

# Crystal structure of deoxycytidylate hydroxymethylase from bacteriophage T4, a component of the deoxyribonucleoside triphosphate-synthesizing complex

Hyun Kyu Song, Se Hui Sohn and Se Won Suh<sup>1</sup>

Department of Chemistry, College of Natural Sciences, Seoul National University, Seoul 151-742, Korea

<sup>1</sup>Corresponding author  
e-mail: sewonsuh@plaza.snu.ac.kr

**Bacteriophage T4 deoxycytidylate hydroxymethylase (EC 2.1.2.8), a homodimer of 246-residue subunits, catalyzes hydroxymethylation of the cytosine base in deoxycytidylate (dCMP) to produce 5-hydroxymethyl-dCMP. It forms part of a phage DNA protection system and appears to function *in vivo* as a component of a multienzyme complex called deoxyribonucleoside triphosphate (dNTP) synthetase. We have determined its crystal structure in the presence of the substrate dCMP at 1.6 Å resolution. The structure reveals a subunit fold and a dimerization pattern in common with thymidylate synthases, despite low (~20%) sequence identity. Among the residues that form the dCMP binding site, those interacting with the sugar and phosphate are arranged in a configuration similar to the deoxyuridylate binding site of thymidylate synthases. However, the residues interacting directly or indirectly with the cytosine base show a more divergent structure and the presumed folate cofactor binding site is more open. Our structure reveals a water molecule properly positioned near C-6 of cytosine to add to the C-7 methylene intermediate during the last step of hydroxymethylation. On the basis of sequence comparison and crystal packing analysis, a hypothetical model for the interaction between T4 deoxycytidylate hydroxymethylase and T4 thymidylate synthase in the dNTP-synthesizing complex has been built.**

**Keywords:** bacteriophage T4/deoxycytidylate hydroxymethylase/deoxyuridylate hydroxymethylase/dNTP-synthesizing complex/thymidylate synthase

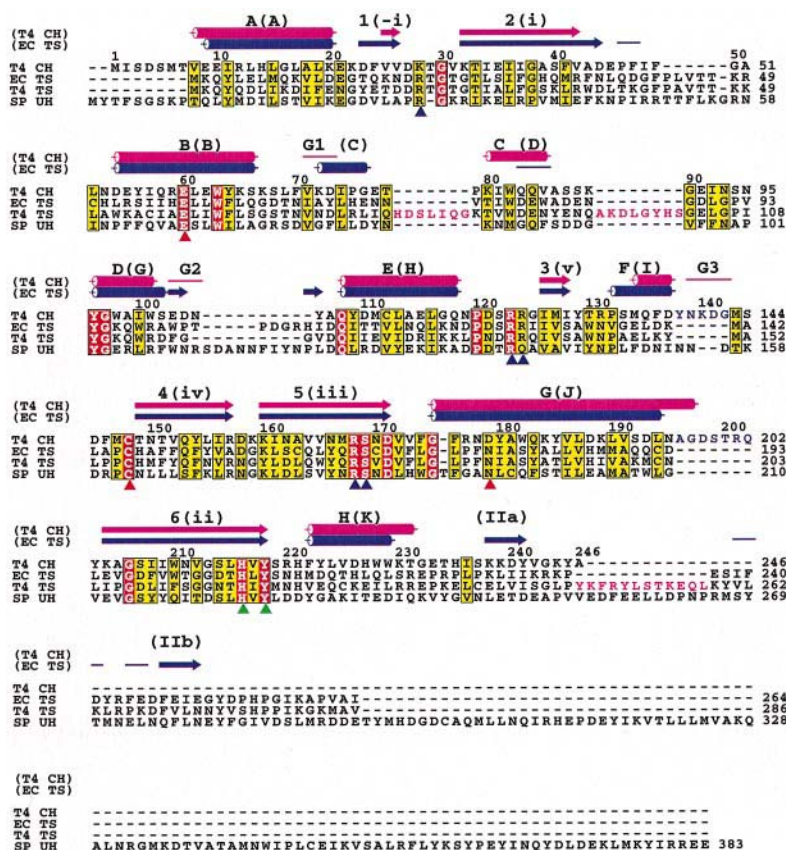
## Introduction

Besides the four standard bases, modified bases are sometimes found in natural DNA (Gommers-Ampt and Borst, 1995). The modifications vary from a simple addition of a methyl or hydroxyl group to the 'very unusual' (or hypermodified) bases such as glucosylated hydroxymethylcytosine in T-even bacteriophage DNA (Warren, 1980) and  $\beta$ -D-glucosyl-hydroxymethyluracil (or J) identified in the unicellular eukaryote *Trypanosoma brucei* (Gommers-Ampt *et al.*, 1993). The modified bases serve various important biological functions, such as protection against degradation by nucleases and regulation of gene expression.

In bacteriophage T4, a specific DNA modification system has been evolved to protect its own genome against phage-encoded nucleases and restriction endonuclease systems of its host *Escherichia coli*. This unusual DNA modification occurs at two levels. First, 2'-deoxycytidylate (or 2'-deoxycytidine-5'-monophosphate, dCMP) is converted into 5-hydroxymethyl-dCMP by a phage-encoded enzyme, deoxycytidylate hydroxymethylase (CH; EC 2.1.2.8). Hydroxymethyl-deoxycytidine triphosphate (Hm-dCTP), subsequently synthesized from Hm-dCMP by the actions of T4 deoxynucleoside monophosphate kinase and *E. coli* nucleoside diphosphate kinase, is incorporated into the phage DNA by T4 DNA polymerase. At the second level, hydroxymethyl-dCMP residues of T4 DNA are glucosylated by glucosyltransferases (Greenberg *et al.*, 1994). The crystal structure of one such bacteriophage T4  $\beta$ -glucosyltransferase has been determined (Vrielink *et al.*, 1994). Glucosylation of phage DNA was also implicated as having a control function in phage-specific gene expression (Cox and Conway, 1973; Ruger, 1978).

T4 DNA replication is tightly coupled to the biosynthesis of its precursors, which appears to be carried out *in vivo* by a multienzyme complex called deoxyribonucleoside triphosphate (dNTP) synthetase (Greenberg *et al.*, 1994). It is vital to supply hydroxymethylcytosine bases instead of unmodified cytosine for the multiplication of bacteriophage T4. T4 CH has been shown to interact, directly or indirectly, with a number of proteins involved in dNTP biosynthesis, including T4 thymidylate synthase (TS), and those involved in T4 DNA replication (Wheeler *et al.*, 1992). Therefore, it plays an important role in the biosynthetic pathway of DNA precursors in bacteriophage T4 as an essential component of the dNTP-synthesizing complex.

T4 CH, encoded by gene 42, exists as a homodimer consisting of polypeptide chains with 246 amino acid residues (subunit  $M_r$  28 450). Its amino acid sequence, as deduced from gene sequencing (Lamm *et al.*, 1987, 1988; Thylen, 1988), shows low (19–21%) but statistically significant identity with those of TSs from T4 and *E. coli* (Figure 1). There is also a limited sequence similarity with the N-terminal two-thirds of deoxyuridylate (dUMP) hydroxymethylase from bacteriophage SPO1 of *Bacillus subtilis* (Wilhelm and Ruger, 1992). T4 CH and TSs, however, exhibit distinct pyrimidine base specificities and differ in the details of their catalyzed reactions. That is, T4 CH hydroxymethylates dCMP, whereas TSs reductively methylate dUMP. The same cofactor 5,10-methylene-5,6,7,8-tetrahydrofolate ( $\text{CH}_2\text{H}_4\text{folate}$ ) is converted into 5,6,7,8-tetrahydrofolate by T4 CH but into 7,8-dihydrofolate by TSs. In the latter case, dihydrofolate is subsequently recycled to tetrahydrofolate by dihydrofolate reductase (DHFR).



**Fig. 1.** Sequence alignment of T4 dCMP hydroxymethylase (T4 CH), *E. coli* thymidylate synthase (EC TS), T4 thymidylate synthase (T4 TS) and SPO-1 dUMP hydroxymethylase (SP UH). The catalytically important residues of T4 CH (Glu60, Cys148 and Asp179) are marked by red triangles, while the deoxyribose-binding residues (His216 and Tyr218) and phosphate-binding residues (Lys28, Arg123', Arg124', Arg168 and Ser169) are marked by green and blue triangles, respectively. Secondary structure elements are indicated above the sequence for T4 CH (magenta) and *E. coli* TS (blue). For T4 CH, six  $\beta$ -strands (arrows) are numbered sequentially 1–6, eight  $\alpha$ -helices (cylinders) are labeled sequentially A–H, and three  $3_{10}$ -helices (thin lines) are labeled G1–G3. The labels for the secondary structure elements of *E. coli* TS are also given in parentheses. This figure was produced with ALSCRIPT (Barton, 1993).

