site cavity, but are neither competitive with dUMP nor do they utilize dUMP as a binding surface or induce closure of the enzyme. In some cases they bind to multiple sites in the enzyme (Stout *et al.*, 1999). Ferrari and coworkers found that the specificity of one of these inhibitors for bacterial TSs could be explained by differences between the dynamic modes of the bacterial and human apoenzymes (Ferrari *et al.*, 2003). In this analysis, the measured binding affinity of the inhibitor was considered to be the weighted average of the binding affinities for each of the conformational states present in the apoenzyme. Bacterial TSs had a greater proportion of conformations in which the inhibitor-binding site was accessible than did eukaryotic TSs. One inhibitor in this series, 3′,3″-dibromophenol-4-chloro-1,8-naphthalein, named GA9, inhibits a variety of prokaryotic and eukaryotic TSs, including *C. neoformans* TS (CnTS), *Escherichia coli* TS (EcTS), *Lactobacillus casei* TS (LcTS) and *Pneumocystis carinii* TS (PcTS), with K_i values in the low micromolar range (Table 1), but it does not inhibit hTS even at the solubility limit for the compound (245 μ*M*; Costi *et al.*, in preparation).

We aimed to determine the crystal structure of a CnTS–dUMP–GA9 complex in order to increase the selectivity and affinity of GA9 for CnTS using structure-based drug design, but were unable to obtain crystals. Therefore, we first determined the crystal structure at 2.1 Å resolution of CnTS in complex with dUMP and the classical antifolate 10-propargyl-5,8-dideazafolate (CB3717; Jones *et al.*, 1981). The structure of this reaction-intermediate analog reveals distinctive features of the CnTS active site that may bear on the feasibility of discovering CnTS-specific antifolates.

We then determined the 1.65 Å structure of EcTS in complex with dUMP and GA9 in order to understand the basis for the selectivity of this inhibitor. Interestingly, the specificity of GA9 cannot be easily rationalized from the sequence differences at the GA9-binding site in EcTS. Following the reasoning of Ferrari and coworkers, we propose that GA9 specificity results from differences in the spectrum of conformational states characteristic of apohTS compared with pathogenic apoTSs. Specifically, apohTS is in equilibrium with an alternate conformation in which the active-site cysteine is flipped out of the active site and a domain is refolded (Fig. 3). In this conformation, the GA9-binding site is disordered, effectively removing some molecules of apohTS from the population available to productively bind GA9 (Phan, Steadman *et al.*, 2001). This conformation has so far not been seen in other TS species and may be responsible for the low affinity of GA9 for hTS.

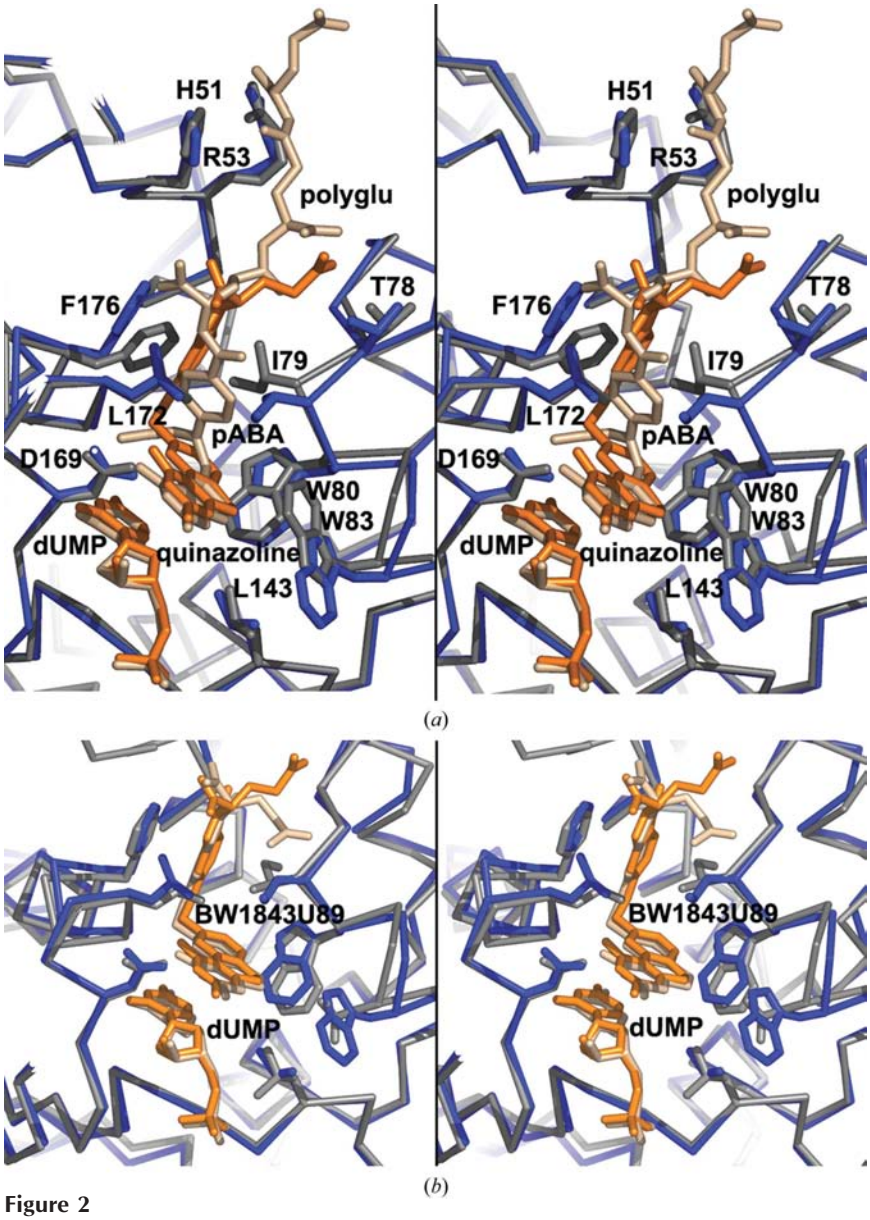

Figure 2
(*a*) Overlay of the structures of EcTS ternary complexes with two antifolates, CB3717 (Kamb *et al.*, 1992) and BW1843U89 (PDB code 1syn), in divergent-eye stereo, illustrating the flexibility of the pABA-glutamyl region of the cofactor-binding site. For clarity, the C-terminal tail covering the binding site is not shown. (*b*) Overlay of the EcTS and PcTS ternary complexes with BW1843U89 (PDB codes 1syn and 1f28, respectively), illustrating the conservation of ligand-induced conformational changes across species.

2. Methods

2.1. CnTS production and crystallization

CnTS was expressed in *E. coli* and purified to homogeneity as previously described (Livi *et al.*, 1994). CnTS was crystallized in

Table 1
GA9 inhibition constants for five species.

	CnTS	EcTS	LcTS	PcTS	hTS
K_i (μM)	4.1	4.1	0.8	28	>>245

complex with 2 m*M* dUMP and 2 m*M* CB3717 by hanging-drop vapor-diffusion techniques using PEG 4000 or 6000 as the precipitant and sodium or potassium acetate as an additive. Initial crystals were obtained under a variety of PEG conditions at room temperature by mixing three parts protein solution (7.3 mg ml^{-1} CnTS, 2 m*M* dUMP, 2 m*M* CB3717 and 5 m*M* DTT) with two parts reservoir solution (20–27% PEG 3K, 6K or 8K in 100 m*M* Tris pH 8.0–8.8 and 150 m*M* sodium or potassium acetate). Room-temperature crystals diffracted to ∼3.5 Å resolution and were tetragonal, belonging to space group $P4_1$, with unit-cell parameters $a = 127.7$, $c = 208.6$ Å and two dimers in the asymmetric unit.

Much better diffracting crystals were obtained at 277 K using microseeding. Seeds from crushed crystals were used to inoculate hanging drops (3 µl 12 mg ml^{-1} protein premixed with 2 m*M* dUMP, 2 m*M* CB3717 and 5 m*M* DTT and 2 µl 23–25% PEG 4K or 6K in 100 m*M* Tris pH 8.4 with 150 m*M* sodium or potassium acetate) that had been pre-equilibrated for 1–2 d. Crystals grew to dimensions of a few hundred micrometres within a week. The 277 K crystals diffracted to 2 Å resolution and were monoclinic, belonging to space group *C*2, with unit-cell parameters $a = 179.4$, $b = 179.5$, $c = 209.1$ Å, $\beta = 89.8^\circ$ and eight in the asymmetric unit.

2.2. CnTS data collection

2.2.1. Tetragonal crystal form. 30 1.5° data frames from each of two fragments of a tetragonal crystal were collected at room temperature on an R-AXIS II image-plate detector using monochromated Cu *Kα* radiation. Because of crystal decay, only a total of 32 frames from the two fragments were included in the final data set. Data were processed, scaled and merged in space group $P4_12_12$ with the *HKL* suite (Otwinowski & Minor, 1997), yielding a 84% complete 3.5 Å data set with an overall $R_{\rm merge}$ of 17.3% and an $\langle I/\sigma(I)\rangle$ of 6.5 (Table 2). The same data scaled and merged in space group $P4_1$ was 65% complete, with $R_{\rm merge} = 15.3\%$.

2.2.2. Monoclinic crystal form. A monoclinic crystal of CnTS, grown at 277 K, was transferred to a cryoprotectant solution containing 15% glycerol and soaked for 10 min before being frozen at 103 K. Data were collected at 103 K on a MAR Research image-plate detector at Stanford Synchrotron Radiation Laboratory beamline 9-1. Data reduction and scaling using *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997) yielded a 97% complete 2.08 Å data with an overall $R_{\rm merge}$ of 10% and an $\langle I/\sigma(I)\rangle$ of 7.4 (see Table 2 and supplementary material[1]).

2.3. CnTS structure solution

The lower resolution tetragonal structure was solved by molecular replacement in space group $P4_12_12$ with *AMoRe* (Navaza, 1994) using a dimer of *Leishmania major* TS (Knighton *et al.*, 1994) as the search model. A single dimer was located in the asymmetric unit. The CnTS sequence, excluding the CnTS-unique inserts, was fitted to the density maps using *CHAIN* and refined by energy minimization in *X-PLOR* (Brünger, 1992).

