

**CRYSTALLOGRAPHIC STUDIES OF THYMIDYLATE SYNTHASE:  
STRUCTURE OF A MAMMALIAN ENZYME AND ANALYSES OF  
INVARIANT NON-CATALYTIC RESIDUES IN A BACTERIAL ENZYME**

by

Rogério Rafael Sotelo-Mundo

---

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF BIOCHEMISTRY

In Partial Fulfillment of the Requirements  
For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

1 9 9 9

## INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

**The quality of this reproduction is dependent upon the quality of the copy submitted.** Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

**UMI<sup>®</sup>**

Bell & Howell Information and Learning  
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA  
800-521-0600



**CRYSTALLOGRAPHIC STUDIES OF THYMIDYLATE SYNTHASE:  
STRUCTURE OF A MAMMALIAN ENZYME AND ANALYSES OF  
INVARIANT NON-CATALYTIC RESIDUES IN A BACTERIAL ENZYME**

by

Rogério Rafael Sotelo-Mundo

---

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF BIOCHEMISTRY

In Partial Fulfillment of the Requirements  
For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

1 9 9 9



**UMI Number: 9946791**

---

**UMI Microform 9946791**

**Copyright 1999, by UMI Company. All rights reserved.**

**This microform edition is protected against unauthorized  
copying under Title 17, United States Code.**

---

**UMI**

**300 North Zeeb Road  
Ann Arbor, MI 48103**

THE UNIVERSITY OF ARIZONA ®  
GRADUATE COLLEGE

As members of the Final Examination Committee, we certify that we have  
read the dissertation prepared by Rogério R. Sotelo-Mundo

entitled Crystallographic Studies of Thymidylate Synthase: Structure  
of a Mammalian Enzyme and Analyses of Invariant Non-catalytic  
Residues in a Bacterial Enzyme

and recommend that it be accepted as fulfilling the dissertation  
requirement for the Degree of Doctor of Philosophy

William R. Montfort

6/25/99  
Date

Gordon Tollin

6/25/99  
Date

Michael A. Wells

6/25/99  
Date

Mark S. Dodson

6/25/99  
Date

John A. Rupley

6/25/99  
Date

Final approval and acceptance of this dissertation is contingent upon  
the candidate's submission of the final copy of the dissertation to the  
Graduate College.

I hereby certify that I have read this dissertation prepared under my  
direction and recommend that it be accepted as fulfilling the dissertation  
requirement.

William R. Montfort  
Dissertation Director


William R. Montfort

8/10/99  
Date

### STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: 

## ACKNOWLEDGEMENTS

I would like to thank my advisor Dr William R. Montfort for his advice and support. I would also like to thank Drs. Frank and Gladys Maley (Wadsworth Center, Albany NY), Dr. Larry Hardy (U. Massachusetts, Worcester MA) and Dr. W. Rode, (Nencki Institute, Warsaw, Poland) for their collaboration in the TS project. I also want to thank Drs. Garth Powis and John Gasdaska from the Arizona Cancer Center for their collaboration on the human thioredoxin project.

To former and current members of the Montfort lab: Sue Roberts, Andrzej Weichsel, Virginia Dress, Jerry Honts, David Hyatt, David Sanders, Bryan Arendall, Pavel Strop, Celia Balfour, John Andersen, thanks for your patience, help and good advice.

I also thank the fellowship received from the National Council on Science and Technology (CONACyT México) and Research Center for Food Science and Technology (CIAD, A.C.), and in particular to Drs. Inocencio Higuera Ciapara, Gloria Yepiz Plascencia and Francisco Vargas Albores, for their sustained support during all this years. Special thanks also to Drs. Alberto Gonzalez, Rosalba Troncoso and Juan Pedro Camou from CIAD.

Finally, to my parents, grandparents and relatives, who have helped us many times and in many ways. It would have been impossible to accomplish this work without your help. Thanks!

## **DEDICATION**

To my wife Maria and my daughter Anna Paula, for their patience and love.

## TABLE OF CONTENTS

	Page
LIST OF ILLUSTRATIONS .....	8
LIST OF ABBREVIATIONS.....	10
ABSTRACT.....	11
 1. INTRODUCTION.....	 13
1.1. TS Function, Fold and Active Site Structure .....	15
1.2. Catalytic Mechanism .....	26
1.3. Ligand-Induced Conformational Change.....	31
1.4. Inhibitors and Anticancer Drugs .....	34
1.5. Mammalian TSs.....	37
1.6. Objectives and Approach.....	38
 2. MATERIALS and METHODS .....	 39
2.1. TS Protein .....	39
2.2. Protein Crystallization .....	40
2.2.1. <i>E. coli</i> TS crystallization.....	40
2.2.2. Rat TS crystallization.....	41
2.3. Data collection .....	42
2.4. Structure determinations .....	42
2.4.1. Difference Fourier Maps.....	43
2.4.2. Molecular Replacement .....	44
2.4.3. Crystallographic refinement.....	46
2.5. Comparison of three-dimensional structures .....	47
 3. CRYSTAL STRUCTURE OF THE WILD-TYPE RAT TS-dUMP-Tomudex COMPLEX .....	 50
3.1. Introduction.....	50
3.2. Results.....	52
3.2.1. Structure of rat TS-dUMP-Tomudex complex (crystal form 1) .....	52
3.2.2. Structure of rat TS-dUMP-Tomudex (crystal form 2) .....	56
3.2.3. Ligand binding in rat TS-dUMP-Tomudex .....	57
3.2.4. Comparison of rat TS with the bacterial homologues .....	63
3.2.5. Comparison of human and rat TS .....	67
3.3. Discussion.....	68
 4. CRYSTAL STRUCTURES OF THE K48Q and R166Q <i>ESCHERICHIA COLI</i> TS MUTANTS .....	 71
4.1. Introduction.....	71
4.2. Structure of <i>E. coli</i> TS K48Q mutant complexes .....	72

## TABLE OF CONTENTS - *Continued*

	Page
4.2.1. K48Q-dUMP Binary Complex .....	74
4.2.2. K48Q-5NO <sub>2</sub> dUMP Binary Complex .....	81
4.2.3. Apo-K48Q crystallized in the presence of dUMP-PDDF and FdUMP-CH <sub>2</sub> H <sub>4</sub> folate .....	86
4.2.4. Discussion .....	87
4.3. Structures of the <i>E. coli</i> TS R166Q mutant complexes .....	90
4.3.1. Apo-R166Q crystallized with dUMP .....	93
4.3.2 R166Q crystallized in the presence of dUMP-PDDF and R166Q- FdUMP-CH <sub>2</sub> H <sub>4</sub> folate .....	97
5. SUMMARY AND CONCLUSIONS .....	104
5.1 Future Studies .....	107
APPENDIX A .....	110
Crystal structure of human thioredoxin oxidized by a potential anti-cancer drug .....	110
A.1. Introduction .....	110
A.2. Materials and methods .....	114
A.2.1. Protein overexpression and purification .....	114
A.2.2. Crystallization and data collection .....	115
A.2.3 Phase Determination, Model Building and Structure Refinement..	116
A.3 Results .....	117
A.3.1. Crystallization of IV-2/TRX .....	117
A.3.2. Determination of the structure of IV-2 oxidized TRX .....	119
A.3.3. Conformational change in TRX produced by oxidation .....	120
A.4 Discussion .....	124
REFERENCES .....	127

## LIST OF ILLUSTRATIONS

	Page
Figure 1.1. Proposed simplified mechanism of catalysis by TS.....	16
Figure 1.2. Chemical structure of TS substrates and inhibitors discussed in the text. ....	17
Figure 1.3. Ribbon representation of TS. ....	20
Figure 1.4. Diagram of the nucleotide-binding site of dUMP on the <i>E. coli</i> TS active site. .....	23
Figure 1.5. Diagram of the cofactor CH <sub>2</sub> H <sub>4</sub> folate binding site in the <i>E. coli</i> TS active site. .....	24
Figure 1.6. Proposed detailed TS mechanism with an alternative folate elimination route, from Montfort and Weichsel (1997). ....	28
Figure 1.7. Ligand-induced conformational change in <i>E. coli</i> TS.....	33
Figure 3.1. Ribbon diagram of the rat TS-dUMP-Tomudex ternary complex. ....	51
Figure 3.2. Stereo view of a SA-omit map of the ligands and catalytic cysteine r189, crystal form 1.....	54
Figure 3.3. Stereo view of a SA-omit map of the ligands and catalytic cysteine r189, crystal form 2.....	58
Figure 3.4. Stereo view of dUMP binding in rat TS. ....	60
Figure 3.5. Stereo view of Tomudex binding in rat TS.....	61
Figure 3.6. Superposition of rat and <i>E. coli</i> TS-dUMP-Tomudex complexes. ....	64
Figure 4.1. Contacts of K48 in the wild type <i>E. coli</i> TS binary and ternary complexes. ..	73
Figure 4.2. Contacts of Q48 on the mutant binary complex. ....	77
Figure 4.3. SA-omit map of the active site in the K48Q-dUMP mutant complex. ....	79
Figure 4.4. Superposition of the wild type and K48Q mutant binary complexes. ....	80
Figure 4.5. SA-omit map for the K48Q mutant binary complex with 5-NO <sub>2</sub> dUMP. ....	83
Figure 4.6 Stereo views of the nucleotide binding site in the binary complex of the K48Q TS mutant with 5-NO <sub>2</sub> dUMP.....	85
Figure 4.7. Stereo view of the mutation site on the apo-R166Q TS. ....	94
Figure 4.8. SA-omit map of the apo-R166Q TS mutant. ....	96
Figure 4.9. SA-omit map of the complex of the R166Q-TS mutant with PDDF and SO <sub>4</sub> .. .....	101
Figure 4.10. Superposition of the R166Q-PDDF and the TS-dUMP-PDDF complexes. .....	102
Figure A.1. Mechanism of disulfide exchange by TRX, and chemical structure of the disulfide TRX-inhibitor IV-2. ....	111
Figure A.2. Ribbon diagram of the oxidized human thioredoxin dimer. ....	121
Figure A.3. Difference electron density maps between reduced and oxidized human TRX. .....	122
Figure A.4. Electron density map of the oxidized human TRX active site.....	123
Figure A.5. Effects on hydrogen bonding at the human TRX dimer interface by changes in pH and oxidation.....	125