Extensive biochemical and structural studies on thymidylate synthases have provided a wealth of information regarding the catalytic mechanism, specific interactions with dUMP and folate analogs, and stability (Stroud and Finer-Moore, 1993; Carreras and Santi, 1995; Stout *et al.*, 1998). dUMP hydroxymethylase from bacteriophage SPO1 has been crystallized (Schellenberger *et al.*, 1995). However, little structural information is available on a pyrimidine hydroxymethylase, as no three-dimensional structure of either dCMP or dUMP hydroxymethylase has been reported to date. In order to gain insight into the substrate specificity, catalysis and interaction with T4 TS in the dNTP-synthesizing complex, we have determined the high-resolution crystal structure of T4 CH, thereby providing the first view of a pyrimidine hydroxymethylase.

## Results and discussion

### Model quality and comparison of different models

Initially, the structure of dCMP-bound T4 CH, fused with a His<sub>6</sub>-tag at its C-terminus, was determined by the multiple isomorphous replacement (MIR) method using three heavy-atom derivatives (Table I) and was refined to 1.6 Å (Figure 2A; native II model in Table II). Subsequently, two more structures were refined at

**Table I.** Heavy atom data collection and phasing statistics

	Native I	KAu(CN) <sub>2</sub>	Hg(AcO) <sub>2</sub>	SeMet
Resolution (Å)	2.3	2.5	2.8	2.5
Unique reflections	24 844	17 837	14 665	16 603
Completeness (%)	85.4	84.9	92.6	80.3
$R_{\text{merge}}(\%)^a/R_{\text{iso}}(\%)^b$	3.5	5.3/14.4	4.4/17.2	5.5/12.1
Number of sites		3	2	14
Phasing power <sup>c</sup> / $R_{\text{Cullis}}^d$		1.31/0.76	1.27/0.79	1.97/0.61
Mean overall figure-of-merit to 3.0 Å = 0.62 (0.80 after DM)				

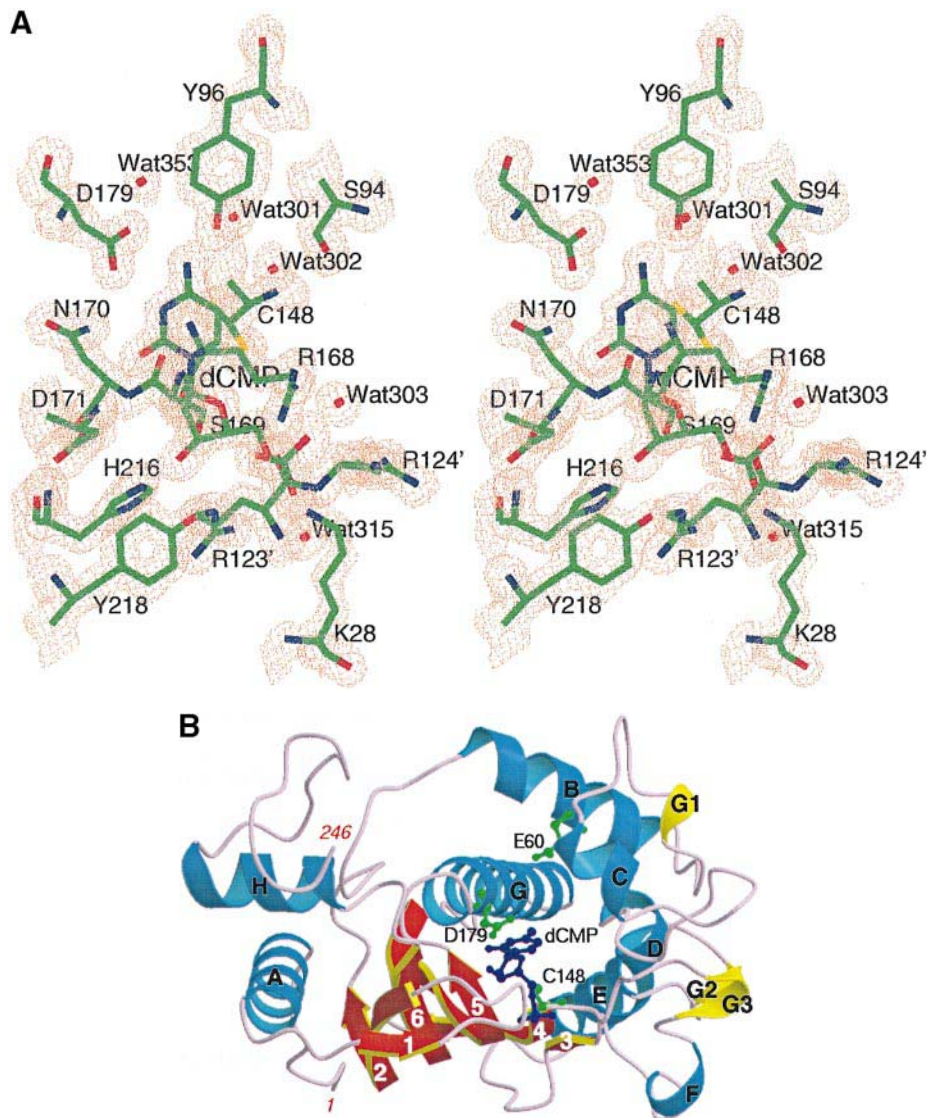
<sup>a</sup> $R_{\text{merge}} = \sum_i \sum_j |I(h,i) - \langle I(h) \rangle| / \sum_i \sum_j I(h,i)$ , where  $I(h,i)$  is the intensity of the  $i^{\text{th}}$  measurement of  $h$  and  $\langle I(h) \rangle$  is the corresponding value of  $I(h)$  for all  $i$  measurements.

<sup>b</sup> $R_{\text{iso}} = \sum ||F_{\text{PH}}| - |F_{\text{P}}|| / \sum |F_{\text{P}}|$ , where  $F_{\text{PH}}$  and  $F_{\text{P}}$  are the derivative and native structure factors, respectively.

<sup>c</sup>Phasing power =  $\langle F_{\text{H}} \rangle / E$ , where  $\langle F_{\text{H}} \rangle$  is the r.m.s. heavy-atom structure factor and  $E$  is the residual lack of closure error.

<sup>d</sup> $R_{\text{Cullis}} = \sum ||F_{\text{PH}} \pm F_{\text{P}}| - |F_{\text{PH}(\text{calc})}|| / \sum |F_{\text{PH}} - F_{\text{P}}|$ , where  $F_{\text{PH}}$  and  $F_{\text{PH}(\text{calc})}$  are the observed and calculated structure factors of a heavy atom derivative.

2.2–2.3 Å (native III and Pi-bound models in Table II). In all of these models, no residues lie in the disallowed region of the Ramachandran plot. The 1.6 Å model of



**Fig. 2.** (A) Final ( $2F_o - F_c$ ) electron density map around the bound dCMP molecule in the T4 CH structure. The map was calculated using 15–1.6 Å data and contoured at  $1.3 \sigma$  ( $0.28 e \text{ \AA}^{-3}$ ). This figure was drawn with CHAIN (Sack, 1988). (B) Ribbon diagram showing the secondary structure elements of T4 CH monomer. Six  $\beta$ -strands (red arrows), eight  $\alpha$ -helices (blue ribbons) and three  $3_{10}$ -helices (yellow ribbons) are drawn. The catalytically important residues (Glu60, Cys148 and Asp179) and a bound dCMP molecule are shown in green and deep blue, respectively.

**Table II.** Crystallographic refinement statistics

Model	Native II	Native III	Pi-bound
X-rays	1.00 Å (PF <sup>a</sup> )	CuK $\alpha$ (SNU <sup>b</sup> )	CuK $\alpha$ (SNU)
Resolution (Å)	8.0–1.6	8.0–2.2	8.0–2.3
Unique reflections, no cut-off	76 944	28 632	25 579
Completeness (%) <sup>c</sup>	97.3 (83.9)	90.0 (65.1)	91.2 (71.3)
No. of residues (His-tagged <sup>d</sup> )	2×241 (yes)	2×246 (no)	2×241 (yes)
Hetero-atoms, non-hydrogen	40 (2 dCMP)	40 (2 dCMP)	10 (2 PO <sub>4</sub> )
Water oxygen atoms	158	151	138
$R_{\text{cryst}}/R_{\text{free}}(\%)^e$	18.9/20.8	16.0/20.8	15.9/20.2
R.m.s. deviations			
Bond lengths/bond angles	0.008 Å/0.921°	0.007 Å/0.868°	0.007 Å/0.881°
Average B-factors (Å <sup>2</sup> )			
Main-chain/side-chain atoms	18.4/23.0	17.5/23.4	20.1/24.3
Waters/hetero-atoms	24.7/18.4	24.3/28.0	22.2/50.6

<sup>a</sup>PF, Photon Factory, Tsukuba, Japan. <sup>b</sup>SNU, Seoul National University, Seoul, Korea. <sup>c</sup>Values in parentheses are for reflections in the highest resolution bin of 0.1 Å thickness. <sup>d</sup>His-tagged, the presence or absence of a C-terminal His<sub>6</sub>-tag in the recombinant protein.