Since the dimer represented only 20% of the unit cell and did not account for extensive protein-like density in adjacent regions of the maps, a second translation search was performed in which the refined dimer contributed a fixed contribution to the calculated structure factors. The second translation search located two additional half-occupied and overlapping positions for the TS dimer that together

[1] Supplementary material has been deposited in the IUCr electronic archive (Reference: SX5032). Services for accessing these data are described at the back of the journal.

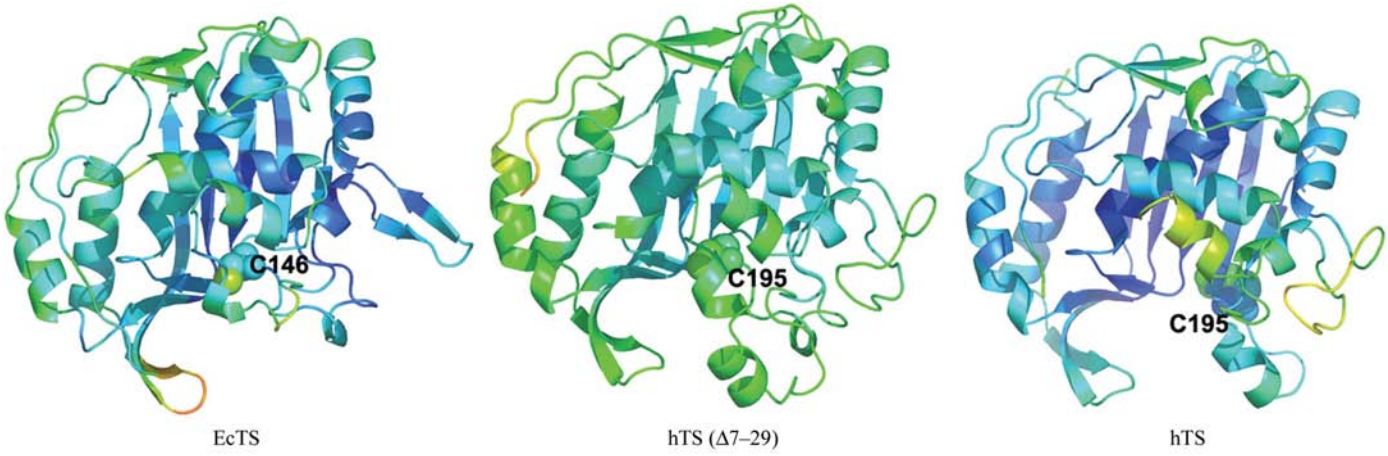

Figure 3
Ribbon drawings from crystal structures of apoEcTS (PDB code 1tjs), apohTS with the N-terminal deletion Δ7–29 (PDB code 1hzw; the N-terminal regions of eukaryotic TSs are unstructured) and full-length apohTS (PDB code 1hw3), color-coded by *B* factor. The color code progresses through the color spectrum with increasing *B* factor from blue for the lowest *B*-factor regions to orange for residues with *B* factors between 60 and 70 Å^2. The full-length apohTS is in a radically different conformation in which the active-site cysteine, shown with spheres and labeled, is flipped out of the active site and a domain of the enzyme is refolded.

Table 2
Crystallographic statistics.

Values in parentheses are for the high-resolution bin.

Protein	CnTS	CnTS	EcTS
Ligands	CB3717, dUMP	CB3717, dUMP	GA9, dUMP
Unit-cell parameters			
a (Å)	127.7	179.4	186.6
b (Å)	127.7	179.5	186.6
c (Å)	208.6	209.1	114.2
β (°)	—	89.8	—
Space group	$P4_1$	$C2$	$R3$
Z	32	64	36
V_M† (Å^3 Da^{-1})	2.6	2.6	3.2
Resolution (Å)	3.5 (3.72–3.5)	2.07 (2.14–2.07)	1.65 (1.71–1.65)
Data cutoff (σ)	−3	−3	−3
Completeness (%)	65 (66)	97.0 (97.9)	93.7 (95.3)
Redundancy	1.5	1.85	2.7
R_{merge}‡ (%)	15.3 (32.2)	10.0 (34.4)	6.1 (69.7)
Average $I/\sigma(I)$	5.8 (2.5)	7.4 (1.7)	15.2 (1.8)
No. of refinement reflections	27314	348493	159427
Refinement resolution (Å)	7.0–3.5	50–2.08	50–1.65
Reflection cutoff (σ)	0	0	0
R_{cryst} (%)	34.2§	29.0	20.7
R_{free} (%)	—	30.5	22.5
No. of waters	—	1219	487
Average B factors (Å^2)			
Protein	—	26.9	25.2
Inhibitor (Å^2)	—	21.9	50.0
Waters (Å^2)	—	30.4	37.2
R.m.s.d. bonds (Å)	—	0.006	0.007
R.m.s.d. angles (°)	—	1.3	1.6
R.m.s.d. $B_{main\ chain}$ (Å^2)	—	1.0	1.3
R.m.s.d. $B_{side\ chain}$ (Å^2)	—	1.7	1.9

† Matthews coefficient (Matthews, 1968). ‡ $R_{merge} = \sum|I - \langle I\rangle|/\sum|\langle I\rangle|$; negative intensities are included as zero. § Full structure not refined.

accounted for the unfitted electron density. This result suggested that the crystal was merohedrally twinned, with a twin fraction close to one half, and this conclusion was confirmed by analysis of the native Patterson map. The true symmetry of the crystal was therefore $P4_1$ with four dimers per asymmetric unit. The R factor for four dimers in a $P4_1$ cell with one pair of dimers in two half-occupied positions was 34.2%. No further refinement of the structure was performed because of the limited resolution, the incompleteness of the data set and the difficulty of accurately deconvoluting the twinned data.

The $C2$ structure was solved by coordinate transformation of a set of crystallographic fourfold-related ordered dimers from the $P4_1$ unit cell into the $C2$ cell. A translation search using *AMoRe* (Navaza, 1994) (with a fixed contribution to F_{calc} values from the ordered dimers) was then performed to locate the disordered partially occupied positions for the four dimers. Calculation and peak searching of native Patterson maps for the CnTS structures and statistical analysis of the intensity distributions were performed using *CNS* (Brünger *et al.*, 1998).

2.4. CnTS structure refinement

The structure of CnTS was refined against the 2.08 Å $C2$ data with *CNS*, utilizing noncrystallographic symmetry (NCS) restraints (Brünger *et al.*, 1998). Initially, only the ordered dimers were included in the model. The sequence for CnTS, except for the N-terminal 12 residues, was fitted to $(2F_o - F_c)\alpha_{calc}$ and $(F_o - F_c)\alpha_{calc}$ maps using *O* (Jones *et al.*, 1991). Density for residues 1–12 was not discernible; presumably these residues are disordered. The ligands dUMP and CB3717 were also fitted to density in each active site. All TS protomers were tightly restrained during refinement to have equivalent structures, as was consistent with the electron density. The crystallographic R factor and R_{free} of a structure consisting of only the four refined ordered dimers were 37 and 38%, respectively.

Copies of the refined set of four ordered dimers were then translated to the four partially occupied sites identified by molecular replacement. The four partially occupied dimer sets were refined as alternate conformations in *CNS*, so that nonbonded interactions between them would be ignored. The position of each disordered set was refined by rigid-body minimization in *CNS* and an overall occupancy was also refined. Refinement proceeded provided all protomers were tightly restrained to have the same structure and B factors. Waters were fitted to electron-density peaks in favorable stereochemical environments. The density in the disordered region of the density maps was not interpretable at atomic resolution; thus, all manual rebuilding was performed only in the ordered regions of the map and the structures of the disordered molecules were updated to reflect manual adjustments to the ordered molecules. The final structure, consisting of four ordered dimers of CnTS and 16 partially occupied dimers, had R = 28.73% and R_{free} = 30.57% (see Table 2 and supplementary material).

2.5. Crystallization and data collection of EcTS–dUMP–GA9

Production of EcTS and crystallization of the apoenzyme have been described (Perry *et al.*, 1990). Cubic crystals of EcTS (space group $I2_13$, unit-cell parameter a = 133 Å) were soaked in a solution containing dUMP and GA9. Soaking the ligands into the apoenzyme crystals converted the crystals to rhombohedral symmetry, with space group $R3$, unit-cell parameters a = 186.6, c = 114.2 Å, and two dimers of the EcTS–dUMP–GA9 complex in the asymmetric unit. Crystals were transferred to cryosolvent prior to data collection at 103 K on a CCD detector at the Advanced Light Source, beamline 5.01. Data reduction and scaling using *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997) yielded 94% complete 1.65 Å data with an overall R_{merge} of 6.1% and an $\langle I/\sigma(I)\rangle$ of 15.2 (Table 2 and supplemental material). Although the data in the highest resolution bin have a high R_{merge}, these data also have a significant $I/\sigma(I)$ (1.8) and are complete (95.3%). Since a significant number of reflections (25%) in the 1.78–1.65 Å shell were observed [$I > 3\sigma(I)$], it was useful to include this shell of data in refinement, even though the effective resolution of the structure is arguably closer to 1.78 Å.

2.6. Structure solution and refinement of EcTS–dUMP–GA9

The structure of EcTS–dUMP–GA9 was solved by molecular replacement with *AMoRe* (Navaza, 1994) using the

coordinates for apoEcTS (Stout *et al.*, 1998; PDB code 1tjs) as a search model. Four protomers (two dimers) were located in the asymmetric unit. One molecule of GA9 and one molecule of dUMP were built into $(F_o - F_c)\alpha_{calc}$ difference density at each of the four active sites in the asymmetric unit and water molecules were fitted to difference density peaks in stereochemically reasonable environments. Several other solvent molecules, including ten molecules of glycerol and 12 molecules of phosphate, and alternate conformations for several side chains in each protomer were also fitted to difference density. Cys-ec146 and/or Cys-ec168 at three of the four active sites in the crystal appeared to have been modified by addition of β-ME and this modification was modeled into density with partial occupancy. In addition, strong density near the side chain Cys-ec50 in each protomer, which was modeled as a glycerol, may have been some (unidentified) covalent modification of the cysteine sulfhydryl. The structure was refined against a maximum-likelihood residual with *CNS* using all data to 1.65 Å (Brünger *et al.*, 1998). Occupancies of dUMP molecules, GA9 molecules, glycerol molecules, phosphate ions, β-ME modifications and disordered side chains were refined. Final occupancies for the inhibitors in the four protomers were 0.82–0.86 and occupancies for dUMP were 0.44–0.48. The final structure had $R = 20.7\%$ and $R_{free} = 22.5\%$ (see Table 2 and supplementary material).