## LIST OF TABLES

	Page
Table 1.1. Secondary structure elements in <i>E. coli</i> TS.....	21
Table 2.1. Sequence alignment of selected TSs. ....	49
Table 3.1 Crystallographic data of the rat TS complexes with dUMP and Tomudex.....	53
Table 4.1 Crystallographic data for the <i>E. coli</i> TS K48Q mutant structures.....	75
Table 4.3 Comparison of wild type ternary complexes and R166 mutant <i>E. coli</i> TS structures .....	103
Table A.1. Crystallographic results for IV-2 oxidized wild type human thioredoxin....	118

## LIST OF ABBREVIATIONS

<b>BME</b>	$\beta$ -mercaptoethanol
<b>CH<sub>2</sub>H<sub>4</sub>folate</b>	5,10-methylene-5,6,7,8-tetrahydrofolate
<b>dCMP</b>	2'-deoxycytidine 5'-monophosphate
<b>DHFR</b>	Dihydrofolate reductase
<b>dTMP</b>	2'-deoxythymidine 5'-monophosphate
<b>DTT</b>	Dithiotreitol
<b>dUMP</b>	2'-deoxyuridine 5'-monophosphate
<b>FdUMP</b>	5'-fluoro-2'-deoxyuridylate
<b>FPGS</b>	Folyl-polyglutamate synthetase
<b>IV-2</b>	1-methylpropyl 2-imidazolyl disulfide
<b>NCS</b>	Non-crystallographic symmetry
<b>PABA</b>	<i>para</i> -aminobenzoic acid
<b>PDDF</b>	10-propargyl-5,8-dideazafolate
<b>RMSD</b>	Root-mean square deviation
<b>SA</b>	Simulated annealing
<b>TRX</b>	Thioredoxin
<b>TS</b>	Thymidylate synthase

## ABSTRACT

Thymidylate synthase (TS, EC 2.1.1.45) is the enzyme responsible for the synthesis of 2'-deoxythymidine 5'-monophosphate (dTMP), using 2'-deoxyuridine 5'-monophosphate (dUMP) as the substrate and 5,10-methylene-5,6,7,8-tetrahydrofolate ( $\text{CH}_2\text{H}_4\text{folate}$ ) as both carbon donor and reductant. Inhibition of TS stops growth of rapidly dividing cells, and for decades TS inhibitors such as 5-fluorouracil, and more recently folyl-analogs, have been used as anticancer drugs. However, prior to my studies, there were no structures available of any ligand-bound mammalian TS. In this dissertation I present the crystal structure of rat TS bound to the substrate dUMP and the anticancer drug Tomudex. Unexpectedly, the enzyme has an open conformation, with ligands bound but no covalent adduct between the catalytic cysteine (Cys 189) and the nucleotide, unlike that reported for the same complex with the *E. coli* enzyme. Three changes in amino acid sequence between the *E. coli* and rat TS proteins, namely ecT78->rR101, ecW83->rN106 and ecV262->rM305, result in loss of van der Waals contacts with Tomudex. These changes coupled with the loss of a hydrogen bond between the Tomudex 2-quinazoline position, which has been changed from the amino group of the cofactor to a methyl group, suggest that Tomudex may inhibit mammalian TS by stabilizing the open conformation.

In a second project, I have studied the role of two conserved residues, K48 and R166 in catalysis of *E. coli* TS. Mutation of each of these residues to glutamine produces nearly inactive proteins. Crystallographic analyses of K48Q and R166Q suggest that the loss of these charged groups reduces binding of the ternary covalent intermediate.

Superposition of the mutated structures with a previously determined wild type structure containing a close analog of this intermediate (TS-FdUMP-CH<sub>2</sub>H<sub>4</sub>folate), reveals that the mutants are either more open or distorted, and therefore unable to contact the ligands properly. Both  $K_m$  and  $k_{cat}$  are altered for the two enzymes, with  $K_m$  increasing about 10-fold, and  $k_{cat}$  reduced 400-fold (K48Q) and 3,400-fold (R166Q). Taken together, these data suggest that K48Q and R166Q bind weakly one or more of the reaction intermediates, leading to near inactivity.

## CHAPTER 1

### INTRODUCTION

The nature of enzyme catalysis can not be separated from protein folding and structure, since the forces that keep the protein structure are the same as those that allow substrates to bind into the enzyme active site (Knowles, 1987). The structural analyses of biological macromolecules have provided important insights into their biological function. In particular, the determination of protein structure is fundamental for identifying residues involved in enzyme catalysis or for ligand binding in transport proteins. The availability of recombinant DNA and protein expression technology, combined with biophysical techniques, allows the biochemist to thoroughly study structure and function of proteins.

Thymidylate synthase (TS) is one of the best examples for which decades of biochemical, structural and clinical research have led to a detailed molecular model for biological catalysis and to the design of anticancer drugs. TS is the only enzyme to produce the DNA-precursor dTMP *de novo* in all species studied so far. Therefore,

inhibition of this pathway impairs cell division by blocking DNA replication, and one TS inhibitor is in clinical use for treating cancer, and several others are in clinical trials. TS also provides a convenient model for the study of the role of conformational change in non-allosteric enzymes, and as a paradigm for structure-based drug design against proliferative diseases such as cancer, bacterial infections and parasitism (Shoichet *et al.*, 1993, Stout *et al.*, 1999).

This dissertation is focused on two key aspects of TS function: the mammalian TS structure and the role of conserved amino acids. For the first aim, I have determined the structure of a ligand-bound mammalian TS complex, to investigate the ligand-induced conformational change and the inhibitor-enzyme interactions in the active site. Before this work, the only structural information available about mammalian TS was the crystal structure of unliganded human TS (Schiffer *et al.*, 1995). Since human TS does not crystallize in the presence of ligands, I have worked with rat TS, which is 90% identical to the human enzyme. I have determined the structure of two crystal forms of rat TS in complexes with the substrate dUMP and the anticancer drug Tomudex (reviewed in Jackman *et al.*, 1996). These structures are the first views of a ligand-bound mammalian TS active site (Sotelo-Mundo *et al.*, 1999).

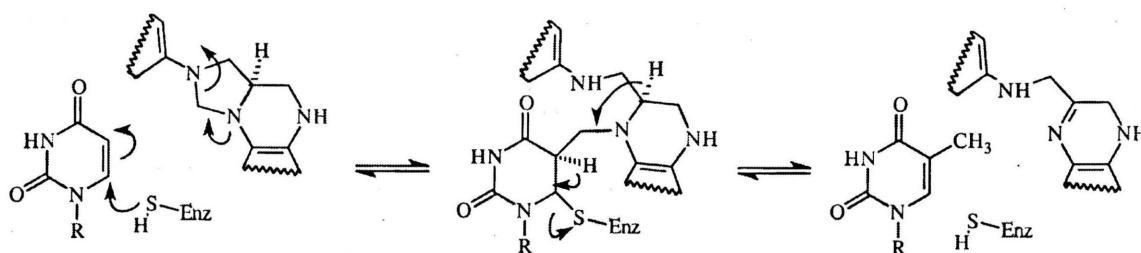
In a second project, I used protein crystallography to investigate the role of conserved residues in the structure and function of TS. I determined the crystal structure of two single-residue mutants of *E. coli* TS, K48Q and R166Q, in complexes with substrates and inhibitor. K48 is a conserved residue that contacts the glutamyl-tail of the cofactor *via* a water molecule, and R166 makes hydrogen bonds to the nucleotide

phosphate of the substrate and is in van der Waals contact with the catalytic cysteine 146. Mutation of each of these residues to glutamine renders a nearly inactive enzyme. Before describing these results in later chapters, I review the TS biochemical function, structure, mechanism and other relevant background information in this introduction.