<sup>e</sup> $R = \frac{\sum ||F_o| - |F_c||}{\sum |F_o|}$ , where  $R_{\text{cryst}}$  and  $R_{\text{free}}$  factors are calculated using the 95% working set and 5% test set of reflections, respectively.

dCMP-bound CH (native II in Table II) accounts for 241 residues in each subunit, as well as ordered solvent molecules and bound dCMP molecules. The C-terminal five residues of CH and six histidine residues showed no electron density. The non-crystallographic symmetry (NCS) between the two subunits of the homodimer in the crystal asymmetric unit is a twofold rotation of 180.0° without any translation. A superposition of NCS-related subunits gives a root-mean-square (r.m.s.) deviation of 0.18 Å for all 1972 atom pairs. Moreover, the two subunits have the same average B-factor of 20.6 Å<sup>2</sup>. Therefore, they are indistinguishable and one of them is arbitrarily chosen for the following discussion, unless otherwise stated.

Subsequently, T4 CH without the C-terminal His-tag was crystallized, also in complex with dCMP, and its structure was refined at 2.2 Å (native III in Table II). It is nearly identical to the 1.6 Å model of His-tagged CH; r.m.s. deviations are 0.15 Å for 964 main-chain atom pairs and 0.19 Å for 1008 side-chain atom pairs in 241 residues, respectively. However, it clearly revealed the missing C-terminal five residues. Therefore, the presence of a His-tag at the C-terminus had little effect on the overall structure, except that the C-terminal end was disordered when the His-tag was attached. The C-terminal five residues from this model have been included in the discussion and model display for the sake of completeness.

T4 CH fused with a C-terminal His-tag was also crystallized in the presence of 5-hydroxymethyl cytosine, a fragment of the product, but its structure did not show the ligand in the active site. Instead, it showed a strong electron density of a bound phosphate ion at the substrate binding pocket (Pi-bound model in Table II). Presumably, the phosphate ion bound tightly to the enzyme during purification. Between the dCMP- and phosphate-bound models, there is no gross structural difference. The r.m.s. deviations are 0.21 Å for 964 main-chain atoms and 0.31 Å for 1008 side-chain atoms in 241 residues, respectively. The segment showing the largest deviation is a loop region around Lys28. The main-chain atoms of residues 28–30 show a r.m.s. deviation of 0.96 Å. The average B-factor of this region is also dramatically lowered upon dCMP-binding (from 85.5 to 30.2 Å<sup>2</sup>), indicating that the flexibility of this loop is reduced upon binding dCMP as a consequence of covalently anchoring the phosphate to the deoxyribonucleoside.

### Subunit and homodimer structure

The homodimer of T4 CH consists of two essentially identical subunits and has approximate dimensions of 72×64×40 Å. The monomer structure consists of a six-stranded β-sheet, surrounded by eight α-helices and three 3<sub>10</sub>-helices (Figure 2B). Each subunit folds into a two-domain structure with approximate dimensions of 55×48×32 Å. A common subunit fold is shared by TSs. Each active site of the dimer is contributed asymmetrically by residues from both subunits. Thus the substrate dCMP bound to the active site is located very close to the dimer interface (Figure 3A and B). All six β-strands within each monomer as well as α-helices A, E and F are involved in dimerization (Figures 2B and 3A). The dimer interface is contributed by the following segments: residues 1–6, 14–18, 23–28, 35–40, 106–138, 144–157, 164–169

and 209–216. The buried surface area in the interface is 1836 Å<sup>2</sup>. The association of the two β-sheets to form the dimer is unique in that the twist of one sheet relative to the other is right-handed (Cohen *et al.*, 1981). A similar pattern of dimerization was previously revealed in *Lactobacillus casei* TS (Hardy *et al.*, 1987). A highly negatively charged surface patch is formed by clustering of 18 acidic residues, Asp4, Glu9, Glu10, Glu21, Asp23, Asp44, Glu45, Asp121 and Asp158 from one monomer and their equivalents from the second monomer, around the dimer twofold axis (Figure 3C). This may serve an important biological function such as protein–protein interactions.

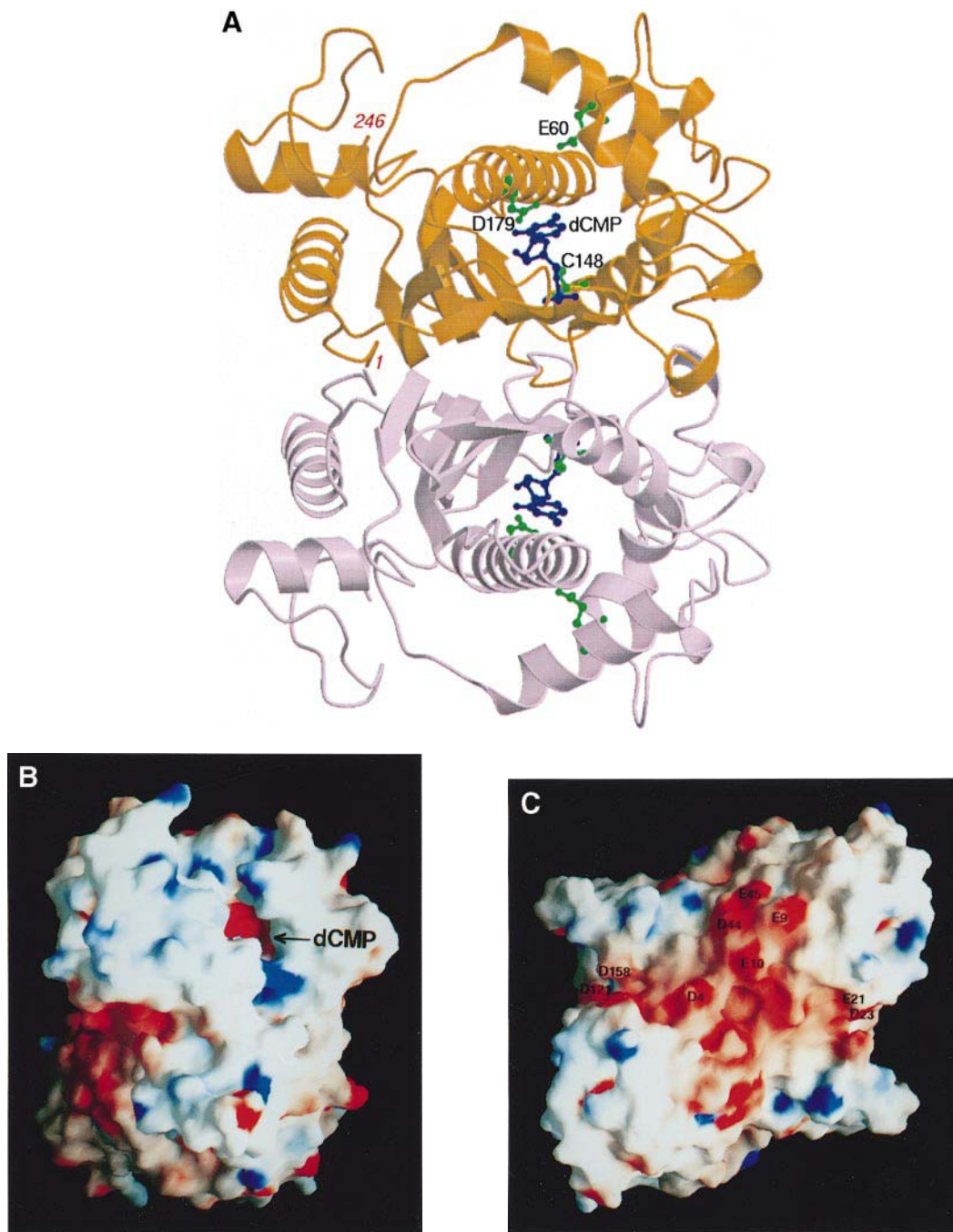
### Similarity to thymidylate synthases

Structural similarity between T4 CH and TSs may be anticipated on the basis of limited sequence identity (Lamm *et al.*, 1988; Wilhelm and Rüger, 1992) and the utilization of the same cofactor. This is confirmed by the PROTEP (Grindley *et al.*, 1993) and DALI (Holm and Sander, 1993) searches, which find a notable structural resemblance only in TSs. Although T4 CH and TS subunits share a common fold, there exist many significant structural differences (Figure 4). Only a limited portion of the CH subunit including the central β-sheet and the key active site residues (Cys148, Asp179) can be superimposed readily with the corresponding part of TSs. A superposition with *E.coli* TS gives a r.m.s. difference of 0.90 Å for 79 C<sub>α</sub> atoms which encompass the central β-sheet β2–β6 and α-helices E and G of T4 CH (Figure 4). On the other hand, a superposition of the remaining structural elements is not straightforward due to large differences in many parts. For instance, the helices A(A), B(B), and H(K) show a shift of ~2.5, ~4.0 and ~2.7 Å, respectively. The labels given in parentheses are those of the secondary structure elements in *E.coli* TS (Montfort *et al.*, 1990). The loops connecting the secondary structure elements show even larger differences. For instance, a loop connecting α-helices C(D) and D(G) and another loop between β-strands 1(i) and 2(i) are shifted by >5.0 Å. Additional structural differences between T4 CH and *E.coli* TS result from the above-mentioned insertions and deletions.