2.7. Structure comparisons

Conformational differences between different species of TS were identified using difference distance matrices produced with the *DDMP* program from the Center for Structural Biology at Yale University, New Haven, CT, USA. Structures were superimposed by overlapping their structural cores (the largest set of contiguous residues that showed no structural differences between the two structures) as identified using the *NEWDOME* program written by Eric Fauman.

To analyze the differences in protomer packing at the dimer interface of CnTS compared with human TS, we first superimposed the CnTS and hTS dimers by least-squares fitting of the common structural cores of one protomer. We then calculated the angles between the respective principal axes of the common structural cores of the second protomers using *GEM* written by Eric Fauman.

Buried surface areas at crystal contact interfaces were determined by comparing the accessible surface area of isolated dimers to the accessible surface area of the same dimers in the crystal structure. Surface accessibility to a probe sphere of radius of 1.4 Å was calculated with the *CCP*4 program *SURFACE* (Collaborative Computational Project, Number 4, 1994). Surface accessibilities to a probe sphere of 2.8 Å were also calculated to estimate the extent of crystallographic interfaces at which there is a layer of bridging waters.

Figures were produced with *BOBSCRIPT* (Kraulis, 1991; Esnouf, 1997; Merritt & Bacon, 1997; Figs. 6, 7, 8, 9, 10*a* and 12), *PyMOL* (Figs. 2*a*, 2*b*, 3, 4, 11, 13 and 14). The sequences in Fig. 9 were aligned using *CLUSTALW* and then edited by hand so that sequence inserts were consistent with the inserts identified by alignment of the three-dimensional structures.

3. Results

3.1. Two crystal forms of CnTS–dUMP–CB3717 both have disordered packing

We crystallized CnTS–dUMP–CB3717 in two crystal forms: a tetragonal form, which diffracted to 3.5 Å, and a centered monoclinic form, which diffracted to 2.1 Å. The monoclinic unit-cell axes are related to the tetragonal unit-cell axes by $c' = c$ and $a' = b' = 2^{1/2}a$ (the diagonal of the tetragonal cell; Fig. 4). Both crystal forms have a similar type of statistical disorder. In both, half the molecules in the asymmetric unit are in fully occupied positions that conform to $P4_12_12$ symmetry. The other half of the molecules are translated with respect to the fully occupied sites by one of two (in the tetragonal cell form) or four (in the monoclinic cell form) slightly different translation vectors; the crystals are composites of crystalline domains with these slightly different packing arrangements.

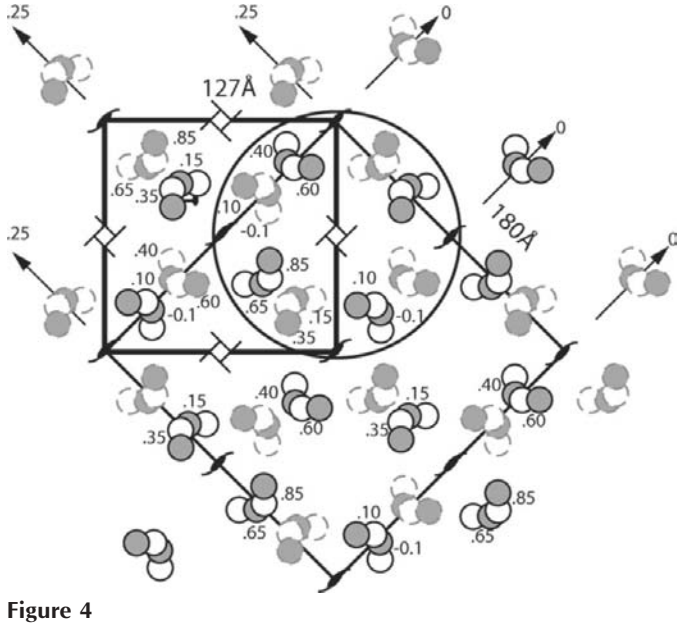

Figure 4
Diagram of the statistical average of the unit cells in the crystal, illustrating the relationship between the tetragonal and monoclinic crystal forms. The two crystal forms have pseudo-$P4_12_12$ and pseudo-$C2$ symmetry, respectively. The *c* axes are the same dimension for both space groups and are perpendicular to the page. The crystallographic pseudo-twofold and fourfold screw axes in the $P4_12_12$ space group are shown by arrows and winged rectangles, respectively. In the diagram, the *a* and *b* axes of the monoclinic cell are translated +1/4 in *z* with respect to the *a* and *b* axes of the tetragonal cell (the *C*2 pseudo-twofolds are in the *ab* plane, parallel to *b*). The paired gray and white spheres schematically depict the paired twofold symmetric TS dimers boxed in the unit-cell plot in Fig. 5 (gray spheres = monomers at higher *z*). Half of these, indicated by fainter color and dashed-line boundaries, are translated as a unit in the crystal by $z = \pm 0.023$ from the position conforming to $P4_12_12$ symmetry; in the *C*2 cell, they are fourfold translationally disordered. The fractional *z* coordinates of the active sites of selected monomers are shown. The asymmetric unit of the *C*2 cell is circled.

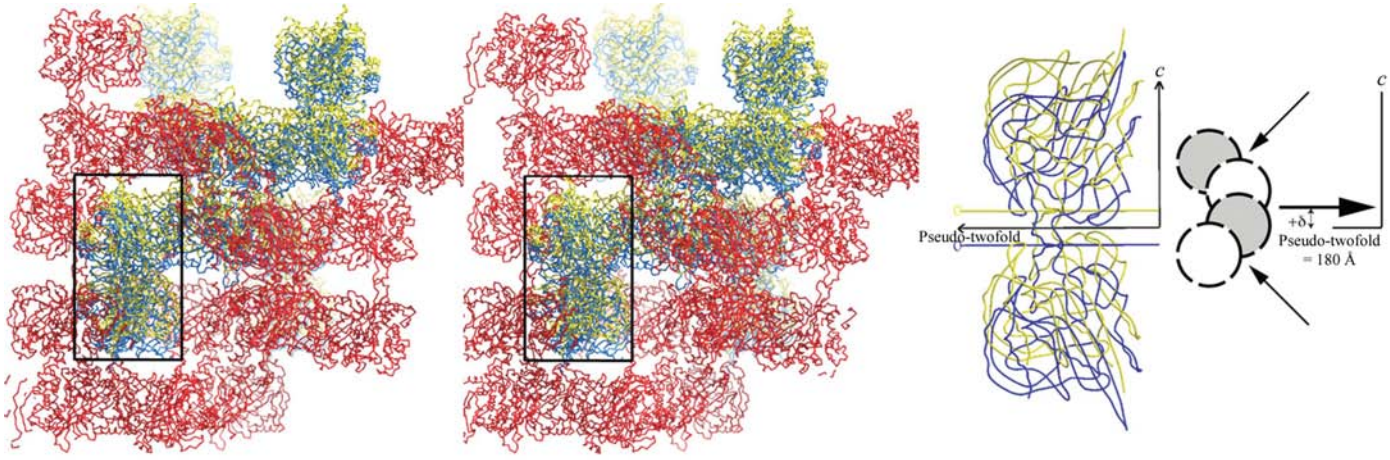

Figure 5
Divergent-eye stereo plot of C^{α} tracings of molecules in the $P4_1$ crystals of CnTS, illustrating the crystal packing. The red 'sublattice' formed by the ordered halves of the asymmetric units is interdigitated alternately with the translationally related yellow or blue 'sublattices' which are formed by disordered halves of the asymmetric units. The disordered half of one asymmetric unit is boxed and shown as an expanded view relative to the crystal *c* axis and to the unit-cell diagonal (180 Å), which is the pseudo-twofold axis. The boxed molecules are dimers of dimers and have local twofold axes relating protomers in a dimer and at the dimer–dimer interface. The schematic representation of a dimer of dimers used in Fig. 4 is also shown, with local twofolds indicated by arrows.

The tetragonal crystal form is best described as a hemihedral twin with pseudo-$P4_12_12$ symmetry. The asymmetric unit consists of four equivalent symmetric dimers. Two dimers at fully occupied sites are related to each other by a pseudo-twofold crystallographic axis coincident with the diagonal of the *ab* plane at $z = 1/4$. The other two dimers are related to the first pair by a translation of 0.5 in x, 0.5 in y and ±0.023 in z; thus, the local twofold relating these dimers to each other is displaced by ±0.023 from the crystallographic pseudo-twofold axis. As expected, the native Patterson shows a strong peak, with a maximum peak height about 30% of the origin peak height, at (x, y, z) = 0.5, 0.5, 0.022 (see supplementary material). A second packing peak at (x, y, z) = 0.0, 0.0, 0.054 had a maximum peak height only 2.5% of the origin peak height, indicating that interactions between dimers at the two alternate positions (+0.023 *versus* −0.023 in z) are infrequent. Therefore, the packing variations are not randomly distributed throughout the crystal, nor do they alternate in a regular way to form a supercell; rather, the crystal is twinned, with a twin fraction close to 0.5. Although the overall second moment of the intensities is typical of an untwinned crystal ($\langle|I|^2\rangle/\langle|I|\rangle^2 = 2.1$), the second moment of the intensities for the 3.8–3.5 Å shell is 1.7, which is close to the value of 1.5 characteristic of perfect twin crystals. The twin fraction estimated from the distribution of $\langle H\rangle$, where $H = (I_1 - I_2)/(I_1 + I_2)$ and I_1 and I_2 are the intensities of reflections related by the twinning operation, is 0.34 (Yeates, 1988). The basis for the twinning is clear from the crystal-packing diagram (Fig. 5). The translationally related halves of the unit cell form identically packed interdigitated 'sublattices' that are nearly independent of one another; the few contacts between sublattices are nonspecific and identical in the alternate packing arrangements in the crystal.