### 1.1. TS Function, Fold and Active Site Structure

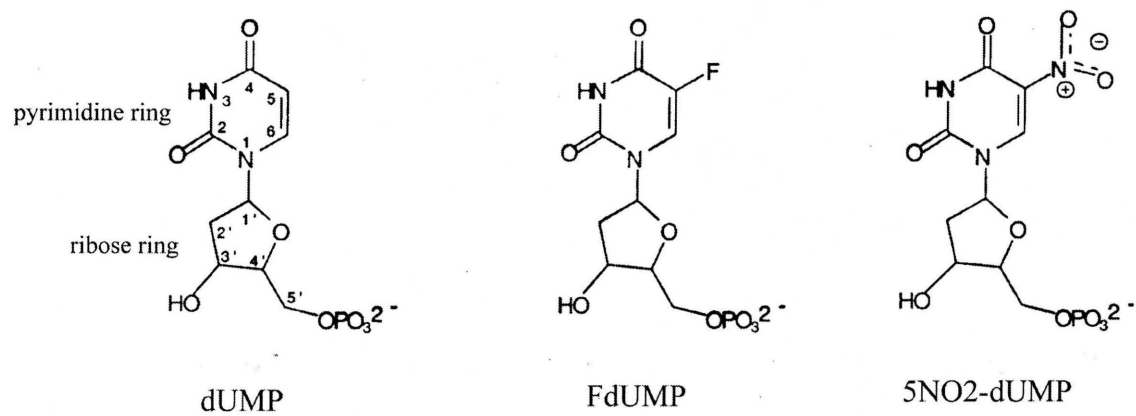
Thymidylate synthase (TS, EC 2.1.1.45) catalyzes the reductive methylation of 2'-deoxyuridine 5'-monophosphate (dUMP) by 5,10-methylene-5,6,7,8-tetrahydrofolate ( $\text{CH}_2\text{H}_4\text{folate}$ ) to produce 7,8-dihydrofolate ( $\text{H}_2\text{folate}$ ) and 2'-deoxythymidine 5'-monophosphate (dTMP), see Figure 1.1. (reviewed in Carreras and Santi, 1995). Since one mole of  $\text{CH}_2\text{H}_4\text{folate}$  is consumed per mole of dTMP produced, the reaction is coupled to the regeneration of folate. This is accomplished by dihydrofolate reductase (DHFR) and serine hydroxymethyltransferase, which convert  $\text{H}_2\text{folate}$  back to  $\text{CH}_2\text{H}_4\text{folate}$  (Mathews and Van Holde, 1996).

The structure of TS has been determined from several species: *E. coli* (Matthews *et al.*, 1990a, Matthews *et al.*, 1990b, Montfort *et al.*, 1990), *L. casei* (Hardy *et al.*, 1987), human (Schiffer *et al.*, 1995), phage T4 (Finer-Moore *et al.*, 1994), bifunctional TS-DHFR from *L. major* (Knighton *et al.*, 1994), *B. subtilis* (Stout *et al.*, 1998, Fox *et al.*, 1999) and rat (Sotelo-Mundo *et al.*, 1999). In total there are more than 70 crystal structures of both wild type and mutant TSs deposited in the Protein Data Bank.

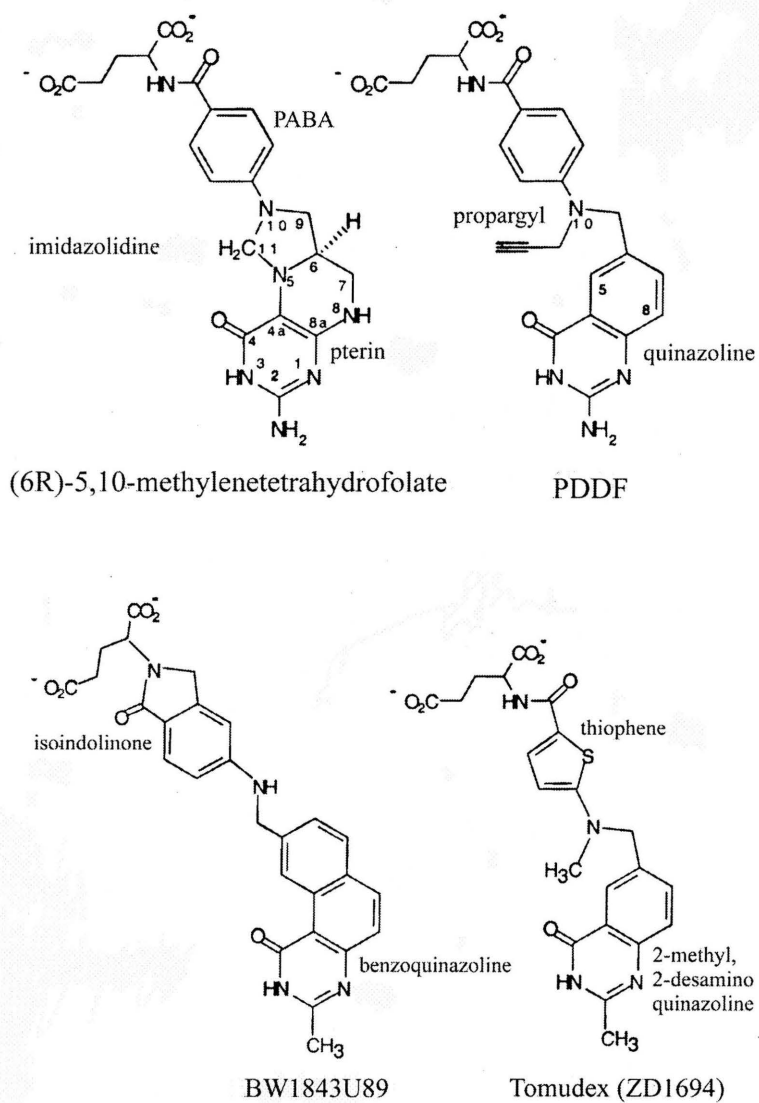


**Figure 1.1. Proposed simplified mechanism of catalysis by TS.** The substrate dUMP is attacked by the active site cysteine, to form a covalent adduct. The five-membered ring of CH<sub>2</sub>H<sub>4</sub>folate opens, leading to a transient ternary covalent complex. A proton is abstracted and the folate reduces the methylene bridge to produce H<sub>2</sub>folate and dTMP (after Carreras and Santi, 1995)





**Figure 1.2. Chemical structures of TS substrates and inhibitors discussed in the text.**



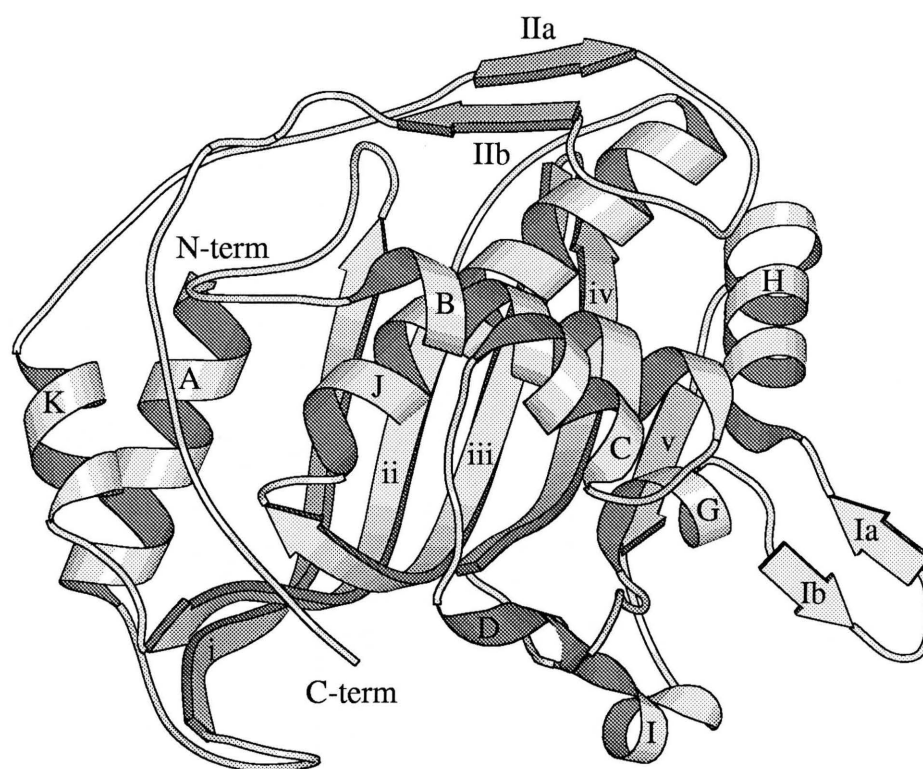
**Figure 1.2. - Continued.**

TS is a functional dimer of identical subunits where residues from each monomer contribute to each of two active sites. Each subunit has a molecular weight of approximately 35 kDa. The ligands are contacted by 19 residues, mostly from a single subunit, however two arginines from one monomer (R126' and R127') contact the nucleotide phosphate in the active site of the other.

All TSs known to date are dimeric, including the bifunctional DHFR-TSs found in protozoa and plants. The dimer interface is formed by a back-to-back apposition of two five stranded  $\alpha$ -sheets with a right-handed twist between them. Three contiguous  $\beta$ -strands bend sharply in a " $\beta$ -kink", and form one wall of the active site cavity. Two layers of helices and loops pack on top of the central  $\alpha$ -sheet (Hardy *et al.* , 1987).

The TS fold, shown in Figure 1.3, has been classified as an  $\alpha$ - $\beta$  2-layer sandwich (Orengo *et al.*, 1997), and it is conserved even in the bifunctional DHFR-TS (Knighton *et al.* , 1994). The secondary structural elements are named according to Montfort and collaborators (1990) and are listed in Table 1.1. This nomenclature will be used in the following chapters when referring to secondary structure.

The nucleotide-binding site is formed by seven amino acid residues in the binary complex with dUMP (shown in Figure 1.4). Two more residues, R21 and R127' contact the nucleotide-phosphate on the closed conformation. Asparagine 177 makes two hydrogen bonds to the pyrimidine ring, and these contacts define the identity of the nucleotide to be methylated.