A major difference between T4 CH and *E.coli* TS is that the C-terminal region of TS is deleted in CH. Among the extra 27 C-terminal residues present in *E.coli* TS, the last six residues move ~4 Å into and partly cover the active site upon binding folate (Matthews *et al.*, 1990a; Montfort *et al.*, 1990). Therefore, the presumed folate binding site of T4 CH is more open than that of *E.coli* TS (Figure 4). Other C-terminal residues of *E.coli* TS which are deleted in T4 CH consist of two short strands (IIa, IIb in Figures 1 and 4) and two 3<sub>10</sub>-helices (Figure 1). Some parts of this region are believed to provide an interaction surface for DHFR (Knighton *et al.*, 1994; Stroud, 1994). In contrast, T4 CH is not functionally required to interact with DHFR, since tetrahydrofolate is produced in the T4 CH-catalyzed hydroxymethylation reaction.

Among the seven extra residues at the N-terminus of T4 CH, four residues Met1, Ser3, Asp4, and Met6 contribute to the dimer association through interaction with the second β-strand of the other subunit. The amino





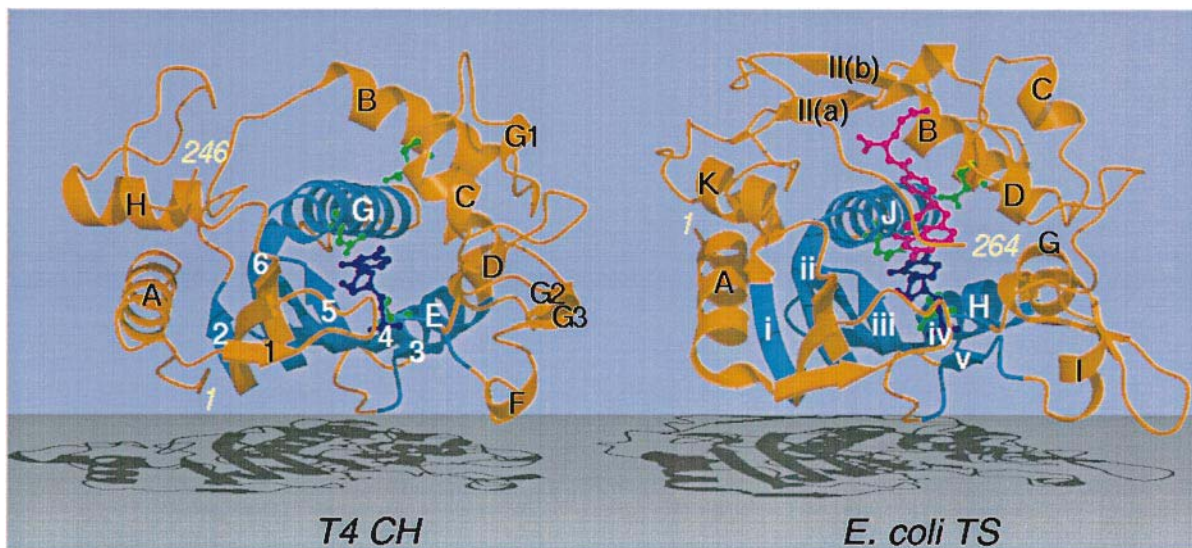
**Fig. 3.** (A) Ribbon diagram showing the T4 CH dimer, with each subunit colored in orange and light pink, respectively. The catalytically important residues and dCMP molecule are shown as in Figure 2B. (B) Diagram showing the electrostatic potential at the molecular surface of T4 CH. A view along the non-crystallographic twofold symmetry axis showing the dCMP binding site. This is the same view as in (A). Negatively charged regions are red and positively charged regions blue. This figure was drawn with GRASP (Nicholls *et al.*, 1991). (C) A view obtained by a 90° rotation of (B) around a vertical axis. The residues forming the highly acidic surface are labeled in the upper subunit only for clarity.

group of the N-terminal methionine residue in *E.coli* TS was found to be modified as carbamate (Fauman *et al.*, 1994). This is stabilized by binding to a threonine pocket, where it is sheltered from the bulk solvent (Fauman *et al.*, 1994). No modification of the N-terminus of T4 CH, which is longer by seven residues than *E.coli* TS, is indicated by the electron density. Consistent with this, Thr46 and Thr47 in the threonine pocket of *E.coli* TS are deleted in T4 CH (Figure 1). This region shows a

very large deviation between the two structures with a maximum shift of ~6.0 Å.

**dCMP binding and pyrimidine base specificity**

The substrate dCMP is bound in a deep active-site pocket of T4 CH, in a manner similar to the binding of dUMP to TSs (Figures 2B and 3A,B). When the nucleotide binding sites of T4 CH and *E.coli* TS are compared, part of the binding site that recognizes the common portion of



**Fig. 4.** Ribbon diagram comparing the subunit folds of T4 CH (left) and *E. coli* TS (right). The structurally similar parts are colored in blue and the remaining parts in orange. The catalytically important residues (Glu60, Cys148 and Asp179 in T4 CH; Glu58, Cys146 and Asn177 in *E. coli* TS) are shown in green. The substrates dCMP in T4 CH and dUMP in *E. coli* TS are drawn in deep blue and the cofactor analog CB3717 bound to *E. coli* TS is in magenta. This figure was drawn by MOLSCRIPT (Kraulis, 1991) and RASTER 3D (Merritt and Murphy, 1994).

the two nucleotides, i.e. 2'-deoxyribose and 5'-phosphate, superimposes very well, whereas the remaining part that confers the distinct pyrimidine base specificity and catalytic activity is more divergent. In T4 CH, 12 residues interact directly with dCMP, while four more residues (Glu60, Tyr64, Trp82 and Ser94) interact indirectly via water molecules or other residues, forming a network of hydrogen bonds, which includes at least five ordered water molecules (Figures 2A and 5). The five residues Tyr96, Cys148, Asn170, Asp171 and Asp179 directly contact the cytosine base and the two residues His216 and Tyr218 make hydrogen bonds to 2'-deoxyribose sugar, while the other five residues (Lys28, Arg123', Arg124', Arg168 and Ser169) and two ordered water molecules form a cage-like binding pocket for the phosphate group (Figures 2A and 5). The primed residues belong to the second subunit within a dimer. The phosphate-binding site of T4 CH is very similar but not identical to that of *E. coli* TS. Lys28 of T4 CH is replaced by Arg23 in *L. casei* TS, for which a precise positioning of the folate is assisted by the C-terminus. More specifically, the C-terminal carboxylate of Val316 forms direct hydrogen bonds with Arg23 and Trp85, while Arg23 and the backbone carbonyl oxygen of Ala315 form direct and water-mediated hydrogen bonds to the folate cofactor (Perry *et al.*, 1993; Carreras and Santi, 1995). These interactions involving the cofactor are not likely to be preserved in a T4 CH ternary complex, as a consequence of the substitution of this arginine with lysine as well as the much shortened C-terminus.

The key catalytic residue Cys148 of T4 CH acts as a nucleophile to attack the C-6 atom of the cytosine base (Graves *et al.*, 1992). Its side-chain adopts a dual conformation in the 1.6 Å model (Figure 2A). In each of the two conformations, S $\gamma$  of Cys148 lies within 3.4 or 3.5 Å from C-6 of dCMP. The corresponding residue Cys146 of *E. coli* TS forms a covalent bond with C-6 of dUMP during catalysis (Montfort *et al.*, 1990; Rutenber and Stroud, 1996). The sulfur atom of Cys148 in one of the dual conformations is also close to the guanidium group

of Arg168 side-chain (Figures 2A and 4), lowering the pK<sub>a</sub> of Cys148 side-chain by stabilizing the thiolate anion. The corresponding residues Arg166 of *E. coli* TS and Arg218 of *L. casei* TS lie similarly in close proximity to the side-chain of the reactive nucleophile cysteine (Matthews *et al.*, 1990b; Stroud and Finer-Moore, 1993; Carreras and Santi, 1995).