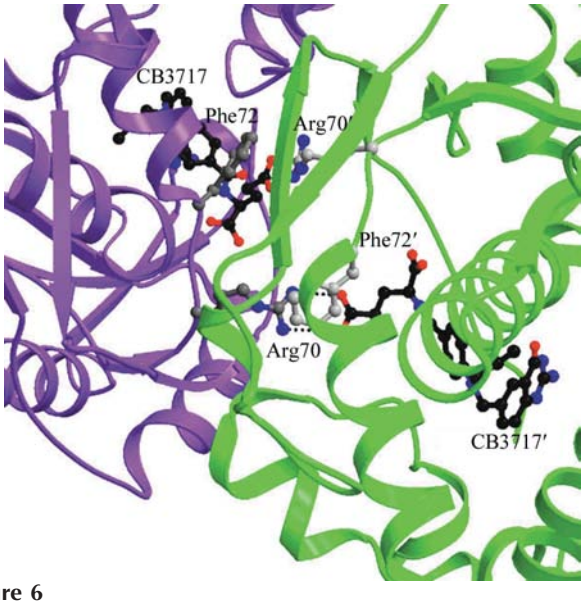

Figure 6
Illustration of one of the two major packing interfaces in the CnTS crystals. The interface is formed by packing together of the polyglutamyl-binding grooves of two CnTS molecules that are related by crystallographic symmetry. The two crystallographically related dimers are plotted with purple and green ribbons, respectively. The glutamyl moiety of CB3717 (plotted in black ball-and-stick form) is hydrogen bonded across the interface to Arg70i (plotted in gray) from the crystallographically related molecule. Contributing to the interaction is Phe72 (plotted in gray), the phenyl ring of which stacks against the CB3717 glutamyl-carboxyl and Arg70 guanidinium group.

Besides the few contacts between molecules in each of the two 'sublattices', there are only two interdimer interfaces in the crystal. Face-to-face packing of dimers at the binding sites for the polyglutamyl moiety of the cofactor forms the first type of interface. Fig. 6 illustrates the interface in detail, showing the twofold-related hydrogen bonds between the glutamyl moiety of CB3717 at one site and the guanidinium group of Arg-cn70′ of the symmetry-related binding site. Phe-cn72, which provides the major contact with the monoglutamyl moiety of CB3717, also stacks against the symmetry-related

Arg-cn70′ guanidinium group at a distance of approximately 3.5 Å. 15 residues contribute to the interface and the total buried surface area is ∼950 Å^2. Packing together of the inserts between helices *G* and *H* (labeled in Fig. 8) from twofold-related dimers generates the second type of interface. This interface is more extensive than the first, involving ∼1341 Å^2 of buried surface area and 24 residues. These two interfaces are conserved in the monoclinic crystal form. The conformations of the interface regions are not significantly different from the conformations of the corresponding regions in other eukaryotic TSs. Therefore, crystal contacts do not perturb the structure.

The monoclinic cell has *C*2 symmetry, with an asymmetric unit comprised of eight equivalent symmetric dimers. The four dimers in fully occupied sites were positioned by coordinate transformation of a fourfold-related set of ordered dimers from the $P4_1$ cell. A translation search located the other set of four dimers in four quarter-occupied positions, related to the fully occupied sites by fractional coordinate translations of 0.5 in *x*, ±0.02 in *y* and ±0.03 in *z*. The translational disorder in *y* (perpendicular to what would have been the fourfold axis in the tetragonal cell) accounts for the breakdown of fourfold symmetry in the crystallographic data and is the only feature that distinguishes the $P4_1$ and *C*2 crystal forms.

We carefully analyzed the *C*2 diffraction patterns for weak reflections that had not been indexed, which might indicate a larger unit cell, but found none. The distribution of the second moments of the intensities for the *C*2 data did not indicate twinning. However, analogously to the $P4_1$ crystal form, the native Patterson showed a peak whose maximum peak height was approximately one eighth of the peak height of the origin peak at (0.5, 0.0199, 0.0319) and only a very weak second packing peak, with peak height 2.6% of the height of the origin peak, at (0.0, 0.0, 0.0646) (see supplementary material). Therefore, the monoclinic crystals are most likely to be comprised of multiple domains with slightly different packing between translationally related components of the asymmetric unit and the twofold axis in the *C*2 cell is a pseudo-twofold arising from averaging intensities of these domains. The additional disorder in *y* increases the contacts between the interdigitated 'sublattices' in the *C*2 cell relative to the $P4_1$ cell, accounting for the better resolution of the *C*2 data. Since there was not an obvious way to separate the intensities contributed from individual domains, the structure was refined in space group *C*2 as a statistical average of the four observed packing arrangements. The relatively high crystallographic *R* factors of $R = 28.6\%$ and $R_{\rm free} = 30.5\%$ probably result from inadequate modeling of crystal disorder. However, density for the ordered region of

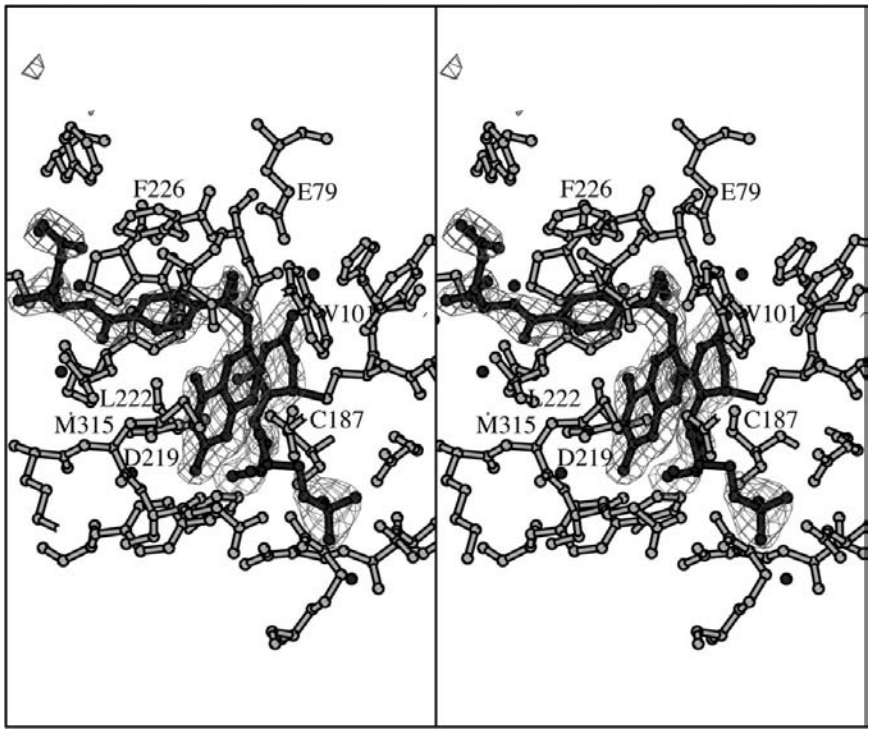

Figure 7
Sections from a 2.08 Å $(F_o - F_c)\alpha_{calc}$ omit electron-density map, contoured at the 3σ level, at the active site of one protomer from the *C*2 crystal form of CnTS, shown in divergent-eye stereo. dUMP and CB3717 were omitted from refinement and calculation of phases.

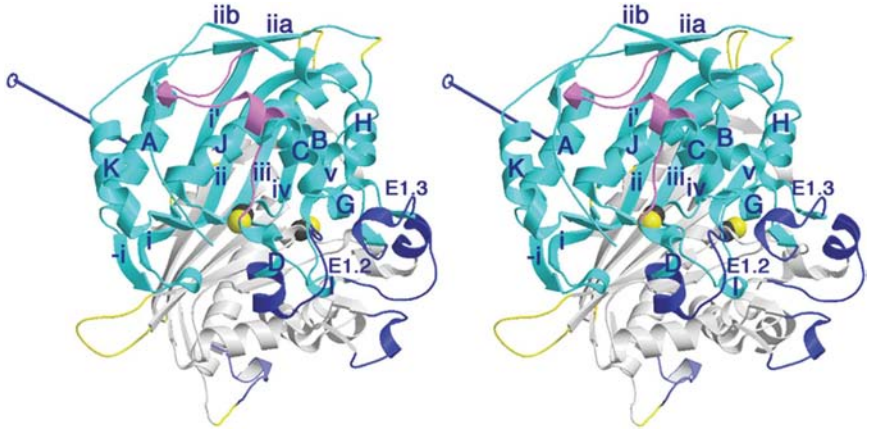

Figure 8
Ribbon plot of one dimer of CnTS, with helices and β-strands in the cyan protomer with labels defined by Montfort *et al.* (1990). Strand i is interrupted by a one-residue insert in CnTS and other fungal TSs and the two parts of this strand are labeled i and i′. The active-site cysteines are shown as space-filling atoms. Inserts conserved in eukaryotic TSs but not in prokaryotic species are shown in dark blue. Inserts unique to fungal TSs or found in fungal and protozoan but not mammalian TSs are in yellow. 'Plastic' regions containing residues that can adjust their conformations to bind a variety of antifolates are shown in pink. The molecular twofold axis is shown by a blue line above the ribbon plot.

the crystal structure was excellent (Fig. 7), indicating that the molecular structure we report is reliable. The descriptions of the CnTS structure that follow refer to the molecules in the fully occupied sites of the *C*2 crystal structure.

3.2. CnTS has the same fold as other eukaryotic TSs, with subtle differences at the dimer interface

CnTS is the seventh eukaryotic TS whose structure has been reported. In addition to the CnTS structure reported here, ternary complex crystal structures of the TSs of rat (Sotelo-Mundo *et al.*, 1999), *Pneumocystis carinii* (Anderson *et al.*, 2000, 2001) and human (Almog *et al.*, 2001; Phan, Koli *et al.*, 2001; Sayre *et al.*, 2001), as well as the bifunctional DHFR–TSs from the protozoa *Leishmania major* (Knighton *et al.*, 1994), *Plasmodium falciparum* (Yuvaniyama *et al.*, 2003) and *Cryptosporidium hominis* (O'Neil *et al.*, 2003), have been published. All of these structures are obligate dimers with the same overall fold as in bacterial TSs (Fig. 8). Three inserts distinguish the eukaryotic TSs (Fig. 9). One is an N-terminal extension that in protozoa connects to the DHFR portion of a DHFR–TS bifunctional enzyme. As is the case for the other monofunctional eukaryotic TS crystal structures, only a few residues of this extension are visible in electron-density maps of CnTS; residues 1–12 are disordered. A second feature of eukaryotic TSs is an insert following helix *C*, residues 111–128 in hTS (Fig. 9), which is conserved in sequence, size and fold and which forms a small domain very different from the highly variable small domains of non-eukaryotic TSs. The third insert, in the loop between helices *G* and *H*, folds back onto helix *G* and makes hydrophobic contacts with both the small domain and helix *G* *via* two highly conserved tyrosines. In CnTS, all residues in the second and third inserts are clearly visible in electron-density maps and have the expected conserved structure.