**Figure 1.3. Ribbon representation of TS.** *E. coli* TS monomer from the structure of the ternary complex with dUMP and PDDF (spacegroup  $P6_3$ ). Ligands were omitted for clarity and secondary structure elements are named according to Montfort *et al.*, (1990)

**Table 1.1. Secondary structure elements in *E. coli* TS**

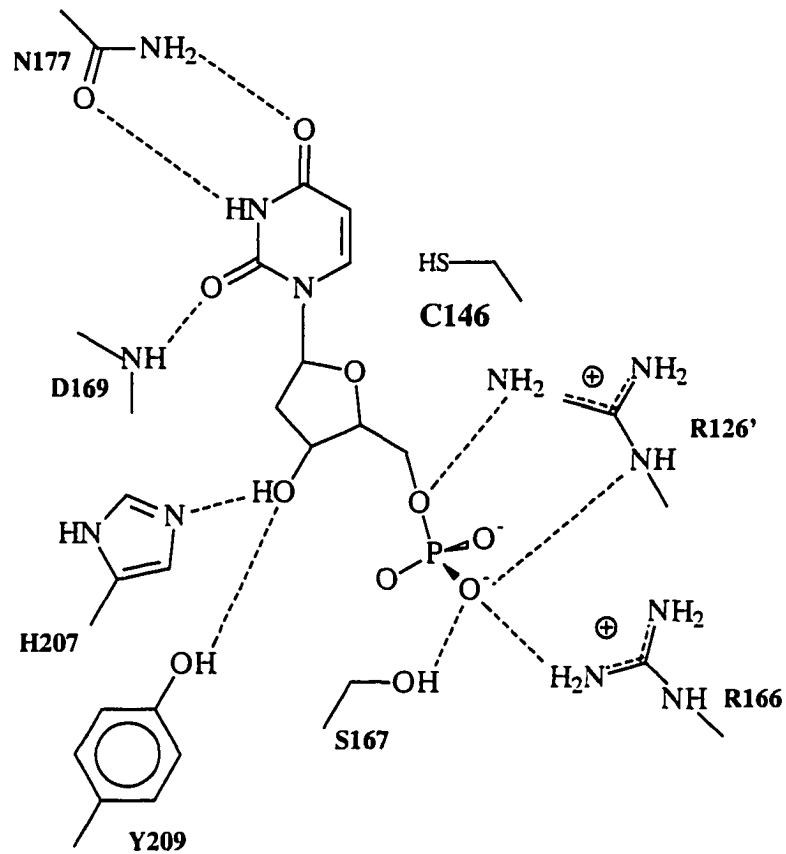
Name	Secondary structure	amino acid sequence
A	$\alpha$ -helix	1-14
B	"	52-65
C	"	69-76
G	"	93-100
H	"	110-122
I	"	137-141
J	"	173-193
K	"	212-219
i	$\beta$ -sheet	25-36
ii	"	198-210
iii	"	158-170
iv	"	146-155
v	"	128-132
Ia	$\beta$ -ribbon	101-103
Ib	"	107-109
IIa	"	230-234
IIb	"	246-250

From Montfort *et al.*, (1990)

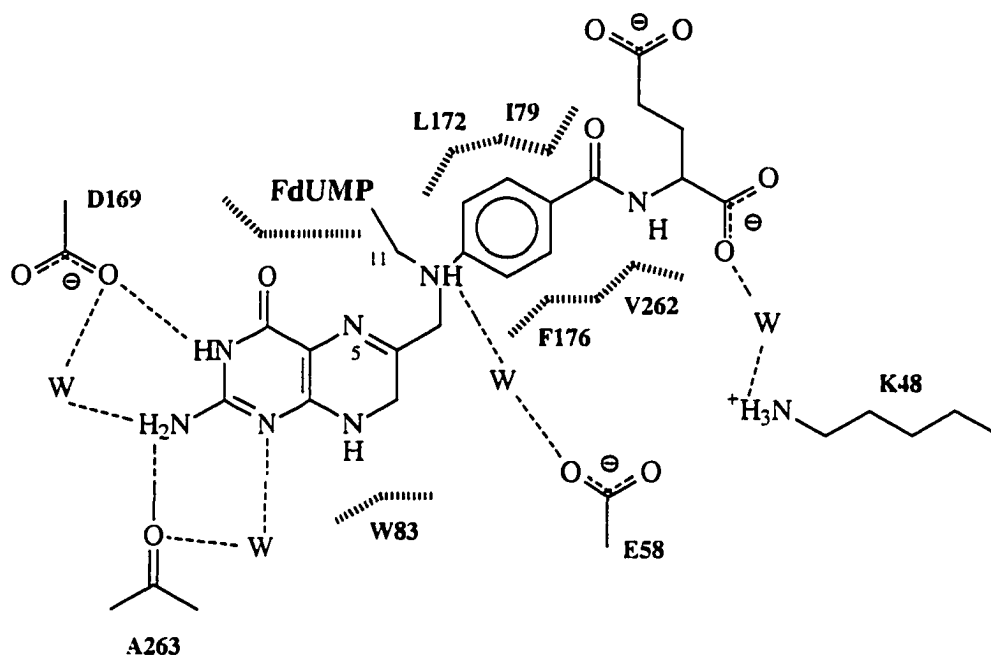
The *E. coli* N177D TS mutant binds and methylates both dCMP and dUMP (Liu and Santi, 1992). Furthermore, additional mutation of H147 to Ala increases eleven orders of magnitude the affinity of the mutant enzyme for dCMP versus dUMP, converting it into a true dCMP methylase (Agarwalla *et al.*, 1997). The main chain of Asp 169 also contacts the uracil ring, contributing one hydrogen bond to the nucleotide-binding site. The ribose 3'-hydroxyl makes two hydrogen bonds to H207 and Y209, and the rest of the contacts are made to the nucleotide phosphate by four arginines: R21, R166, R126' and R127'.

The role of these arginine residues has been investigated using site-directed mutagenesis. Substitution of R126' for the acidic residue glutamate increases  $K_m$  600-fold and decreases  $k_{cat}$  2,000-fold. The structure of the R126'E shows that the mutant binds dUMP very poorly, and when the nucleotide and the cofactor-analog PDDF are included the nucleotide-phosphate shifts by 1 Å compared to the wild type complex (Strop *et al.*, 1997). However, the adjacent R127', despite being invariant, is not critical for catalysis (Santi *et al.*, 1990). As for R166, mutation for any of the other 19 amino acids does not allow rescue of a thymidine auxotroph *L. casei* strain (Michaels *et al.*, 1990). Moreover, its mutation to glutamine (R166Q) reduces  $k_{cat}$  by more than 3,000 times compared to the wild type enzyme.

In contrast to dUMP, the TS cofactor CH<sub>2</sub>H<sub>4</sub>folate, or its close analogue PDDF, bind to the TS active site largely through van der Waals contacts with the pyrimidine ring of the nucleotide and nearby hydrophobic residues (Figure 1.5.).



**Figure 1.4. Diagram of the nucleotide-binding site of dUMP on the *E. coli* TS active site.** Based on the *E. coli* TS binary complex with dUMP (Roberts and Montfort, in preparation). Hydrogen bonds represented as dotted lines. R21 and R127' move towards the nucleotide-phosphate upon binding of the folate and active site closure.



**Figure 1.5. Diagram of the cofactor CH<sub>2</sub>H<sub>4</sub>folate binding site in the *E. coli* TS active site.** Based on the structure crystal structure of the covalent complex of *E. coli* TS with FdUMP and CH<sub>2</sub>H<sub>4</sub>folate (Hyatt *et al*, 1997). The five-membered ring is open and in covalent bond with FdUMP. Hydrogen bonds represented as dotted lines, and van der Waals contacts with hash lines. FdUMP has been omitted for clarity. Water molecules are labeled as "W".



The CH<sub>2</sub>H<sub>4</sub>folate-pterin or PDDF-quinazoline rings stack with the pyrimidine of the nucleotide, and the folate is further shielded from solvent by W80 and W83. The folyl-PABA ring makes van der Waals contacts with I79, L172 and F176. All these interactions restrain the folate to a "bent" conformation, compared to the extended structure of the folate in solution (Poe *et al.*, 1979).

The folate only makes two direct hydrogen bonds to the enzyme, between Asp 169 and the carbonyl of Ala 263 and the pterin rings of the folates (Matthews *et al.*, 1990a, Fauman *et al.*, 1994, Hyatt *et al.*, 1997). TS also contacts the folate indirectly through two ordered water molecules. The penultimate-residue of the C-terminus, Ala 263, makes a hydrogen bond to a water molecule that also contacts N1 of the folate, and the pterin 2-amino group of CH<sub>2</sub>H<sub>4</sub>folate. A water-mediated contact with E58 to the folate N-10 atom, observed in the crystal structure of the ternary complex of *E. coli* TS-FdUMP-CH<sub>2</sub>H<sub>4</sub>folate (Hyatt *et al.*, 1997), supports the role of Glu 58 as a general acid catalyst (discussed below).

The glutamate group attached to the folate-PABA ring is expected to be charged at physiological pH, and once the folates are transported into the cell, several glutamate molecules are added on  $\gamma$ -linked amide bonds, in a process called polyglutamylation (reviewed in McGuire and Coward, 1984). This modification increases the folate affinity for TS (Galivan *et al.*, 1976), likely due to the increased electrostatic interactions with the enzyme. Early work using chemical crosslinking of the folyl-analog pteroylheptaglutamate with *L. casei* TS have shown that K48 is in contact with the glutamate tail of the folate (Dunlap *et al.*, 1971), and later confirmed on the structure of

the *E. coli* TS-dUMP-PDDF complex (Montfort *et al.*, 1990). Interestingly, the folyl-glutamate contacts K48 through a hydrogen bond to a water molecule..

The conservative mutation K48R in the bacteriophage T4 decreases  $k_{cat}$  and increases  $K_m$  for CH<sub>2</sub>H<sub>4</sub>folate by two orders of magnitude compared to wild type (LaPat-Polasko *et al.*, 1990), and the K48Q mutant in *E. coli* TS, drops  $k_{cat}$  approximately 400 fold (F. Maley, unpublished).