Asp179 of T4 CH was suggested to be the primary determinant of the pyrimidine base specificity by substituting it with Asn, which lowered the value of  $k_{cat}/K_M$  for dCMP by  $1.5 \times 10^4$ -fold but increased it for dUMP by 60-fold (Graves *et al.*, 1992). As a result, the D179N mutant of T4 CH shows a slight preference for dUMP. The equivalent residue in dUMP hydroxymethylase from bacteriophage SPO1 is Asn194 (Figure 1); the corresponding residue is invariably Asn in all TSs. Site-directed mutagenesis of the corresponding residue Asn229 of *L. casei* TS has also indicated a role in determining the pyrimidine base specificity, as the N229D mutation resulted in a 40-fold increase in the specificity for dCMP over dUMP (Liu and Santi, 1993; Agarwalla *et al.*, 1997). Recent structural studies on Asn229 mutants suggest, however, that Asn229 does not contribute substantially to substrate binding energy and may contribute to catalysis through orientation of dUMP and through hydrogen-bond stabilization of reaction intermediates (Finer-Moore *et al.*, 1998). This suggests that it is not the cyclic hydrogen bond network bridging atoms at positions 3 and 4 that plays a major role in determining the pyrimidine base specificity of *L. casei* TS. But it is rather the difference in charge properties of the reaction intermediates that discriminates one side-chain over the other. With dUMP as substrate, a negative charge is developed at O-4, when Cys198 adds covalently to C-6 of dUMP during methylation by *L. casei* TS (Stroud and Finer-Moore, 1993; Figure 1), and hydrogen bonds from Asn229 and water molecules were proposed to stabilize the developing negative charge at O-4 (Finer-Moore *et al.*, 1998). The negatively charged reaction intermediate will be

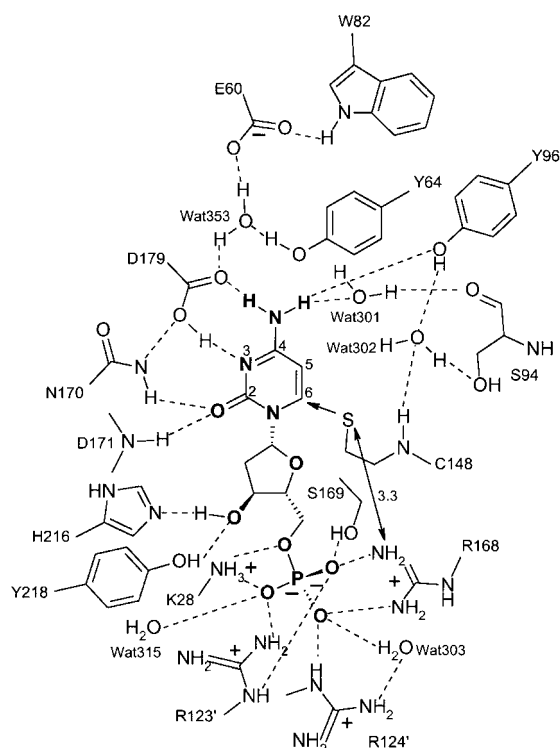


Fig. 5. Schematic diagram of the hydrogen-bond network in the active site of T4 CH. The distance between the S $\gamma$  atom of Cys148 and the guanidium group of Arg168 (3.3 Å) is indicated.

destabilized by replacing Asn229 with Asp. Our structure shows that the orientation of the Asp179 side-chain of T4 CH is constrained by the hydrogen bond network (Figure 5), which is analogous to the case of Asn229 in *L.casei* TS (Stroud and Finer-Moore, 1993). This network will be disturbed when the side-chain of Asp179 is rotated to make hydrogen bonds with dUMP. If it is mutated to Asn, the electron-deficient reaction intermediates (Graves *et al.*, 1992; Scheme 1) will not be stabilized as well. Therefore we propose, in analogy with *L.casei* TS, that Asp179 of T4 CH discriminates for dCMP over dUMP by achieving a proper orientation of the pyrimidine base for nucleophilic attack by Cys148 through a hydrogen bond network and a better stabilization of the reaction intermediates.

#### Catalysis of hydroxymethylation reaction

A catalytic mechanism has been proposed for CH in analogy with that for TS (Graves *et al.*, 1992). During the initial steps, they both catalyze the transfer of a methylene group from CH<sub>2</sub>H<sub>4</sub>folate to C-5 of either dCMP or dUMP (Carreras and Santi, 1995; Figure 3). <sup>18</sup>O-exchange studies indicated the existence of a 5-exocyclic methylene intermediate in the reaction mechanism of T4 CH (Butler *et al.*, 1994). The same intermediate was also proposed for TSs (Stroud and Finer-Moore, 1993). The catalytic activities of these enzymes are differentiated by the last step of the proposed mechanisms. During this critical step, T4 CH catalyzes the hydroxylation of the methylene group to produce Hm-dCMP without oxidizing the cofactor, whereas TS catalyzes the reduction of the exocyclic methylene to a methyl group with a concomitant oxidation of the cofactor to dihydrofolate. In the TS-catalyzed

reaction, the C-6 hydrogen of the pterin ring of the folate cofactor is donated to the exocyclic methylene group (Carreras and Santi, 1995). In contrast, a water molecule is added in the T4 CH-catalyzed reaction (Graves *et al.*, 1992).

Glu60/Trp82/Tyr96 of T4 CH, interacting directly or indirectly with the cytosine base through the hydrogen bonding network, are conserved in the *E.coli* TS sequence, corresponding to Glu58/Trp80/Tyr94. However, the locations of their side-chains relative to the pyrimidine base are shifted by ~3.8, ~4.8 and ~2.6 Å, respectively, when compared with those in the binary complex of *E.coli* TS (PDB ID 1bid). Our structural comparison indicates that it is unjustifiable to assume a similar functional role for all of these residues of T4 CH as in TSs on the basis of sequence conservation alone. Specifically, Tyr96 of T4 CH, equivalent to Tyr94 in *E.coli* TS, is slightly shifted in its position to allow a water molecule (Wat302 in Figures 2A and 5) to occupy nearly the same site (within 0.6 Å) as the hydroxyl group of Tyr94 in the ternary complex of *E.coli* TS (PDB ID 1kce) (Sage *et al.*, 1996). We suggest that this water molecule is likely to be added to the exocyclic methylene group in the last step of the hydroxymethylation reaction and that Tyr94 of *E.coli* TS plays a role in displacing this water molecule to a position unfavorable for attacking the exocyclic methylene group in the methylation reaction. This water is anchored by three hydrogen bonds to Ser94 O $\gamma$ , Tyr96 O $\eta$  and Cys148 backbone NH (Figures 2A and 5). Ser94 is not conserved and is substituted by Pro92 in *E.coli* TS or by Leu144 in *L.casei* TS. The main-chain trace around this region shows a large deviation between T4 CH and TSs. The proposed role of Wat302 in hydroxymethylation by T4 CH may be tested by replacing Ser94 with appropriate amino acids, as the hydration step is expected to be impaired upon displacing this water molecule.

In a predicted mechanism of pyrimidine hydroxymethylases, it was considered that a water molecule called water<sup>C7</sup>, located 3.4 Å from C-7 of dTMP in the product ternary complex of *E.coli* TS, or a nearby water molecule could be the source of the hydroxyl group transferred to the exocyclic methylene intermediate (Fauman *et al.*, 1994). This water<sup>C7</sup> in *E.coli* TS occupies a site 2.5 Å away from Wat302 in T4 CH, when the pyrimidine bases and the two pairs of catalytically important residues Cys148/Cys146 and Asp179/Asn177 are overlapped. If the exocyclic methylene group is built into dCMP bound to T4 CH, the water molecule Wat302 is 3.2 Å from the C-7 atom of the exocyclic methylene and is nearly aligned with a  $\pi$  orbital, making an angle of 98° between C-5 to C-7 and C-7 to Wat302, in a good position to add to the C-7 methylene. In *E.coli* TS, this water position is occupied by the side-chain hydroxyl group of Tyr94. It suggests that the strictly conserved tyrosine residue in TSs could play a role in preventing hydroxymethylation by displacing this crucial water molecule. That is, Wat302 in our T4 CH structure is likely to be added to the C-7 methylene but is probably displaced in TSs to the site of water<sup>C7</sup> in *E.coli* TS and it is then poorly positioned for a hydroxyl group transfer.