CnTS has four inserts that are not conserved in eukaryotic TSs. These include two one-residue inserts, Pro-cn52, interrupting β-strand i, and Asn-cn61 in the loop following it, which are present in other fungal TSs, a nine-residue insert in the loop connecting β-strands iii and iv (residues cn198–cn206) and a three-residue insert in the long C-terminal tail (residues cn286–cn288). None of these inserts are in sequences of the protein that contribute to the active-site cavity and they do not affect the structure of the conserved core of the protomer. However, three of the four inserts are either in or at the edge of the dimer interface and may be responsible for the altered packing of the protomers relative to non-fungal species. The packing differences are small at the active site and propagate to larger differences at the edges of the interface. Relative to human TS, the difference in protomer packing can be described as a 1.5 Å shift and 4° twist of one protomer with respect to the other, as well as a ~6° rotation of one protomer about an axis in the plane of its β-sheet, which slightly opens one side of the dimer interface (Fig. 10*a*).

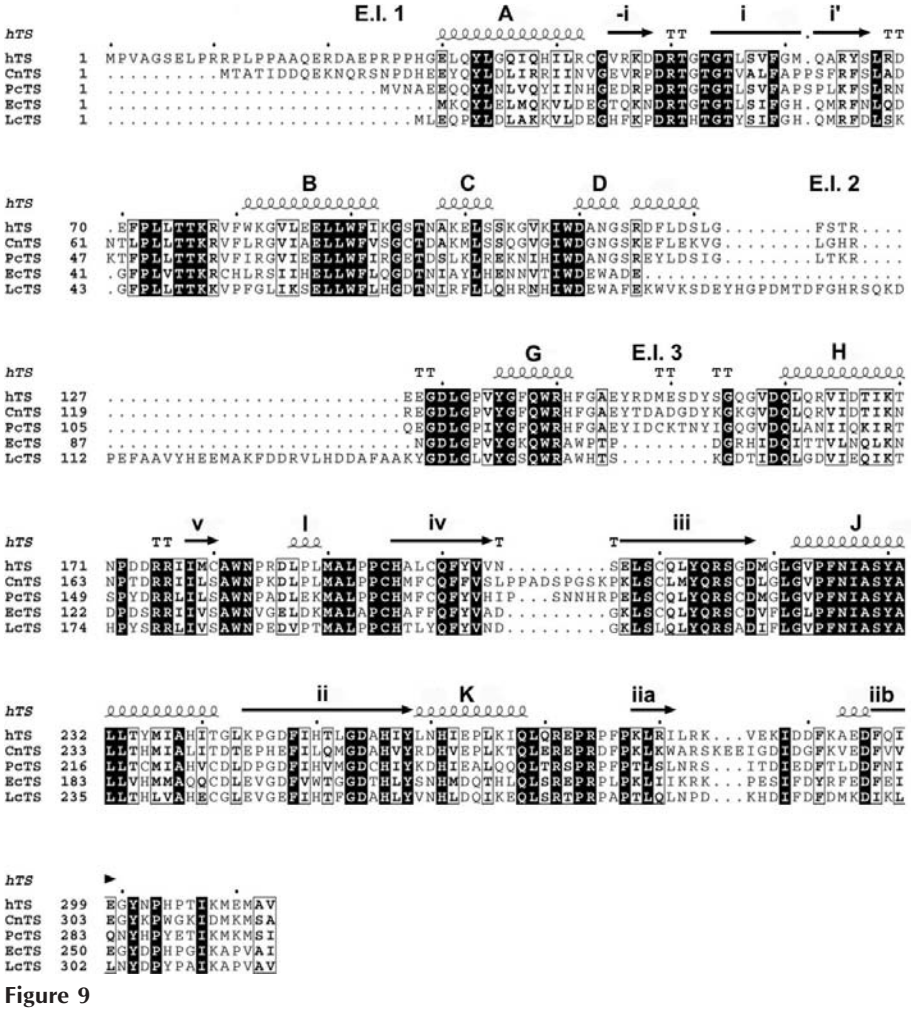

Figure 9
Aligned sequences of hTS, CnTS, PcTS, EcTS and LcTS. Conserved residues are printed with white letters. Secondary-structure elements, as derived from hTS coordinates (PDB code 1ju6), are indicated above the sequences with coils for helices and arrows for β-strands and are labeled as in Fig. 6. Sequence numbers are shown on the left side of each line.

3.3. CnTS–dUMP–CB3717 is in the closed conformation

The crystal structures of bacterial TS ternary complexes with substrate and cofactor analogs have all been in closed enzyme conformations (Stroud & Finer-Moore, 2003). Complexes of eukaryotic TSs with cofactor analogs have crystallized in both open and closed conformations or in a partially open conformation in which the C-terminus has moved into the active site but other protein segments have not shifted and there is no covalent bond between dUMP and the catalytic thiol (Stroud &

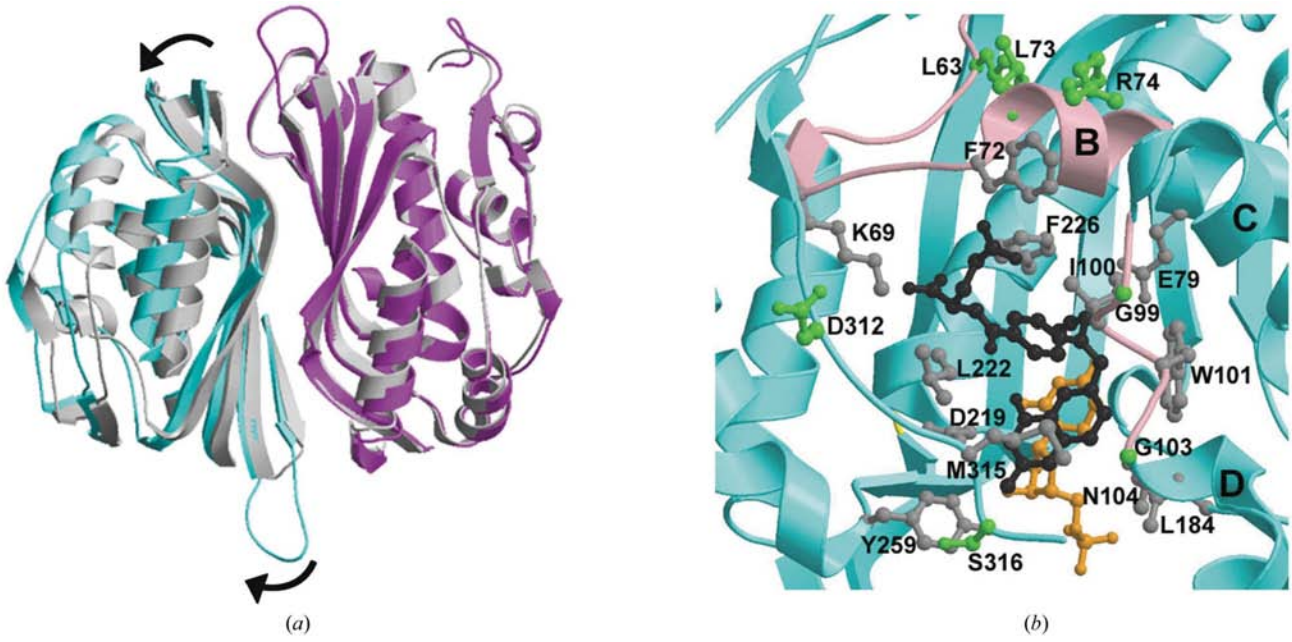

Figure 10
(*a*) Ribbon plot showing the β-sheet dimer interface of CnTS (magenta and cyan ribbons) overlapped with that of hTS (gray ribbons). The interfaces of the two TS species overlap closely at the active-site cavity, but at the edge of the sheet adjacent to helix *A* (shown) the protomers are rotated apart in CnTS relative to human TS, as indicated by the arrows. (*b*) Ribbon drawing showing the CnTS active site with CB3717 (black), dUMP (gold) and CB3717-binding residues (gray) plotted in ball-and stick form. The CB3717-binding residues are conserved across all eukaryotic species, but some bordering residues or residues in the polyglutamate-binding groove, plotted in green, are different in CnTS and human TS, offering an avenue for drug design.

Finer-Moore, 2003). The conformation of the complex in the crystal structure appears to be a function of the inhibitor and of crystallization conditions (Almog *et al.*, 2001; Phan, Koli *et al.*, 2001). Difference distance matrices of CnTS–dUMP–CB3717 compared with other TS ternary complex structures show unequivocally that CnTS is in a fully closed conformation and density maps clearly show there is a covalent bond between the 6-position of dUMP and the catalytic thiol. We therefore compare ligand binding in the CnTS structure with other closed ternary complexes.

3.4. Substrate and cofactor binding residues of CnTS are highly conserved, but the cofactor-binding site of CnTS is less constrained than that of EcTS and hTS

The ligand–protein contacts in the CnTS ternary complex are very similar to those seen in the closed ternary complexes of other TS species. As expected, CnTS makes all of the conserved hydrogen bonds with dUMP; the highly conserved dUMP-binding site offers few avenues for species-specific drug design.