The TS active site is highly conserved, and mutation of almost any invariant active site residue results in decreased ligand binding and catalysis, suggesting that the enzyme has evolved to maintain a precise and optimal environment for ligand binding and catalysis (Carreras and Santi, 1995, Hyatt, 1997, Montfort and Weichsel, 1997). Unless noted otherwise, all references to amino acid sequence correspond to the *E. coli* TS numbering.

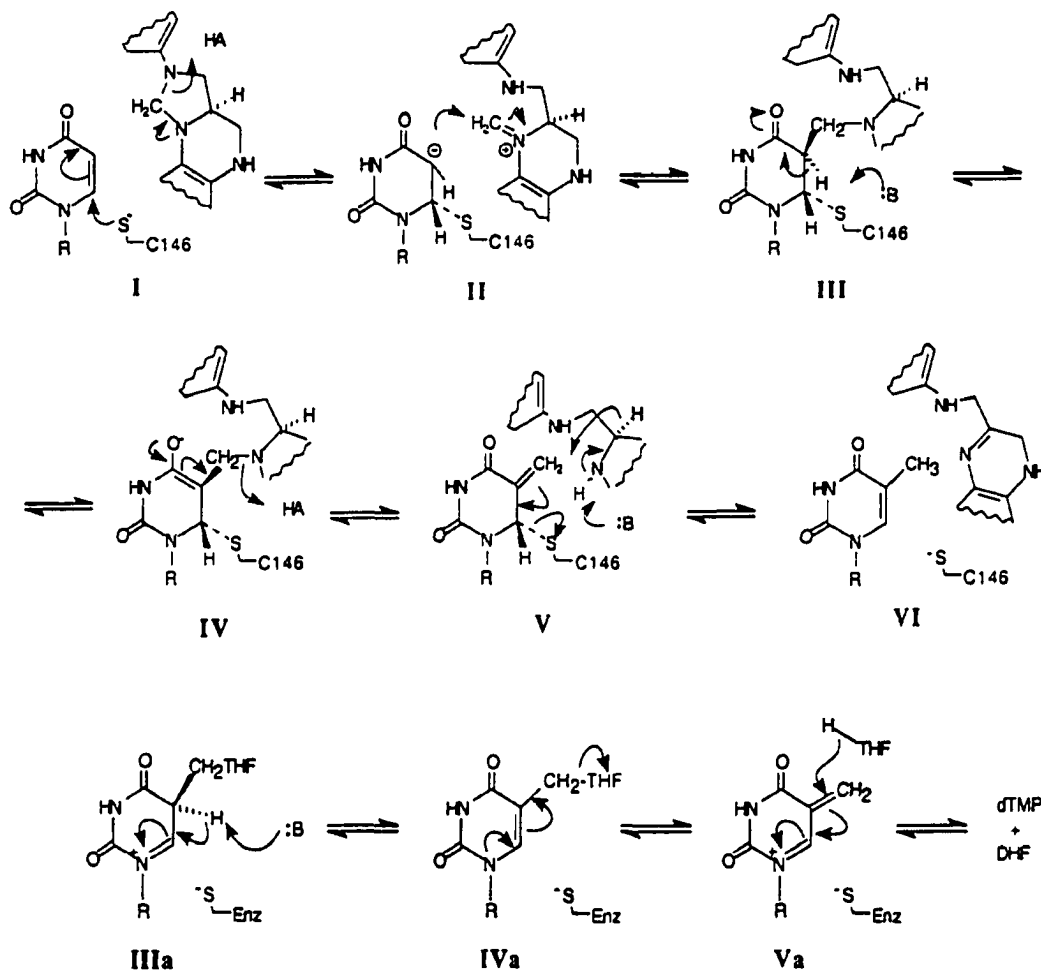
## 1.2. Catalytic Mechanism

TS is a methyltransferase which uses CH<sub>2</sub>H<sub>4</sub>folate both as a reductant and as the carbon source for the conversion of dUMP into dTMP (reviewed in Carreras and Santi, 1995). The results of many studies have led to a detailed model for the mechanism of the TS reaction, shown in Figure 1.6. The binding of the substrate and cofactor is ordered, binding dUMP first (Lorenson *et al.*, 1967). The catalytic reaction is initiated by the Michael addition of Cys 146 to the nucleotide pyrimidine ring at position C6 (intermediate I in Figure 1.6.).

The formation of this covalent adduct has been observed by using the nucleotide-analog and TS-inhibitor FdUMP (Bosch *et al.*, 1958). The fluorine atom at position 5 in the pyrimidine ring does not allow further reaction, trapping the transient covalent bond between C6 and the catalytic Cys 146 (Bellisario *et al.*, 1976). The covalent bond has been also detected by  $^{19}\text{F}$ -NMR spectroscopy (Byrd *et al.*, 1978), in the crystal structure of the ternary complex with FdUMP and  $\text{CH}_2\text{H}_4\text{folate}$  (Matthews *et al.*, 1990a, Hyatt *et al.*, 1997), the ternary complex with dUMP and PDDF (Montfort *et al.*, 1990), and in the binary complex with 5- $\text{NO}_2\text{dUMP}$  (Arendall and Montfort, in preparation).

Once the nucleotide has reacted with the active site cysteine 146, the keto/enol equilibrium is displaced towards formation of the enolate, and C5 becomes a reactive nucleophile (intermediate **II**). Evidence of the formation of the enolate comes from hydrogen-exchange experiments using  $[5\text{-}^3\text{H}]\text{dUMP}$  and the dehalogenation of 5-BrdUMP and IdUMP (Wataya *et al.*, 1973, Pogolotti *et al.*, 1979). The C5 nucleophile reacts with the activated folate, forming a methylene bridge that now links the nucleotide with the cofactor.

That the formation of the methylene bridge between C5 of the pyrimidine ring and C11 from the folate occurs via formation of an N5-iminium cation (intermediate **II**), was initially suggested by studies using model organic molecules (Kallen and Jencks, 1966), and later supported by resonance Raman spectroscopy studies of ternary complex of  $\text{CH}_2\text{H}_4\text{folate}$  and 5- $\text{NO}_2\text{dUMP}$  with *E. coli* TS (Austin *et al.*, 1995).



**Figure 1.6. Proposed detailed TS mechanism with an alternative folate elimination route, from Montfort and Weichsel, (1997). Alternative steps IIIa, IVa and Va are from Hyatt *et al*, (1997). Dashed lines for the thiol adduct indicates instability in this bond.**

However, the authors suggest that the observed signal assigned to the opening of the five-membered ring could be also due to a free-iminium cation product of elimination of formaldehyde to the solution. A more thorough study, in which the lifetime of iminium cations was estimated to be approximately  $2.5 \times 10^{-7}$  s, suggests that the enolate can rapidly attack the imidazolidine ring in a concerted bimolecular mechanism, without need of an unstable iminium ion in solution (Eldin and Jencks, 1995).

There are several proposals as for how the imidazolidine N10 gets protonated, and changes to a  $sp^3$  hybridization. It has been suggested that Glu 58 protonates N10 directly or through a water molecule (Huang and Santi, 1994, Huang and Santi, 1997), or alternatively that the  $pK_a$  of N10 is increased and that it gets protonated from the bulk solvent before ring opening (Arendall and Montfort, in preparation).

In any case, a methylene bridge is formed between C5 of the nucleotide and C11 of the folate, in addition to the covalent bond previously formed between the active site Cys 146 and C6 of the nucleotide, to form intermediate **III** (see Figure 1.6.). Substitution of the substrate by the fluorinated substrate-analog 5-FdUMP, traps the ternary covalent adduct in solution, and the structure of that complex has been determined by X-ray crystallography (Matthews *et al.* , 1990a, Hyatt *et al.* , 1997).

The elimination of  $H_4$ folate has been postulated to occur via a C5-methylene intermediate (Carreras and Santi, 1995). The existence of that intermediate has been proposed by experiments with model compounds (Santi *et al.*, 1974) and its existence was recently demonstrated by detection of a  $\beta$ -mercaptoethanol-adduct of the exocyclic

C5-methylene intermediate (Barrett *et al.*, 1998). Reduction of the exocyclic methylene group (intermediate **V** or **Va**) is accomplished by the folyl-C6-hydride. Only the (6*R*) stereoisomer of CH<sub>2</sub>H<sub>4</sub>folate is in the proper orientation to attack the nucleotide; the (6*S*) stereoisomer is therefore inactive.

Recent data suggests that the strength of the Cys146-C6 dUMP bond is reduced on intermediate **III**. Poor electron density of the thiol bond in the crystal structure of the covalent ternary complex with FdUMP and CH<sub>2</sub>H<sub>4</sub>folate has led to the proposal of an alternative mechanism where the C5 proton is made more acidic due to the strain between the Cys146-C6 bond leading to elimination, as shown in intermediate **IIIa** (Hyatt *et al.* , 1997). The C146 thiolate anion would be stabilized by the nearby Arg 166, thus postulating a catalytic role for that residue, in addition to its postulated role in nucleotide binding, due to the contacts made with the nucleotide phosphate. Further methylene reduction would be accomplished in the absence of the covalent bond to the enzyme, to produce dTMP and H<sub>2</sub>folate (Hyatt *et al.* , 1997).