In our attempts to grow crystals of ternary complexes of T4 CH, it was observed that the enzyme shows a greater affinity for the folate cofactors when dCMP is first

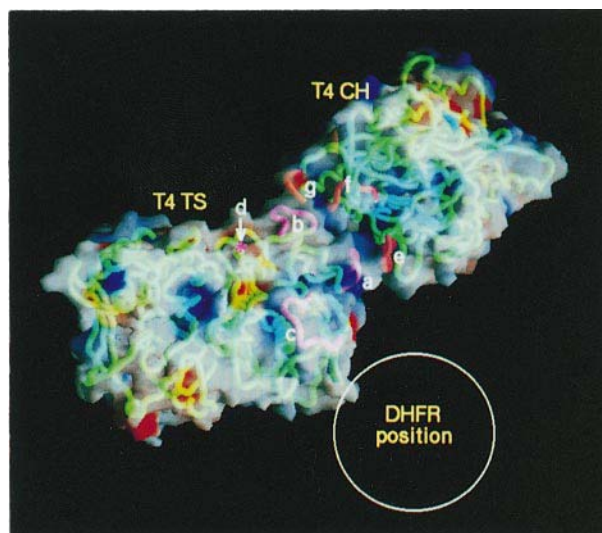


bound. This suggests that the folate cofactor binding site is partially contributed by the bound dCMP. This is similar to the case of TSs, in which dUMP binds to the enzyme and then the pterin ring of  $\text{CH}_2\text{H}_4\text{folate}$  or its analog, the quinazoline ring of CB3717, binds with its face parallel to the face of dUMP (Stroud and Finer-Moore, 1993). However, we have not been able to see a strong electron density for the folate cofactors in the presumed ternary complex crystals. This difficulty is possibly attributable to the weak binding of the folate cofactor to T4 CH. This is consistent with the expectation that the cofactor would not be covered by the shortened C-terminus in the ternary complex of T4 CH (Figure 4). Due to the more open folate binding pocket in T4 CH, the remaining part of the cofactor, *p*-aminobenzoic acid ring and glutamate, may be less constrained by the T4 CH residues and may adopt multiple conformations in the ternary complex. The C-terminus plays a key role in positioning the folate cofactor in *L. casei* TS (Perry *et al.*, 1993), which is critical for the hydride transfer from C-6 of the pterin moiety to the exocyclic methylene intermediate. In T4 CH, the pterin ring of  $\text{CH}_2\text{H}_4\text{folate}$  cofactor is expected to stack with the base of dCMP in a slightly altered fashion, as a subtle change in the positioning of the folate will be required to make the hydride transfer unfavorable without affecting the methylene transfer.

#### Interaction with T4 thymidylate synthase in dNTP-synthesizing complex

During T4 phage reproduction, at least eight T4 enzymes (including CH and TS) and two host enzymes (nucleoside diphosphate kinase and adenylate kinase) involved in *de novo* synthesis of dNTPs form an organized multi-enzyme complex called dNTP synthetase, efficiently channeling DNA precursors to the replication apparatus (Allen *et al.*, 1983; Greenberg *et al.*, 1994). The interactions between dNTP biosynthetic enzymes and replication proteins have been shown (Wheeler *et al.*, 1992). Several of the T4 enzymes involved in dNTP biosynthesis, including TS, DHFR, dCTPase-dUTPase and ribonucleotide reductase, were found to adsorb specifically to a column of immobilized T4 CH (Wheeler *et al.*, 1992). An anti-anti-T4 CH antibody was found to be specific in binding to T4 TS, with no detectable affinity for *E. coli* TS (Young and Mathews, 1992). This indicates a direct association of T4 CH with T4 TS but not with the host TS, although T4 and *E. coli* TSs share highly similar sequence and structure (48.6% sequence identity and 0.85 Å r.m.s. deviation for 175  $\text{C}_\alpha$  atom pairs). The AT-rich T4 genome requires thymine nucleotides to be synthesized at approximately twice the rate of hydroxymethylcytosine nucleotides. The 2:1 ratio was maintained over a wide range of physiological conditions, suggesting either physical or functional associations between T4 CH and T4 TS, which could help regulate their activities relative to each other (Mathews *et al.*, 1993).

Interestingly, when a dimeric molecule of T4 CH in our crystal lattice is replaced by T4 TS through a superposition, the two regions of T4 CH (residues 138–142 of the first subunit and residues 196'–202' of the second) interact with the two unique insertion regions of T4 TS (residues around 91–102 and 78–82), respectively. The latter two regions are also involved in the



**Fig. 6.** A view showing a possible interaction between T4 CH and T4 TS in the proposed complex model. A transparent electrostatic potential surface along with the backbone as green tubes is shown. Three unique insertions and a deletion present in T4 TS compared with *E. coli* TS are marked by a–d. Three segments of T4 CH involved in the interaction are marked by e–g. A circle is drawn to indicate the position of DHFR domain in the bifunctional DHFR-TS from *L. major*, corresponding to the approximate dimension of monomeric DHFR from *E. coli*.

crystal packing of T4 TS (Finer-Moore *et al.*, 1994). The two regions of T4 CH at the interface correspond to the unique insertions compared with *E. coli* and T4 TSs (Figure 1). Prompted by this observation and the encouraging results of other investigators in deducing biologically relevant protein–protein interactions from an inspection of the crystal packing (Story *et al.*, 1992; Rice and Steitz, 1994; Janin and Rodier, 1995; Shapiro *et al.*, 1995), a hypothetical model of the complex between T4 CH and T4 TS was built using MULTIDOCK (Jackson *et al.*, 1998). This model shown in Figure 6 is largely consistent with the proposal that three insertions and one deletion in the primary sequence, unique to T4 TS, possibly provide T4 TS-specific intermolecular interaction sites (Finer-Moore *et al.*, 1994).

The proposed interaction site on T4 TS encompasses the two inserted sequences at 76–83 and 95–102 (marked a and b in Figure 6). The remaining insertion at 237–248 (marked c in Figure 6) and a deletion between 117 and 118 (marked d in Figure 6) are located adjacent to this interface, possibly providing an interaction surface for other components of the dNTP-synthesizing complex. On the T4 CH side, the interaction site is formed by the two segments covering the insertion regions (marked e and g in Figure 6 for residues 136–144 and 197'–201', respectively) and a third segment covering residues 117'–124' (marked f in Figure 6). This model of the complex has an interface accessible surface area (ASA) of  $\sim 440 \text{ \AA}^2$ , when four side-chains each of T4 CH and T4 TS are manually adjusted for a better fit (<http://www.biochem.ucl.ac.uk/bsm/PP/server>). These are substantially smaller than the mean value for non-obligatory heterodimer complexes,  $804 \text{ \AA}^2$  (Jones and Thornton, 1996). A relatively small interface area between T4 CH and T4 TS in the proposed complex suggests a weak binding of the two enzymes, as



the binding energy correlates well with the buried surface area in the protein–protein interface (Horton and Lewis, 1992). T4 TS was also shown to interact with several T4 proteins including dCMP deaminase, DHFR and single-stranded DNA binding protein (gp32) (McGaughey *et al.*, 1996; Wheeler *et al.*, 1996), as well as CH. As a result, the interaction between T4 CH and T4 TS needs to be limited in its extent and sufficient surface areas of T4 TS outside the interface with T4 CH should be available for interactions with other proteins.

Furthermore, the proposed model for the complex between T4 CH and T4 TS does not seem to be ruled out by the structure of the bifunctional DHFR-TS from *Leishmania major* (Knighton *et al.*, 1994). A common mechanism of substrate channeling across the surface of the protein between the two active sites was indicated by this structure (Knighton *et al.*, 1994; Stroud, 1994; Elcock *et al.*, 1996). Therefore, it seems reasonable to assume that the general position of T4 DHFR bound to T4 TS will not be too different from that of the DHFR domain in the bifunctional enzyme. This enables us to envision a crude model of the complex between T4 TS and T4 DHFR. Since the structure of T4 DHFR, a homodimer of 193-residue subunits, has not yet been determined, the model of *E.coli* DHFR, a monomer with 159 residues, was superimposed on the DHFR domain of the bifunctional enzyme (Figure 6). It shows that T4 DHFR, even as a dimer, will not overlap T4 CH. Therefore, the proposed interaction between T4 CH and T4 TS appears to be consistent with the data available. Nevertheless, it needs to be validated by structural analysis of the complex.