The binding sites for the pterin and pABA moieties of the cofactor are lined with residues that are completely conserved with respect to hTS and highly conserved with respect to EcTS. The only cofactor-binding residues that are not conserved between CnTS and EcTS are Met-cn315 and Asn-cn104. These two residues are conserved among all eukaryotic TSs, while in bacterial TSs Asn-cn104 is substituted by Trp and Met-cn315 is usually substituted by Val (Fig. 9). The penultimate residue in the protein, Ser-cn316, is a serine in other fungal TSs and in *B. subtilis* TS, but is an Ala in all other TSs, including hTS. Although the side chain of this residue faces

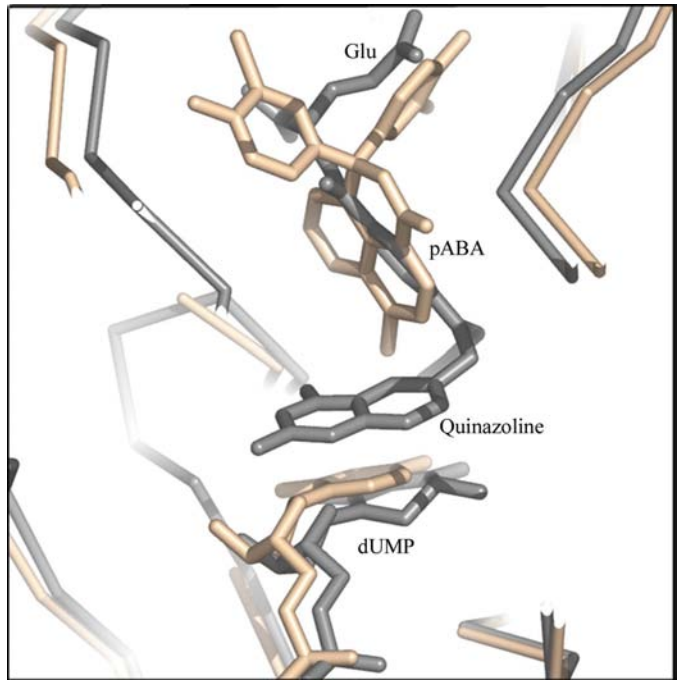

Figure 11
Comparison of bound 5-fluoro-2′-deoxyuridine 5′-monophosphate (FdUMP) and CH_2H_4folate in EcTS–FdUMP–CH_2H_4folate (gray) with dUMP and GA9 bound to EcTS–dUMP–GA9 (tan) after the two complexes were superimposed by aligning their structural cores.

away from the active-site cavity in CnTS, a different side-chain rotamer would put the serine hydroxyl in a position to hydrogen bond to an inhibitor in the active site. Thus, this residue has the potential to provide a unique interaction with a fungal-specific inhibitor.

Folates and antifolate inhibitors are converted intracellularly into polyglutamylated forms that bind to TS with higher affinity and prevent efflux from cells (Shane, 1989). The polyglutamate moiety (Fig. 1) binds to a shallow groove on the enzyme surface defined by the C-terminal tail, the loop connecting helix *C* and helix *D* and the so-called folate-binding loop leading into helix *B* (the *C*–*D* loop and folate-binding loop are colored pink in Figs. 8 and 10*b*). The residues lining the groove are less conserved than the residues lining the rest of the active-site cavity (Kamb *et al.*, 1992). In structures of EcTS bound to polyglutamylated antifolates, the polyglutamate contacts were largely water-mediated hydrogen bonds involving Lys-ec48, Arg-ec49, His-ec51, Arg-ec53, Ser-ec54, Asn-ec76, Thr-ec78 and Lys-ec259 (Kamb *et al.*, 1992; Sayre et al., 2001). The corresponding residues in CnTS are Lys-cn69, Arg-cn70, Phe-cn72, Arg-cn74, Gly-cn75, Gly-cn97, Gly-cn99 and Asp-cn312 (Fig. 9). Three of these residues are not conserved between CnTS and hTS: Arg-cn74 is a Lys, Thr-cn99 is a Gly and Asp-cn312 is a Lys in the human enzyme.

There are several residue substitutions in CnTS relative to hTS that lie in flexible segments surrounding the folate-binding site. These substitutions may indirectly affect inhibitor selectivity by altering the potential for ligand-induced conformational change. Folate-induced conformational changes maximize binding interactions with the protein and are essential for high binding affinity. The loop connecting helices *C* and *D* contains three binding residues for folates or their analogs: Ile-cn100, Trp-cn101 and Asn-cn104. Adjacent to these important binding residues is the CnTS-specific residue Gly-cn99, as well as Gly-cn103, which are Lys and Ala, respectively, in hTS and Thr and Glu in EcTS. A greater range of conformations is available to glycines than to residues with side chains. In addition, the residue in hTS that is equivalent to Gly-cn99, Lys-h107, donates a hydrogen bond from its N^{ζ} atom to Ser-h102 O, which adds additional rigidity to this segment in hTS. There are also some substitutions of smaller residues in CnTS at the interfaces between helix *B* and the *C*–*D* and folate-binding loops that may allow the loops more freedom to adapt their conformations to new antifolates. For example, Ala-cn78 is a Glu in hTS and His in EcTS and two residues at the helix *B*–folate loop interface, Leu-cn63 and Leu-cn73, are Phe and Trp in hTS. The latter interface residues are close-packed in hTS (shortest distance, 3.2 Å) but not in CnTS (closest distance, 4.5 Å).

3.5. GA9 binds to the open conformation of EcTS at the pABA-glutamyl binding site

GA9 preferentially inhibits CnTS (as well as EcTS and LcTS) over hTS and therefore affords the opportunity to investigate novel strategies for developing drugs with specificity for pathogenic species over humans. In order to understand the basis for selective inhibition of bacterial TSs and CnTS over hTS by GA9, we determined the crystal structure of the EcTS–dUMP–GA9 ternary complex formed by soaking dUMP and GA9 into crystals of the apoenzyme. We aimed to then use this structure to deduce the basis for selective inhibition of CnTS. This strategy was based on the high conservation of the TS binding site and the similar K_i values for EcTS and CnTS. Crystal structures of several TS ternary complexes with dUMP and antifolates have shown that these inhibitors bind analogously in *E. coli* and eukaryotic TSs (Anderson *et al.*, 2001; Sayre *et al.*, 2001) and human cancer drugs have been successfully designed using crystal structures of EcTS complexes to model inhibitor–hTS interactions (Webber *et al.*, 1993).

Binding of the ligands to the apoEcTS crystals converted the crystal symmetry from the cubic space group $I2_13$, with one protomer of the dimer in the asymmetric unit, to the rhombohedral space group *R*3 with two dimers in the asymmetric unit. Although the crystal symmetry changed, the conformation of the ligand-bound enzyme was essentially identical to that of apoEcTS. The overall r.m.s.d.s between the positions of the C^{α} atoms of the aligned apoEcTS and EcTS–dUMP–GA9 structures is 0.4 Å.

The refined occupancies and relative bound positions of dUMP and GA9 were the same within experimental error in all four protomers in the asymmetric unit. The binding mode of dUMP was the same as in other substrate complexes of TS; however, the occupancy of dUMP was reduced to ∼0.45 because of partial oxidation of Cys146 and/or Cys168 in the active-site cavities. In TS ternary complexes with folates or antifolates, dUMP usually provides a binding surface for the pterin or quinazoline ring of the folate or inhibitor. In contrast, there are no close contacts between dUMP and GA9 in TS–dUMP–GA9. The tricyclic GA9 binds to the pABA-glutamyl binding site of the cofactor, with one bromophenyl group overlapping the binding site for the cofactor glutamate tail and the naphthalide moiety overlapping the pABA ring-binding site (Fig. 11). The second bromophenyl group is directed towards the C-terminal tail of the enzyme, preventing the C-terminus from moving into the active-site cavity. Both bromophenyl rings of the inhibitor are statistically disordered: each ring alternates between two orientations related by a 180° rotation about their $C^{\alpha}-C^{\beta}$ bonds. The binding mode of GA9 in the crystal structure is consistent with the fact that GA9 is a competitive inhibitor with respect to CH_2H_4folate (Costi *et al.*, in preparation).

The binding determinants for GA9 include residues in the helix *C*–helix *D* loop and the folate-binding loop, but there is no significant deformation of these flexible loops relative to the apoEcTS structure. The naphthalide moiety is in van der Waals contact with conserved residues Ile-ec79 and Leu-ec172 (Fig. 12), which are also major contact residues for antifolates. In each protomer, GA9 is in a crystal contact region and one of its bromophenyl hydroxyl groups makes a hydrogen bond with His-ec108i from a symmetry-related loop. The second bromophenyl hydroxyl group is hydrogen bonded to Ser-ec54 O^{γ} (Fig. 12).

4. Discussion

4.1. Variability in flexibilities of the cofactor-binding sites of different TS species as the basis for species-specific antifolates

Thymidylate synthase is a flexible molecule that relies on extensive conformational changes to tightly bind and orient its cofactor against its substrate, dUMP. The conformational changes coincide with a slow and energetically costly step in cofactor binding. For example, the free-energy barrier between the adsorption complex, in which cofactor is loosely docked to the TS–dUMP binary complex, and the final tight complex resulting from the conformational change is approximately 68 kJ mol^{-1} in *Lactobacillus casei* TS (Santi *et al.*, 1987). Thus, the cofactor apparently takes an active role in binding, using hydrophobic interactions with the protein to drive the protein into closed conformations that are probably not present to a significant extent in the TS–dUMP complex. Our structure of CnTS bound to substrate and the cofactor analog CB3717 provides further evidence that the cofactor-binding mechanism is highly conserved across species. Both the conformation of the active site and the identity of ligand-contact residues in CnTS–dUMP–CB3717 are nearly identical to those seen in hTS ternary complexes with cofactor analogs and very similar to those seen in ternary complexes of EcTS.

Crystallographic studies of TS bound to a variety of drug leads have shown that the same protein segments that envelop the cofactor as it binds can similarly conform to novel ring systems (Anderson *et al.*, 2001; Stout & Stroud, 1996; Weichsel & Montfort, 1995; Sayre *et al.*, 2001) (Fig. 2*a*). Novel inhibitors can be accommodated not only by the protein dynamic modes that open and close the active-site cavity, but also by local rearrangements in structure that optimize protein–inhibitor contacts. Perhaps the most striking example of how the inherent flexibility of TS allows it to bind novel inhibitors is the complex of EcTS with dUMP and *N*,*O*-didansyl-L-tyrosine (DDT). This inhibitor, which is somewhat similar in shape and structure to CH_2H_4folate but is much larger, containing three ring systems, binds to the pterin and pABA-glutamyl binding sites of EcTS by inducing rotamer changes in several active-site residues, local changes in the backbone conformations of the folate-binding loop and the helix *C*–helix *D* loop and rigid shifts in protein segments, which opened up the active-site cavity (Fritz *et al.*, 2001). In contrast to most antifolates, which are not species-specific, DDT selectively inhibits bacterial TS species and it was postulated that bacterial-specific residues, brought into contact with DDT by the induced conformational changes, were responsible for selectivity (Fritz *et al.*, 2001). Thus, one avenue to designing species-specific inhibitors of TS is to harness the flexibility of the cofactor-binding site to conform to novel molecules that access nonconserved residues bordering the active site.