The reaction mechanism described above proposes the existence of proton donors and acceptors during intermediate steps. Several ordered water molecules in contact with the ligands have been postulated as these general acid catalysts (Huang and Santi, 1997, Hyatt *et al.* , 1997). As an example, mutation of Glu 58 to Gln, which contacts an ordered water molecule in hydrogen bond with folyl-N10 from intermediate **III**, drops  $k_{cat}$  27,000 fold. The  $K_m$  for dUMP was reduced slightly (32 fold) and there was no change in the  $K_m$  for CH<sub>2</sub>H<sub>4</sub>folate (Huang and Santi, 1997), which the authors argue supports this hypothesis. Overall, the positioning of the ordered water molecules and postulated

hydrogen-bond donors and acceptors requires a substantial ligand-induced conformational change in the enzyme, which will be discussed as follows.

### 1.3. Ligand-Induced Conformational Change

An extensive but somewhat subtle ligand-induced conformational change accompanies TS catalysis (Montfort *et al.*, 1990). Binding of nucleotide and cofactor induces structural changes that were first identified using spectroscopic (Danenberg *et al.*, 1974), chromatographic and sedimentation analyses (Lockshin and Danenberg, 1980). Another suggestion that ligands produce a structural change in the enzyme came from the observation that a tight-binding cofactor analog protects TS against proteolytic inactivation with carboxypeptidase (Galivan *et al.*, 1977). Binding of ligands produces a decrease on the Stokes radius of the enzyme (Lockshin and Danenberg, 1980), and further analysis of the unliganded (Hardy *et al.*, 1987) and substrate and cofactor bound TS (Montfort *et al.*, 1990) have provided a structural explanation for this observation.

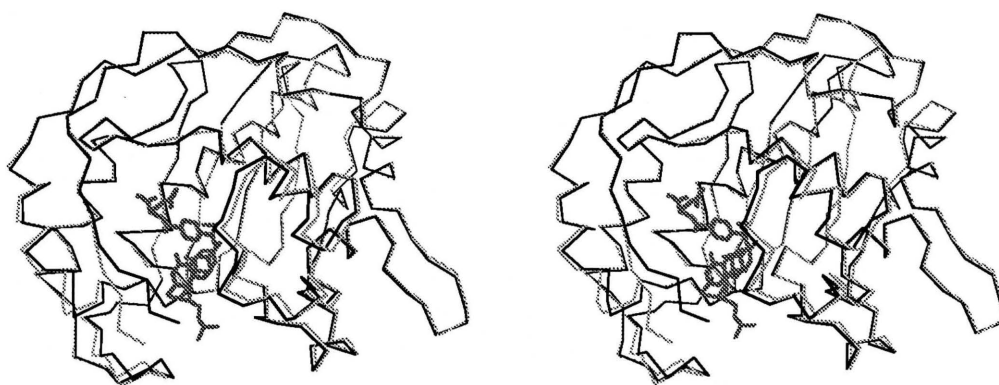
The ligand-induced conformational change in TS is defined by several structural movements. The C-terminus shifts inward towards the active site by about 5 Å, and the loop after helix A moves about 1 Å (Figure 1.3. and 1.7.), to contact the nucleotide phosphate and the C-terminal carboxylate. These shifts lead to a hydrogen bond between the penultimate residue and the pterin ring of CH<sub>2</sub>H<sub>4</sub>folate (Hyatt *et al.*, 1997) or the quinazoline ring the folyl-analog PDDF (Montfort *et al.*, 1990). Also, there are other smaller rearrangements in the enzyme that allow many van der Waals interactions with

the cofactor. In general, the conformational shift has been called "segmental accommodation", since there is a core of residues that are kept rigid while others move towards the ligands (Montfort *et al.* , 1990).

The crystal structures of analogs of TS reaction intermediates, such as the covalent complex of *E. coli* TS-FdUMP and CH<sub>2</sub>H<sub>4</sub>folate (intermediate **III**), show that the conformational change brings many residues in contact with the ligands that would otherwise be too distant to interact. Also, the conformational change probably helps distort the folate from its extended conformation observed in solution (Poe *et al.* , 1979) to the "bent" conformation observed repeatedly on the crystal structure of TS ternary complexes (reviewed in Montfort and Weichsel, 1997)

The mechanism of the ligand-induced conformational change has remained elusive. Many residues located away from the active site are conserved, suggesting they may be involved in the conformational change mechanism. However, TS proteins mutated at these positions have mostly only minor reductions in activity or conformational mobility (Hyatt, 1997, Montfort and Weichsel, 1997). Overall, the ligand-induced conformational change has been suggested to aid TS during catalysis by allowing the substrate and cofactor to bind in their ground-state conformations, to stabilize the higher energy "bent" conformer of CH<sub>2</sub>H<sub>4</sub>folate, to prevent the occurrence of reactive intermediates until they are in position to react with each other, and to prevent the thiol adduct from adopting a stable conformation and falling into a "thermodynamic hole" (Hyatt *et al.* , 1997).





**Figure 1.7. Ligand-induced conformational change in *E. coli* TS.** Stereo view of a superposition of the backbone traces of the structures of the TS binary complex with dUMP, (Roberts and Montfort, in preparation), in grey; and ternary complex with dUMP and PDDF (Montfort *et al*, 1990), in black. The ligands from the ternary complex are included.

### 1.4. Inhibitors and Anticancer Drugs

TS is inhibited by certain chemical analogs of its substrate dUMP or cofactor  $\text{CH}_2\text{H}_4\text{folate}$ . The importance of the nucleotide-analog drugs reside in their unbeaten efficacy against colorectal cancer since their clinical introduction 40 years ago, until the advent of folate-based drugs in the last years (reviewed in Bertino, 1997). The synthesis of 5-fluorouracil was first reported in 1957, based on the observation that rodent tumor cells salvaged uracil more efficiently than normal tissues (Heidelberger *et al.*, 1957). The fluoro-substituted pyrimidine was further metabolized to FdUMP, and proved later to be a potent inhibitor of TS (Bosch *et al.*, 1958). Other analogs such as 5-fluorouridine and 5-fluorodeoxyuridine, which are also converted to FdUMP are also used in the clinic (Bertino, 1997). It was later found that FdUMP required the presence of  $\text{CH}_2\text{H}_4\text{folate}$  to form a covalent ternary complex involving a cysteine residue in the enzyme, shown as intermediate **III**, shown in Figure 1.6 (Santi *et al.*, 1974). In the clinic, 5-fluorouracil is administered in conjunction with DNA-damaging agents, ionizing radiation, antifolates, and more recently topoisomerase inhibitors to stop cell growth by TS inhibition and misincorporation of dUMP into DNA (reviewed in Bertino, 1997).

Folic acid was isolated in 1941 as an antianemic factor and synthesized in 1945. It was observed that it seemed to increase the growth of leukemia cells. Therefore, it was deduced that folic acid analogs might work as cytotoxic agents (reviewed in Kamen, 1997). For example, aminopterin, a 4-amino analog of folic acid, was found to produce temporary remission in children with acute lymphoblastic leukemia. Another folic-acid

analog is methotrexate, which is a competitive inhibitor of DHFR and has been used as an anticancer drug for the last 50 years (reviewed in Kamen, 1997).

Several issues came to light when a folate rather than a pyrimidine based TS-inhibitor strategy was followed to design new anticancer drugs (reviewed in Jackman *et al.*, 1996). Enzymes in the dUMP-biosynthetic pathway such as dCMP deaminase are negatively regulated by dTTP. Therefore, upon TS inhibition by FdUMP, the dUMP pool is increased, alleviating the inhibitory effects of the fluoro-nucleotide. Moreover, since the levels of  $\text{CH}_2\text{H}_4\text{folate}$  in the cell are sub-optimal for stable ternary covalent complex formation with FdUMP, addition of folinic acid (leucovorin) potentiates the biological activity of 5-fluorouracil. The interest in folate-based TS inhibitors was increased by the realization that folate analogs would not require metabolic activation, and should not be catabolized. Later it was found that folate analogs were also processed by addition of multiple  $\gamma$ -linked glutamyl groups to the  $\alpha$ -glutamate attached to the PABA ring, as occurs with the natural cofactor  $\text{CH}_2\text{H}_4\text{folate}$ . This process, called polyglutamylation, is catalyzed by folyl-polyglutamate synthetase, prevents folates from leaking back out of the cell, and increases their affinity for TS binding (McGuire and Coward, 1984, Kamen, 1997).

Numerous folate analogs have been made and screened for biological activity against TS. A folate-based TS inhibitor designed to optimize active site binding is 10-propargyl-5,8,-dideazafolate (PDDF), which inhibits human TS with a  $K_i$  of 10 nM and has antitumor activity in vivo (Jones *et al.*, 1981, Jackman and Calvert, 1995). The quinazoline ring makes hydrogen bonds to the active site on the closed conformation, and

the propargyl group makes hydrophobic contacts with the active site (Montfort *et al.* , 1990). Although promising in toxicological evaluations in mice, it was discontinued from clinical trials due to renal and hepatic toxicity resulting from its low solubility (Jackman and Calvert, 1995).

Many quinazoline-based antifolates have been made and screened for improved solubility (Jackman *et al.* , 1996). Among them, Tomudex has higher solubility, but a lower  $K_i$  towards the free enzyme (670 nM). Tomudex is a substrate of the enzyme folylpolyglutamate synthase *in vitro*, and its biological activity is potentated by polyglutamylolation (Jackman *et al.* , 1996). Tomudex is the only TS-directed antifolate drug approved for clinical use in Europe and it is in clinical trial in the US (Smith *et al.*, 1996).