## Materials and methods

### Expression and purification of recombinant T4 dCMP hydroxymethylase

Construction of the expression vector (pET-22b-CH) and purification of His<sub>6</sub>-tagged T4 CH have been described elsewhere (Sohn *et al.*, 1999). For the selenomethionine-substituted protein, the expression vector was transformed into the methionine auxotroph B834(DE3)pLysS cells. M9 cell culture medium containing extra amino acids was used instead of Luria–Bertani (LB) medium and 5 mM DTT was added right after metal chelate chromatography on Ni-NTA resin. After the initial structure determination by MIR, the enzyme without the C-terminal His-tag was expressed by amplifying the DNA-containing translation termination codon at the 3'-position and inserting it into *NdeI*–*XhoI*-digested pET-22b. The intact enzyme was highly overexpressed in soluble form in B834(DE3) cells upon induction by 0.5 mM isopropyl-β-D-galactopyranoside (IPTG) at 30°C. Instead of His-tag affinity chromatography, ammonium sulfate fractionation in the range 45–70% saturation and DEAE ion exchange chromatography were employed. Other purification steps were the same as for the His-tagged enzyme.

### Crystallization and X-ray data collection

Crystallization of His-tagged T4 CH in the presence of dCMP and X-ray data collection have been reported elsewhere (Sohn *et al.*, 1999). The crystals belong to the space group C2, with unit cell dimensions of  $a = 174.22$ ,  $b = 53.12$ ,  $c = 75.17$  Å and  $\beta = 115.29^\circ$  (determined using native II data in Table II). Selenomethionine-substituted His-tagged enzyme and the native, intact enzyme without the His-tag crystallized isomorphously under the same conditions as the native His-tagged protein.

### Structure determination and refinement

Three heavy atom derivatives were used for phasing by the MIR method (Table I). Mercury and gold sites were located by the difference Patterson interpretation with RSPS (CCP4, 1994) and SOLVE (Terwilliger and Berendzen, 1996). Minor heavy atom sites and selenium sites were located in the cross-phase difference Fourier maps. Heavy atom para-

meters were refined with MLPHARE (CCP4, 1994). Initial MIR phases were improved by solvent flattening and two-fold non-crystallographic symmetry averaging with DM (CCP4, 1994). The electron density map was of sufficient quality to allow building of a nearly complete subunit model using program O (Jones *et al.*, 1991). Bound dCMP was also clearly visible. The second subunit was generated by applying the NCS relationship. The structure was refined with X-PLOR (Brünger, 1993) against native II data, extending the high-resolution limit to 1.6 Å in steps. The NCS between the two independent subunits in the asymmetric unit was maintained with tight restraint during the early stages of the refinement but was relaxed in the final rounds. Individual isotropic B-factors, initially set to 20 Å<sup>2</sup>, were refined in the last stages of the refinement with restraint. Solvent molecules were placed by searching the model-phased ( $F_o - F_c$ ) maps and a bulk solvent correction was applied. Other models were also refined with X-PLOR (Brünger, 1993), after a rigid-body refinement at 3.5 Å resolution.

### Structure analysis and atomic coordinates

The stereochemistry of the model was assessed with PROCHECK (Laskowski *et al.*, 1993). Models were displayed with program O (Jones *et al.*, 1991) and CHAIN (Sack, 1988). The secondary structure elements were assigned by PROCHECK (Laskowski *et al.*, 1993). Model superpositions were done by LSQKAB in CCP4 (CCP4, 1994). Coordinates and structure factors have been deposited in the Brookhaven Protein Data Bank for immediate release [accession codes: 1b5e for dCMP-bound His-tagged enzyme; 1b49 for Pi-bound His-tagged enzyme; and 1b5d for dCMP-bound His-untagged enzyme].

## Acknowledgements

We thank Dr P.Artymiuk for running the PROTEP search and Dr D.Matthews for kindly providing us with the coordinates of bifunctional dihydrofolate reductase-thymidylate synthase from *L.major*. We also thank Professor N.Sakabe, Dr N.Watanabe and Dr M.Suzuki for assistance during data collection at BL-6B of Photon Factory, Japan, as part of the TARA Sakabe project (98G363). H.K.S. is supported by the Postdoctoral Fellowship from Korea Ministry of Education. This work was supported by grants from the Korea Ministry of Education and the Korea Science and Engineering Foundation through the Center for Molecular Catalysis at Seoul National University. The publication cost was partially supported by the Research Institute of Molecular Science, SNU.

## References

- Agarwalla,S., LaPorte,S., Liu,L., Finer-Moore,J., Stroud,R.M. and Santi,D.V. (1997) A novel dCMP methylase by engineering thymidylate synthase. *Biochemistry*, **36**, 15909–15917.
- Allen,J.R., Lasser,G.W., Goldman,D.A., Booth,J.W. and Mathews,C.K. (1983) T4 phage deoxyribonucleotide-synthesizing enzyme complex. Further studies on enzyme composition and regulation. *J. Biol. Chem.*, **258**, 5746–5753.
- Barton,G.J. (1993) ALSCRIPT: a tool to format multiple sequence alignments. *Protein Eng.*, **6**, 37–40.
- Brünger,A.T. (1993) *X-PLOR version 3.1*. Yale University Press, New Haven, CT.
- Butler,M.M., Graves,K.L. and Hardy,L.W. (1994) Evidence from <sup>18</sup>O exchange studies for an exocyclic methylene intermediate in the reaction catalyzed by T4 deoxycytidylate hydroxymethylase. *Biochemistry*, **33**, 10521–10526.
- Carreras,C.W. and Santi,D.V. (1995) The catalytic mechanism and structure of thymidylate synthase. *Annu. Rev. Biochem.*, **64**, 721–762.
- CCP4 (1994) Collaborative Computational Project Number 4. The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr.*, **D50**, 760–763.
- Cohen,F.E., Sternberg,M.J.E. and Taylor,W. (1981) Analysis of the tertiary structure of protein β-sheet sandwiches. *J. Mol. Biol.*, **148**, 253–272.
- Cox,G.S. and Conway,T.W. (1973) Template properties of glucose-deficient T-even bacteriophage DNA. *J. Virol.*, **12**, 1279–1287.
- Elcock,A.H., Potter,M.J., Matthews,D.A., Knighton,D.R. and McCammon,J.A. (1996) Electrostatic channeling in the bifunctional enzyme dihydrofolate reductase-thymidylate synthase. *J. Mol. Biol.*, **262**, 370–374.
- Fauman,E.B., Rutenber,E.E., Maley,G.F., Maley,F. and Stroud,R.M. (1994) Water-mediated substrate/product discrimination: the product