Since conformational change is usually part of the binding mechanism of an inhibitor, species' differences in degree of flexibility will directly affect the inhibition profile. Thus, Taylor and coworkers discovered a viral sialidase inhibitor that inhibited the influenza virus type A sialidase with an ∼1000-fold lower K_i than for the highly homologous influenza virus type B sialidase (Taylor *et al.*, 1998). The inhibitor made identical protein contacts in both type A and type B sialidase complexes, but inhibitor binding required a conformational change in an active-site residue that was less energetically costly to make in the type A enzyme than in the more constrained type B enzyme (Taylor *et al.*, 1998). CnTS is distinguished from hTS by several residue substitutions that impart greater flexibility to cofactor-binding loops or decrease the number of van der Waals and hydrogen-bond contacts between these loops and neighboring parts of the proteins. Because of these differences, there may in principle be antifolates that preferentially bind CnTS over hTS, in spite of the high conservation of their active sites.

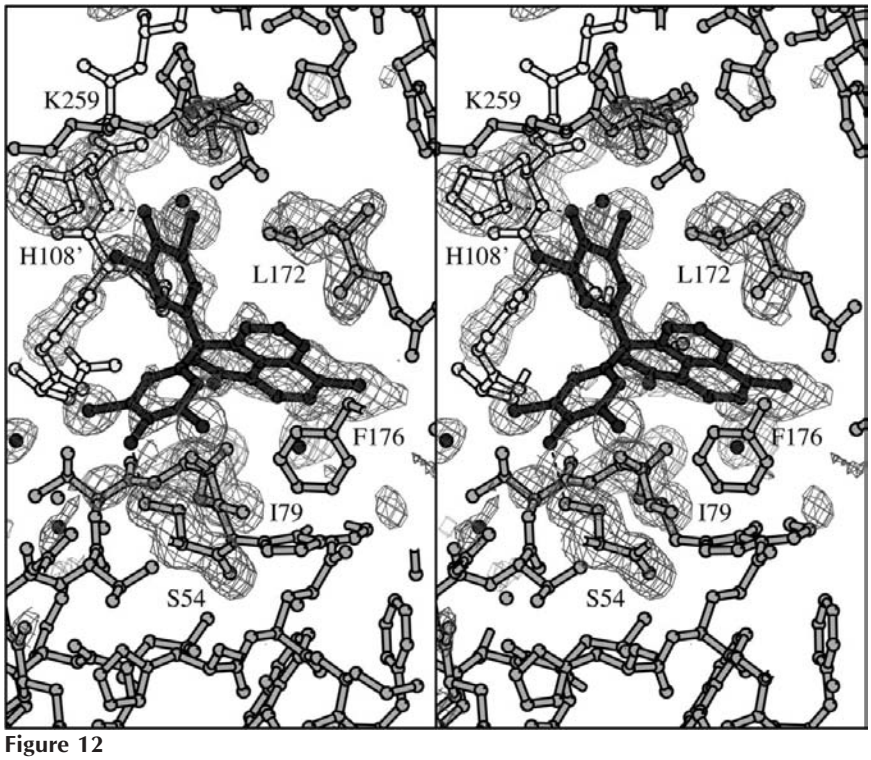

Figure 12
Binding site for GA9 in the crystal structure of EcTS–dUMP–GA9 superimposed on sections from a simulated annealing $(F_o - F_c)\alpha_{calc}$ omit map contoured at the 3σ level shown in divergent-eye stereo. GA9 and all residues with atoms within 3.5 Å of GA9 were omitted from simulated-annealing refinement and map calculation. Residues from a molecule related by crystallographic symmetry are plotted in light gray and have primed residue numbers.

4.2. Polyglutamyl-binding groove as a target for species-specific inhibitors

In an EcTS complex, GA9 overlaps the binding site for the pABA-glutamyl moiety of the cofactor. The site includes part of the relatively poorly conserved binding groove

for the polyglutamate tails of polyglutamylated folates and folate analogs, which is formed by the C-terminal tail, the 'folate-binding loop' and the loop connecting helices *C* and *D*. This region has long been recognized as a potential target for species-specific inhibitors (Kamb *et al.*, 1992). However, GA9 is the first species-specific TS inhibitor shown to bind to this site. Specificity of GA9 for EcTS could be enhanced by modifications that made additional specific contacts with the nonconserved residues in the groove.

There are a number of differences among TS species at the GA9-binding site that may influence the GA9 inhibition profile shown in Table 1 (Fig. 9). The only intramolecular hydrogen bond between a side chain and GA9 in the crystal structure is between Ser-ec54 O^{γ} and the hydroxyl group of one of the bromophenol moieties. The analogous residue in eukaryotic TSs, including both hTS and CnTS, is Gly, while in LcTS the analogous residue is Leu-lc56. Three residues away, His-ec51 is substituted by a Phe in eukaryotes and a Pro in LcTS. The substituted Leu and Phe side chains of apoLcTS and apohTS (or CnTS), respectively, overlap the GA9-binding site and would have to change conformation in order for the enzymes to bind GA9 in the same mode as in EcTS–dUMP–GA9. The second bromophenol group binds next to Lys-ec259 in the C-terminal tail, causing the side chain of this residue to change from its apoEcTS conformation. In hTS, PcTS and LcTS this residue is also a Lys, but in CnTS it is Asp-cn312, which forms an ion pair with Lys-cn310. Lys-cn310 is a Thr in hTS and PcTS, a Gly in EcTS and an Ala in LcTS. In general, each TS species we have analyzed would require some minor conformational adjustment, presumably at an energetic cost, in order to bind GA9 at the site defined by the EcTS–dUMP–GA9 complex. These changes do not appear to be more dramatic for hTS than for the other TS species; thus, the much higher K_i value for GA9 inhibition of hTS compared with the other species cannot be easily rationalized on the basis of the single binding mode seen in our structure. Most puzzling is the fact that although the residues lining the glutamyl binding groove are highly conserved between CnTS and hTS, GA9 only inhibits the former enzyme.

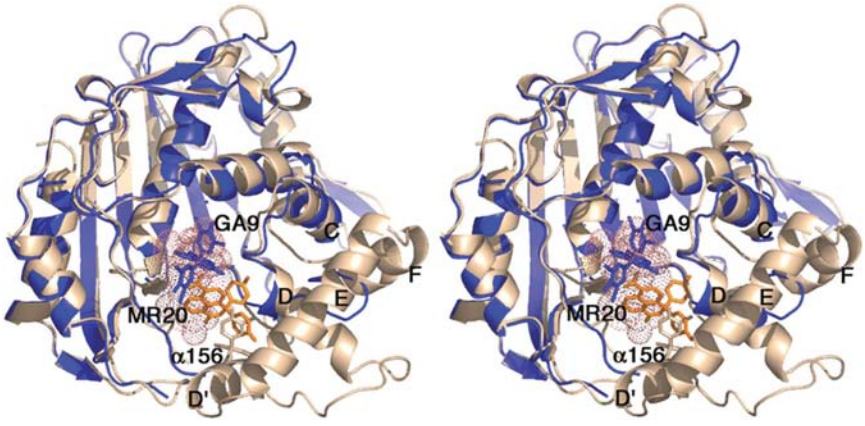

Figure 13
Divergent-eye stereoview of the overlaid ribbon drawings of the EcTS–dUMP–GA9 complex (blue) and the LcTS–α156 complex (gold) with stick representations for GA9 and α156. The inhibitor MR20 is also shown with orange stick bonds at its binding site in the LcTS–MR20 crystal structure. The space-filling spheres with dotted surface indicate the binding sites for dUMP and CH_2H_4folate. Helices in the variable small domain of TS are labeled. Helices *D*′, *E* and *F* and some of helix *D* are not present in EcTS or eukaryotic TSs, but are comprised of a long insert found in TSs from *L. casei* and a few other bacterial species. In crystal structures, MR20 and α156 bind against the small domain of LcTS.

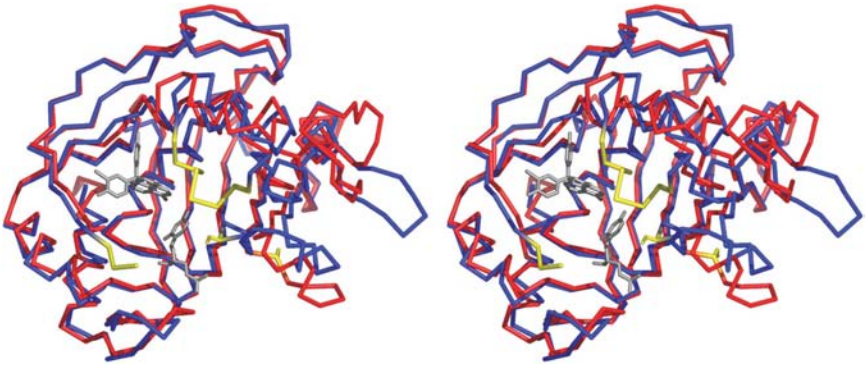

Figure 14
Divergent-eye stereo plot of the α-carbon tracings of two aligned crystal structures of apohTS (PDB codes 1hw3 and 1hzw), illustrating two different conformations of the enzyme. The active-site cysteines are shown with yellow sticks. Regions in the red structure that are unresolved in the blue structure are shown in yellow ribbon. dUMP and GA9 are plotted with gray sticks at binding sites derived from the EcTS–dUMP–GA9 structure.

4.3. Conformational heterogeneity in the target apoprotein as a possible basis for GA9 bioprofile

One possible explanation for the low affinity of GA9 for hTS is that the ligand-induced conformational changes required to optimize binding are less likely in hTS. The GA9-binding site is formed by some of the flexible cofactor-binding loops that we have argued are more constrained in hTS than in CnTS. Although GA9 binding does not induce protein conformational changes in the crystal of EcTS, we cannot rule out the possibility that GA9 binding induces conformational changes in other TSs. A second possible explanation is that the binding site(s) may vary from one species to another and/or there may be more than one GA9-binding site in any TS species. GA9 has relatively few hydrogen-bonding groups and a compact spherical shape with a roughly threefold symmetric arrangement of ring systems. Thus, it may be able to bind to several regions within the spacious hydrophobic active-site cavity. There is a single well ordered (except for the statistical disorder of the bromophenol rings) GA9 site seen in the EcTS–dUMP–GA9 structure. However, this site may be stabi-

lized by crystal contacts and multiple binding modes may exist in solution.