Another promising antifolate is BW1843U89, a true TS inhibitor which has good therapeutic index, and for which its toxicity can be treated with folate supplementation (Smith *et al.*, 1995). Structurally, BW1843U89 distorts the TS active site and binds in an unexpected conformation that could not be predicted by the current computational docking methods used in drug design (Weichsel and Montfort, 1995, Stout and Stroud, 1996). It is of great interest to determine the nature and extent of the conformational change induced in the mammalian system due to the differences in amino acid sequence compared with the bacterial TS.

### 1.5. Mammalian TSs

The majority of the structural and biochemical studies of TS, which began about 40 years ago (Friedkin, 1959), have been performed with protein isolated from bacterial sources. Since the protein is highly conserved, such studies have enabled the development of potent inhibitors of human TS, some of which have entered the clinic as anticancer drugs (reviewed in Takimoto, 1997). However, an understanding of several mammalian-specific TS functions has not been possible since the mammalian proteins have been difficult to crystallize.

The mammalian TSs have other functions not identified so far in the lower eukaryotic or bacterial homologues. It has been observed that the TS expression is induced by the presence of FU, and it has been proposed that mammalian TS regulates its own expression by a mechanism where the enzyme binds to its own mRNA (Chu and Allegra, 1996). Moreover, it has been observed that TS binds to the mRNA of other proteins like p53 *in vitro* (Chu *et al.*, 1999). This mechanism was postulated to be similar to the iron metabolism regulation mediated by aconitase and the iron responsive element of ferritin mRNA (Hentze *et al.*, 1987).

However, more recent data suggest that protein instability and not mRNA binding provide the mechanism responsible for the mammalian TS posttranslational regulation by nucleotide analogs (Kitchens *et al.*, 1999). Furthermore, TS may be regulated through phosphorylation and compartmentalization between the cytosol and nucleus (Samsonoff

*et al.*, 1997). Once in the nucleus, mammalian TS may be a component of a larger 'replisome' (veer Reddy and Pardee, 1983).

There is also a growing effort to design genetically engineered drug-resistant isoforms of human TS that can be used to transfect bone-marrow stem cells of cancer patients, who then will be able to tolerate higher doses of anticancer drugs during treatment (Fantz *et al.*, 1998, Tong *et al.*, 1998a, Tong *et al.*, 1998b). It is of interest to determine the interactions of TS-directed anticancer drugs bound to the drug target, so that any differences between inhibition of mammalian and bacterial TSs can be uncovered and exploited in the design of new and more specific anticancer, antibiotic or antiparasitic drugs.

## 1.6. Objectives and Approach

This dissertation addresses three questions regarding the function of TS: the ligand-induced conformational change; the ligand-enzyme interactions of a mammalian TS; and the role of conserved active-site residues K48Q and R166. To address the first question, I have determined the crystal structures of two rat TS ternary complexes with dUMP and Tomudex. To study the role of conserved active-site residues on ligand binding and catalysis, I have determined several structures of the *E. coli* TS mutants K48Q and R166Q crystallized in the presence with dUMP, dUMP-PDDF, 5-NO<sub>2</sub>dUMP and FdUMP-CH<sub>2</sub>H<sub>4</sub>folate.

## CHAPTER 2

### MATERIALS AND METHODS

#### 2.1. TS Protein

The wild-type and mutant *E. coli* TS proteins were provided by Dr. Frank Maley (Wadsworth Center, NY) as a frozen-ammonium sulfate precipitate (Maley and Maley, 1988). The protein was prepared for further use by dialysis against 20 mM potassium phosphate buffer (pH 8.0), 0.5 mM EDTA, 5 mM DTT, with at least four buffer exchanges of 500 mL each. Dialyzed TS was stored at 4°C until further use.

The wild-type rat TS protein was provided by Dr. W. Rode (Nencki Institute, Warsaw, Poland) as an ammonium sulfate precipitate (Ciesla *et al.*, 1995). The protein was dialyzed for crystallization experiments against 50 mM MES (pH 7.4), 5 mM DTT and stored at 4°C until use. The concentration of all TS solutions was estimated

spectroscopically by absorbance at 280 nm, using an extinction coefficient of  $0.87 \times 10^5 \text{ M}^{-1} \text{ cm}^{-1}$  for the *E. coli* and rat enzyme (Maley *et al.*, 1995).

## **2.2. Protein Crystallization**

### **2.2.1. *E. coli* TS crystallization**

Crystals of *E. coli* TS binary and ternary complexes were obtained by the hanging drop technique (Giegé and Ducruix, 1992). All crystallization conditions were found by varying pH and precipitant concentration. The well buffer contained 20 mM  $\text{KH}_2\text{PO}_4$ , 0.2 mM EDTA and 5 mM DTT, and ammonium sulfate as the precipitant. The ammonium sulfate concentration was varied from 2.0 to 2.6 M in 0.1-M increments, and the pH was varied from 7.5 to 8.5 in 0.1-pH unit increments. The pH of the well buffers was adjusted after the addition of ammonium sulfate. The protein concentration in the protein stock solution was 10-15 mg/mL, and the ligands were added to a final concentration of 3 mM. The protein stock solution was mixed 1:1 with the well buffer on the glass coverslip and the crystallization trials were done at room temperature (22°C). If crystals were not obtained after one week, the experiment was repeated including a 1-hr equilibration period at 37°C just after setting up the hanging drops. Crystallization plates containing  $\text{CH}_2\text{H}_4\text{folate}$  or folate analogs were kept in the dark during the crystallization process due to their photosensitivity. Usually, the binary complexes of *E. coli* TS with nucleotide produced cubic crystals, except when 5- $\text{NO}_2\text{dUMP}$  was used, which produced hexagonal



rods. The ternary complexes crystallized in a hexagonal or trigonal form. However, the K48Q and R166Q mutants produced only hexagonal rods in the presence of ligands.

### **2.2.2. Rat TS crystallization**

Two crystal forms of the rat TS ternary complex with Tomudex and dUMP were obtained, both in the space group C2. Dr. Larry Hardy (University of Massachusetts), with whom we collaborate, obtained crystal form 1. The protein stock solution he used contained a final protein concentration of 35 mg/ml. The ligands dUMP and Tomudex (a gift from Dr. Tom Boyle, Zeneca Pharmaceuticals) were added to a final concentration of 2.3 mM. Crystals were obtained by the hanging drop method at 4° C after mixing 1:1 with a well buffer solution of 21% PEG 4000 (Fluka), 0.1 M Tris HCl buffer (pH 8.5), and 0.2 M magnesium chloride.

I obtained crystal form 2 at room temperature, using a different precipitant, which also produced monoclinic crystals, but of a different form. The rat TS protein was extensively dialyzed against a buffer containing 5 mM DTT and 20 mM TES, pH 7.4. The protein stock solution contained a final protein concentration of 32 mg/mL, dUMP (6 mM), and Tomudex (6 mM).

The crystallization conditions for crystal form 2 were initially obtained from a sparse matrix grid (Jancarik and Kim, 1991) and further refined. The final conditions were 20% PEG 5000 MME, 200 mM ammonium sulfate, 5 mM DTT and 100 mM MES, pH 6.0 in the well buffer. The hanging drops were prepared by mixing 1:1 the protein

stock with the well buffer described above, and crystals of approximately 0.7 x 0.5 x 0.20 mm were obtained after a few days at room temperature.

### **2.3. Data collection**

All data sets except crystal form 1 of rat TS were collected at room temperature in house using an Enraf Nonius area detector (FAST) and a FR571 rotating Cu K $\alpha$  rotating anode X-ray generator. The data were processed with MADNES (Messerschmidt and Pflugrath, 1987), PROCOR (Kabsch, 1988), and CCP4 (CCP4, 1994). The diffraction data from crystal form 2 of the rat TS-dUMP-Tomudex were measured from two crystals.

For crystal form 1 of the rat TS ternary complex, the data were measured from a single flash-frozen crystal (-170° C in the original buffer) using a MAR imaging plate and Cu K $\alpha$  radiation at the University of Massachusetts by Dr. Hardy. Data were processed with Denzo and Scalepack (Otwinowski, 1993).

### **2.4. Structure determinations**

All the structures reported here were determined using either the molecular replacement or the difference Fourier method. The *E. coli* TS crystal structures were determined by the difference Fourier map method (Chambers and Stroud, 1977), since in all cases the unit cell of the crystals were isomorphous with a previously determined crystal structure. The rat TS crystal structures were solved by the molecular replacement

method since no previous structure of this protein was available. Since crystal form 2 was not isomorphous with crystal form 1, the latter was also solved by the molecular replacement method.

#### **2.4.1. Difference Fourier Maps**

The crystal structures of the complexes of *E. coli* TS with different ligands were determined by the difference Fourier map method. In this method, a difference electron density map is calculated using the coefficients  $|F_1 - F_2|$ , where  $F_1$  corresponds to the observed structure factors for the new crystal and  $F_2$  corresponds to the calculated structure factors for a similar known structure. The phases used in the Fourier synthesis are from the known structure, after removal of the part of the model to be rebuilt, and removal of model bias through use of simulated annealing. The electron-density difference map is displayed on a graphics workstation and used to rebuild the initial model to fit the electron density. The new model is refined against the observed structure factors, and a new rebuilding cycle is performed until the model agrees with the observed electron density, the stereochemistry of the model agrees with the known values, and all of the electron density is accounted for.

### 2.4.2. Molecular Replacement

The structure of the rat TS crystal form 1 was determined by molecular replacement using an *E. coli* TS ternary complex (Montfort *et al.* , 1990) as the starting model, with ligands and solvent molecules removed, using the program X-PLOR (Brunger, 1993).