- complex of thymidylate synthase at 1.83 Å. *Biochemistry*, **33**, 1502–1511.
- Finer-Moore, J.S., Maley, G.F., Maley, F., Montfort, W.R. and Stroud, R.M. (1994) Crystal structure of thymidylate synthase from T4 phage: component of a deoxynucleoside triphosphate-synthesizing complex. *Biochemistry*, **33**, 15459–15468.
- Finer-Moore, J.S., Liu, L., Birdsall, D.L., Brem, R., Apfeld, J., Santi, D.V. and Stroud, R.M. (1998) Contributions of orientation and hydrogen bonding to catalysis in Asn229 mutants of thymidylate synthase. *J. Mol. Biol.*, **276**, 113–129.
- Gommers-Ampt, J.H. and Borst, P. (1995) Hypermodified bases in DNA. *FASEB J.*, **9**, 1034–1042.
- Gommers-Ampt, J.H., Van Leeuwen, F., de Beer, A.L.J., Vliegthart, F.G., Dizdaroglu, M., Kowalak, J.A., Crain, P.F. and Borst, P. (1993) β-D-Glucosyl-hydroxymethyluracil: A novel modified base present in the DNA of the parasitic protozoan *T. brucei*. *Cell*, **75**, 1129–1136.
- Graves, K.L., Butler, M.M. and Hardy, L.W. (1992) Roles of Cys148 and Asp179 in catalysis by deoxycytidylate hydroxymethylase from bacteriophage T4 examined by site-directed mutagenesis. *Biochemistry*, **31**, 10315–10321.
- Greenberg, G.R., He, P., Hilfinger, J. and Tseng, M.J. (1994) Deoxyribonucleoside triphosphate synthesis and phage T4 DNA replication. In Karam, J.D. (ed.), *Molecular Biology of Bacteriophage T4*. American Society of Microbiology, Washington, DC, pp. 14–27.
- Grindley, H.M., Artymiuk, P.J., Rice, D.W. and Willett, P. (1993) Identification of tertiary structure resemblance in proteins using a maximal common subgraph isomorphism algorithm. *J. Mol. Biol.*, **229**, 707–721.
- Hardy, L.W., Finer-Moore, J.S., Montfort, W.R., Jones, M.O., Santi, D.V. and Stroud, R.M. (1987) Atomic structure of thymidylate synthase: target for rational drug design. *Science*, **235**, 448–455.
- Holm, L. and Sander, C. (1993) Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.*, **233**, 123–138.
- Horton, M. and Lewis, M. (1992) Calculation of the free energy of association for protein complexes. *Protein Sci.*, **1**, 169–181.
- Jackson, R.M., Gabb, H.A. and Sternberg, M.J.E. (1998) Rapid refinement of protein interfaces incorporating solvation: application to the docking problem. *J. Mol. Biol.*, **276**, 265–285.
- Janin, J. and Rodier, F. (1995) Protein–protein interaction at crystal contacts. *Proteins*, **23**, 580–587.
- Jones, S. and Thornton, J.M. (1996) Principles of protein–protein interactions. *Proc. Natl Acad. Sci. USA*, **93**, 13–20.
- Jones, T.A., Zou, J.-Y., Cowan, S.W. and Kjeldgaard, M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr.*, **A47**, 110–119.
- Knighton, D.R., Kan, C.-C., Howland, E., Janson, C.A., Zuzana, H., Welsh, K.M. and Matthews, D.A. (1994) Structure of and kinetic channelling in bifunctional dihydrofolate reductase-thymidylate synthase. *Nature Struct. Biol.*, **1**, 186–194.
- Kraulis, P.J. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.*, **24**, 946–950.
- Lamm, N., Tomaschewski, J. and Rüger, W. (1987) Nucleotide sequence of the deoxycytidylate hydroxymethylase gene of bacteriophage T4 (g42) and the homology of its gene product with thymidylate synthase of *E.coli*. *Nucleic Acids Res.*, **15**, 3930.
- Lamm, N., Wang, Y., Mathews, C.K. and Rüger, W. (1988) Deoxycytidylate hydroxymethylase gene of bacteriophage T4. *Eur. J. Biochem.*, **172**, 553–563.
- Laskowski, R.A., MacArthur, M.W., Moss, D.S. and Thornton, J.M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.*, **26**, 283–291.
- Liu, L. and Santi, D.V. (1993) Exclusion of 2'-deoxycytidine 5'-monophosphate by asparagine 229 of thymidylate synthase. *Biochemistry*, **32**, 9263–9267.
- Mathews, C.K., Wheeler, L.J., Ungermann, C., Young, J.P. and Ray, N.B. (1993) Enzyme interactions involving T4 phage-coded thymidylate synthase and deoxycytidylate hydroxymethylase. *Adv. Exp. Med. Biol.*, **338**, 563–570.
- Matthews, D.A., Appelt, K., Oatley, S.J. and Xuong, N.H. (1990a) Crystal structure of *Escherichia coli* thymidylate synthase containing bound 5-fluoro-2'-deoxyuridylate and 10-propargyl-5,8-dideazafofolate. *J. Mol. Biol.*, **214**, 923–936.
- Matthews, D.A., Villafranca, J.E., Janson, C.A., Smith, W.W., Welsh, K. and Freer, S. (1990b) Stereochemical mechanism of action for thymidylate synthase based on the X-ray structure of the covalent inhibitory ternary complex with 5-fluoro-2'-deoxyuridylate and 5,10-methylene-tetrahydrofolate. *J. Mol. Biol.*, **214**, 937–948.
- McGaughy, K.M., Wheeler, L.J., Moore, J.T., Maley, G.F. and Mathews, C.K. (1996) Protein–protein interactions involving T4 phage-coded deoxycytidylate deaminase and thymidylate synthase. *J. Biol. Chem.*, **271**, 23037–23042.
- Merritt, E.A. and Murphy, M.E.P. (1994) Raster3D Version 2.0—a program for photorealistic molecular graphics. *Acta Crystallogr.*, **D50**, 869–873.
- Montfort, W.R., Perry, K.M., Fauman, E.B., Finer-Moore, J.S., Maley, G.F., Hardy, L., Maley, F. and Stroud, R.M. (1990) Structure, multiple site binding and segmental accommodation in thymidylate synthase on binding dUMP and anti-folate. *Biochemistry*, **29**, 6964–6977.
- Nicholls, A., Sharp, K.A. and Honig, B. (1991) Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins*, **11**, 281–293.
- Perry, K.M., Carreras, C.W., Chang, L.C., Santi, D.V. and Stroud, R.M. (1993) Structures of thymidylate synthase with a C-terminal deletion: role of the C-terminus in alignment of 2'-deoxyuridine 5'-monophosphate and 5,10-methylenetetrahydrofolate. *Biochemistry*, **32**, 7116–7125.
- Rice, P.A. and Steitz, T.A. (1994) Model for a DNA-mediated synaptic complex suggested by crystal packing of γδ resolvase subunits. *EMBO J.*, **13**, 1514–1524.
- Rüger, W. (1978) Transcription of bacteriophage T4 DNA *in vitro*: Selective initiation with dinucleotides. *Eur. J. Biochem.*, **88**, 109–117.
- Rutenber, E.E. and Stroud, R.M. (1996) Binding of the anticancer drug ZD1694 to *E.coli* thymidylate synthase: assessing specificity and affinity. *Structure*, **4**, 1317–1324.
- Sack, J.S. (1988) CHAIN—A crystallographic modelling program. *J. Mol. Graph.*, **6**, 244–245.
- Sage, C.R., Rutenber, E.E., Stout, T.J. and Stroud, R.M. (1996) An essential role for water in an enzyme reaction mechanism: The crystal structure of the thymidylate synthase mutant E58Q. *Biochemistry*, **35**, 16270–16281.
- Schellenberger, U., Livi, L.L. and Santi, D.V. (1995) Cloning, expression, purification and characterization of 2'-deoxyuridylate hydroxymethylase from phage SPO1. *Protein Expr. Purif.*, **6**, 423–430.
- Shapiro, L. et al. (1995) Structural basis of cell–cell adhesion by cadherins. *Nature*, **374**, 327–337.
- Sohn, S.H. et al. (1999) Crystallization and preliminary X-ray crystallographic analysis of deoxycytidylate hydroxymethylase from bacteriophage T4. *Acta Crystallogr.*, **D55**, in press.
- Story, R.M., Weber, I.T. and Steitz, T.A. (1992) The structure of the *E.coli* recA protein monomer and polymer. *Nature*, **335**, 318–325.
- Stout, T.J., Schellenberger, U., Santi, D.V. and Stroud, R.M. (1998) Crystal structure of a unique thermal-stable thymidylate synthase from *Bacillus subtilis*. *Biochemistry*, **37**, 14736–14747.
- Stroud, R.M. and Finer-Moore, J.S. (1993) Stereochemistry of a multistep/bipartite methyl transfer reaction: thymidylate synthase. *FASEB J.*, **7**, 671–677.
- Stroud, R.M. (1994) An electrostatic highway. *Nature Struct. Biol.*, **1**, 131–134.
- Terwilliger, T.C. and Berendzen, J. (1996) Correlated phasing of multiple isomorphous replacement data. *Acta Crystallogr.*, **D52**, 749–757.
- Thylen, C. (1988) Expression and DNA sequence of the cloned bacteriophage T4 dCMP hydroxymethylase gene. *J. Bacteriol.*, **170**, 1994–1998.
- Vrieland, A., Rüger, W., Driessen, H.P.C. and Freemont, P.S. (1994) Crystal structure of the DNA modifying enzyme β-D-glucosyltransferase in the presence and absence of the substrate uridine diphosphoglucose. *EMBO J.*, **13**, 3413–3422.
- Warren, R.A.J. (1980) Modified bases in bacteriophage DNAs. *Annu. Rev. Microbiol.*, **34**, 137–158.
- Wheeler, L., Wang, Y. and Mathews, C.K. (1992) Specific associations of T4 bacteriophage proteins with immobilized deoxycytidylate hydroxymethylase. *J. Biol. Chem.*, **267**, 7664–7670.
- Wheeler, L., Ray, N.B., Ungermann, C., Hendricks, S.P., Bernard, M.A., Hanson, E.S. and Mathews, C.K. (1996) T4 phage gene 32 protein as a candidate organizing factor for the deoxyribonucleoside triphosphate synthetase complex. *J. Biol. Chem.*, **271**, 11156–11162.
- Wilhelm, K. and Rüger, W. (1992) Deoxyuridylate-hydroxymethylase of bacteriophage SPO1. *Virology*, **189**, 640–646.
- Young, J.P. and Mathews, C.K. (1992) Interactions between T4 phage-coded deoxycytidylate hydroxymethylase and thymidylate synthase as revealed with an anti-idiotypic antibody. *J. Biol. Chem.*, **267**, 10786–10790.

Received December 10, 1998; revised and accepted January 13, 1999