Crystallography and docking experiments indicate that other inhibitors from the diphenol-1,8-naphthalein class bind to LcTS using multiple binding modes (Stout *et al.*, 1999). In crystal structures, two diphenol-1,8-naphthalein derivatives, 3′,3′′-dichlorophenyl-1,8-naphthalein (α156) and 5-nitrodiphenol-1,8-naphthalein (MR20) (Figs. 1 and 2), bound near the variable small domain of LcTS, displaced from the cofactor and substrate-binding sites (Stout *et al.*, 1999; Fig. 13). Although these inhibitors were found at single binding sites in LcTS, their binding modes were judged to be nonspecific since mutation of protein contact residues did not alter binding affinity in a predictable way (Stout *et al.*, 1999). These inhibitors could be docked into the LcTS active site in multiple binding modes with similar energies. Sites predicted by docking included those seen in the crystal structures (near the small domain) and others overlapping the cofactor site, such as the one we see in the GA9 complex (Stout *et al.*, 1999). If GA9 has multiple binding modes in any one TS species, then the measured K_i values are apparent K_i values derived from ensembles of weakly bound GA9–TS complexes. In this case, the single binding mode seen in our crystal structure cannot by itself account for the inhibition profile.

A third possibility is that GA9 does not make enough specific contacts with the protein to stabilize a unique complementary enzyme conformation; rather, its binding affinity is dependent on the proportion of enzyme conformations with accessible binding sites, as proposed by Ferrari *et al.* (2003). The apo form of hTS is unusual because it is able to refold into an inactive conformation in which the loop containing the catalytic cysteine is rotated 180°, burying the catalytic cysteine so that it cannot react with substrate (Schiffer *et al.*, 1995; Phan, Steadman *et al.*, 2001; Almog *et al.*, 2001). This refolding, which was postulated to be part of a mechanism for regulating hTS activity (Phan, Steadman *et al.*, 2001), destroys the GA9-binding site seen in the EcTS–GA9 complex (Fig. 14). Thus, if this unusual disorder is unique to hTS, it may explain the aberrantly high K_i of GA9 for this TS species, since it effectively removes a subset of apohTS molecules from the pool available to bind GA9.

5. Conclusions

The high homology between the active-site cavities of CnTS and hTS complicate the design of CnTS-specific inhibitors that might be useful drugs for treating *C. neoformans* infections. Nevertheless, we have previously shown that specific inhibitors can be made. The analysis of the CnTS and EcTS complex crystal structures presented here suggest two mechanisms by which specific inhibition can occur. Firstly, sequence differences in residues bordering the cofactor-binding site may make ligand-induced conformational changes easier in CnTS than in hTS. Secondly, differences in the ensembles of conformations adopted by the CnTS and hTS apoenzymes could result in different overall K_i values for compounds like the diphenol-1,8-napthaleins, which make few specific contacts with the protein and do not induce protein conformational changes.

These observations show how a 100-fold difference in ligand affinity is encoded in dynamics of the target protein. Correspondingly, it is axiomatic that any computational scheme to advance structure-based drug discovery cannot predict differences in affinity of this magnitude unless the flexibility of the target is built into the premise for the docking trials. To an important degree, this emphasizes the importance of current efforts to incorporate target flexibility into docking algorithms (Carlson, 2002; Ferrari *et al.*, 2004; Halperin *et al.*, 2002). Without this feature such algorithms cannot correctly rank affinities of lead designs within 100-fold.

This work was supported by NIH Grants CA 63081 (to RMS), GM 067542 (to ACA) and MIUR, COFIN 2002 and COSTI (to MPC). We acknowledge the assistance of Professor Barlocco Daniela.

References

Almog, R., Waddling, C. A., Maley, F., Maley, G. F. & Van Roey, P. (2001). *Protein Sci.* **10**, 988–996.
Anderson, A. C., O'Neil, R. H., Surti, T. S. & Stroud, R. M. (2001). *Chem. Biol.* **8**, 445–457.
Anderson, A. C., Perry, K. M., Freymann, D. M. & Stroud, R. M. (2000). *J. Mol. Biol.* **297**, 645–657.
Brouwer, A. E., Rajanuwong, A., Chierakul, W., Griffin, G. E., Larsen, R. A., White, N. J. & Harrison, T. S. (2004). *Lancet*, **363**, 1764–1767.
Brünger, A. T. (1992). *X-PLOR Version* 3.1. *A System for X-ray Crystallography and NMR.* New Haven: Yale University Press.
Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Carlson, H. A. (2002). *Curr. Pharm. Des.* **8**, 1571–1578.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Costi, M. P., Rinaldi, M., Tondi, D., Pecorari, P., Barlocco, D., Ghelli, S., Stroud, R. M., Santi, D. V., Stout, T. J., Musiu, C., Marangiu, E. M., Pani, A., Congiu, D., Loi, G. A. & La Colla, P. (1999). *J. Med. Chem.* **42**, 2112–2124.
Esnouf, R. M. (1997). *J. Mol. Graph.* **15**, 133–138.
Ferrari, A. M., Wei, B. Q., Costantino, L. & Shoichet, B. K. (2004). *J. Med. Chem.* **47**, 5076–5084.
Ferrari, S., Costi, P. M. & Wade, R. C. (2003). *Chem. Biol.* **10**, 1183–1193.
Fritz, T. A., Tondi, D., Finer-Moore, J. S., Costi, M. P. & Stroud, R. M. (2001). *Chem. Biol.* **8**, 981–995.
Halperin, I., Ma, B., Wolfson, H. & Nussinov, R. (2002). *Proteins*, **47**, 409–443.
Heidelberger, C., Danenberg, P. V. & Moran, R. G. (1983). *Adv. Enzymol. Relat. Areas Mol. Biol.* **54**, 58–119.
Jones, T., Zou, J.-Y., Cowan, S. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Jones, T. R., Calvert, A. H., Jackman, A. L., Brown, S. J., Jones, M. & Harrap, K. R. (1981). *Eur. J. Cancer*, **17**, 11–19.
Kamb, A., Finer-Moore, J., Calvert, A. H. & Stroud, R. M. (1992). *Biochemistry*, **31**, 9883–9890.
Knighton, D. R., Kan, C. C., Howland, E., Janson, C. A., Hostomska, Z., Welsh, K. M. & Matthews, D. A. (1994). *Nature Struct. Biol.* **1**, 186–194.
Kraulis, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.

Livi, L. L., Edman, U., Schneider, G. P., Greene, P. J. & Santi, D. V. (1994). *Gene*, **150**, 221–226.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Merritt, E. A. & Bacon, D. J. (1997). *Methods Enzymol.* **277**, 505–524.
Mitchell, T. G. & Perfect, J. R. (1995). *Clin. Microbiol. Rev.* **8**, 515–548.
Montfort, W. R., Perry, K. M., Fauman, E. B., Finer-Moore, J. S., Maley, G. F., Hardy, L., Maley, F. & Stroud, R. M. (1990). *Biochemistry*, **29**, 6964–6977.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
O'Neil, R. H., Lilien, R. H., Donald, B. R., Stroud, R. M. & Anderson, A. C. (2003). *J. Biol. Chem.* **278**, 52980–52987.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Perry, K. M., Fauman, E. B., Finer-Moore, J. S., Montfort, W. R., Maley, G. F., Maley, F. & Stroud, R. M. (1990). *Proteins*, **8**, 315–333.
Phan, J., Koli, S., Minor, W., Dunlap, R. B., Berger, S. H. & Lebioda, L. (2001). *Biochemistry*, **40**, 1897–1902.
Phan, J., Steadman, D. J., Koli, S., Ding, W. C., Minor, W., Dunlap, R. B., Berger, S. H. & Lebioda, L. (2001). *J. Biol. Chem.* **276**, 14170–14177.
Pogolotti, A. L. Jr, Danenberg, P. V. & Santi, D. V. (1986). *J. Med. Chem.* **29**, 478–482.
Santi, D. V. & Danenberg, P. V. (1984). *Folates and Pterins*, Vol. 1, edited by R. L. Blakley & S. J. Benkovic, pp. 345–398. New York: John Wiley & Sons.
Santi, D. V., McHenry, C. S., Raines, R. T. & Ivanetich, K. M. (1987). *Biochemistry*, **26**, 8606–8613.
Sayre, P. H., Finer-Moore, J. S., Fritz, T. A., Biermann, D., Gates, S. B., MacKellar, W. C., Patel, V. F. & Stroud, R. M. (2001). *J. Mol. Biol.* **313**, 813–829.
Schiffer, C. A., Clifton, I. J., Davisson, V. J., Santi, D. V. & Stroud, R. M. (1995). *Biochemistry*, **34**, 16279–16287.
Shane, B. (1989). *Vitam. Horm.* **45**, 263–335.
Sotelo-Mundo, R. R., Ciesla, J., Dzik, J. M., Rode, W., Maley, F., Maley, G. F., Hardy, L. W. & Montfort, W. R. (1999). *Biochemistry*, **38**, 1087–1094.
Stout, T. J., Sage, C. R. & Stroud, R. M. (1998). *Structure*, **6**, 839–848.
Stout, T. J. & Stroud, R. M. (1996). *Structure*, **4**, 67–77.
Stout, T. J., Tondi, D., Rinaldi, M., Barlocco, D., Pecorari, P., Santi, D. V., Kuntz, I. D., Stroud, R. M., Shoichet, B. K. & Costi, M. P. (1999). *Biochemistry*, **38**, 1607–1617.
Stroud, R. M. & Finer-Moore, J. S. (2003). *Biochemistry*, **42**, 239–247.
Takemura, Y. & Jackman, A. L. (1997). *Anticancer Drugs*, **8**, 3–16.
Taylor, N. R., Cleasby, A., Singh, O., Skarzynski, T., Wonacott, A. J., Smith, P. W., Sollis, S. L., Howes, P. D., Cherry, P. C., Bethell, R., Colman, P. & Varghese, J. (1998). *J. Med. Chem.* **41**, 798–807.
Webber, S. E. *et al.* (1993). *J. Med. Chem.* **36**, 733–746.
Weichsel, A. & Montfort, W. R. (1995). *Nature Struct. Biol.* **2**, 1095–1101.
Yeates, T. O. (1988). *Acta Cryst.* A**44**, 142–144.
Yuvaniyama, J., Chitnumsub, P., Kamchonwongpaisan, S., Vanichtanankul, J., Sirawaraporn, W., Taylor, P., Walkinshaw, M. D. & Yuthavong, Y. (2003). *Nature Struct. Biol.* **10**, 357–365.