The protein content of the rat TS crystal was estimated using the Matthews coefficient (Matthews, 1968),

$$V_M = V/M \cdot Z \quad (1)$$

where  $V_M$  corresponds to the Matthews coefficient,  $V$  is the unit cell volume,  $M$  is the molecular weight and  $Z$  is the number of molecules per asymmetric unit. It has been observed that  $V_M$  in protein crystals usually varies from 1.5 to 4.0. Applying formula (1), crystal form 1 corresponds to 1 dimer per asymmetric unit ( $V_M = 3.57$ ) and crystal form 2 to two dimers per asymmetric unit ( $V_M = 2.59$ ).

Once the contents of the unit cells were known, a strategy for the determination of each structure was devised. For crystal form 1, rotation and translation searches were performed to obtain the rotation matrix and the translation vector that would correctly position the initial model, in this case an *E. coli* TS dimer in the closed conformation without ligands, into the monoclinic cell of the rat TS crystal. Omission of the ligands, and the side-chains of those amino acids that differ between the rat and *E. coli* TSs was included as a control to verify the correctness of the molecular replacement solution. For crystal form 2, an additional step was included, the determination of the relationship

between the two dimers within the asymmetric unit. Since this symmetry is not explained by the space-group, it is called non-crystallographic symmetry and it is obtained from a self-rotation search.

Optimal placement of the starting model in the unknown crystal unit cell requires six parameters: rotation along each of the three axes, and translation along these axes. The rotation function search for crystal form 1 using the modified *E. coli* TS model, produced two-fold related peaks corresponding to the two monomers in the dimer. After refining the position of these two-fold related rotations, they had a height five-fold over background. The translation search of the rotated model and rigid body refinement resulted in an initial model with  $R_{\text{crys}} = 0.45$  for either of the two-fold rotations. An electron density map calculated at this point showed clear electron density for the omitted ligands, and some disordered density corresponding to the sequence insertions. These results confirmed the correctness of the molecular replacement solution.

The final model was obtained after multiple rounds of model building with O (Jones *et al.*, 1991), and crystallographic refinement with X-PLOR (Brunger, 1987, Engh and Huber, 1991). Two overall temperature factors were refined for each residue, one each for the main chain and side chain atoms, and non-crystallographic symmetry restraints were applied throughout refinement, due to the low-resolution obtained for this crystal form (3.3 Å).

For crystal form 2, a low-resolution data set was used in the molecular replacement search, using the rat TS model from crystal form 1, except that ligands and residues 20 - 30 were omitted. A 3.0 Å model was obtained by first placing one dimer in

the orientation indicated by the highest peak obtained in the cross-rotation function with the model, followed by a translation search ( $R_{\text{crys}} = 0.44$ ). The second dimer was positioned based on the highest peak obtained from the self-rotation function search, followed by a translation search in the absence of the first dimer, and rigid body refinement of both dimers ( $R_{\text{crys}} = 0.33$ ). Clear electron density was observed for omitted residues 20 - 30 in the initial difference electron density maps, but electron density was absent for the ligands until the structure was better refined. The final model was refined against a second data set collected from 8.0 to 2.6-Å (Table 3.1). Non-crystallographic symmetry was initially restrained during refinement but was eventually left unrestrained, and individual isotropic temperature factors were refined for each atom.

### 2.4.3. Crystallographic refinement

Once the initial phases were estimated, the starting model was subject to a simulated-annealing refinement (Brunger, 1987, Brunger *et al.*, 1990, Brunger and Rice, 1997). Simulated annealing is an optimization technique well suited for multiple-minima problems such as the crystallographic refinement problem. It can explore a larger volume of the conformational space, it reduces the phase-bias towards the initial model in difference-Fourier and molecular replacement methods, and it helps to reduce the crystallographic parameter  $R$ -free more rapidly compared to conventional positional refinement (Brunger and Rice, 1997).  $R$ -free corresponds to the crystallographic residual for a small subset (5-10%) of reflections omitted from the refinement process from the

beginning, and serves as a cross-validated parameter to guide the advancement and correctness of the refinement (Branden and Jones, 1990).

The atomic model was displayed on a graphics workstation along with the newly calculated electron density maps using the program O (Jones, 1978). Manual rebuilding was undertaken as required to fit the electron density, to modify the residues that were mutated and to include and modify the ligands and water molecules that were found in the electron density. The corrected model was used to calculate a new bulk-solvent model, then the model was refined with a least-squares positional algorithm. The thermal B-factors were refined isotropically, one per atom if the data-resolution was higher than 2.8 Å, otherwise two B-factors were refined per residue, one for the main-chain atoms and another for the side-chain. For low-resolution models, non-crystallographic restraints were included, to increase the data/parameters ratio in the crystallographic refinement (Kleywegt and Jones, 1997b), and it reflected in lower R-free values compared to unrestrained refinement.

## **2.5. Comparison of three-dimensional structures**

Comparison of the structural models was done using the least-squares minimization algorithm of Kabsch (Kabsch, 1976) as implemented in the CCP4-LSQKAB (CCP4, 1994) or Uppsala-LKSQMAN (Kleywegt and Jones, 1994) programs. The reported RMS deviations were calculated between all the corresponding C $\alpha$  atoms after an amino acid sequence alignment. The alignment between rat (Ciesla *et al.*, 1995)

and *E. coli* (Belfort *et al.*, 1983) TS was based on the reported amino acid sequences. The structural alignment of bacterial and mammalian TSs was done after a sequence alignment of diverse species (Table 2.1).



**Table 2.1. Sequence alignment of selected TSs.** The numbering is from *E. coli* TS. Conserved residue K48, and invariant residues C146 and R166 are indicated in bold in the sequence alignment.

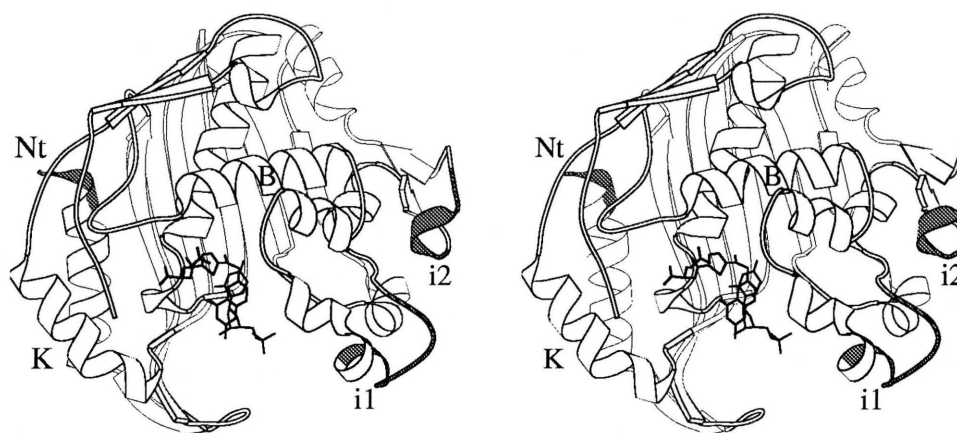
				1	21
Rat	-----MLVEG	SELQSGAQQP	RTE.APQHGE	LQYLRQVEHI	MRCGFKKEDR
Human	MPVAGSELPR	RPLFPAAQER	DAEPRPPHGE	LQYLGQIQHI	LRCGVRKDDR
<b>E. coli</b>	-----	-----	-----M	KQYLELMQKV	LDEGTQKNDR
<i>L. casei</i>	-----	-----	-----MLE	QPYLDLAKKV	LDEGHFKPDR
phage T4	-----	-----	-----M	KQYQDLIKDI	FENGYETDDR
<i>B. subt.</i> A	-----	-----	-----	-----	-----
	22				70
Rat	TGTGTLSVFG	MQARYSLRDE	FPLLTTKRVF	WKGVLEELLW	FIKGS.TNAK
Human	TGTGTLSVFG	MQARYSLRDE	FPLLTTKRVF	WKGVLEELLW	FIKGS.TNAK
<b>E. coli</b>	TGTGTLSIFG	HQMRFNLDG	FPLVTTKRCH	LRSIHELW	FLQGD.TNIA
<i>L. casei</i>	THGTYSIFG	HQMRFDLSKG	FPLLTTKKVP	FGLIKSELLW	FLHGD.TNIR
phage T4	TGTGTIALFG	SKLRWDLTKG	FPAVTTKKLA	WKACIAELIW	FLSGS.TNVN
<i>B. subt.</i> A	-----M	TQFDKQYNSI	IKDIINNGIS	DEEFDVRTKW	DSDGTPAHTL
	71				98
Rat	ELSSKGVRIW	DANGSRDFLD	S.....	.LGF SARQEG	DLGPVYGFQW
Human	ELSSKGVKIW	DANGSRDFLD	S.....	.LGFSTREEG	DLGPVYGFQW
<b>E. coli</b>	YLHENNVTIW	DEWADEN...	.....	.....G	DLGPVYKQW
<i>L. casei</i>	FLLQHRNHIW	DEWAFEKWVK	SDEYHGPDMT	DFGHRSQKDP	EFAAVYHEEM
phage T4	DLR....LIQ	HDSLIIQKTV	WDENYENQAK	DLGYHS...G	ELGPIYKQW
<i>B. subt.</i> A	SVISKQMRFD	NSEVPILTTK	KVAWKTAIKE	LLWIWQLKSN	DVNDLNMGMV
	99				120
Rat	R.....	.HFGADYKDM	DSDY.....	....SGQGVD	QLQKVIDTIK
Human	R.....	.HFGAEYRDM	ESDY.....	....SGQGVD	QLQRVIDTIK
<b>E. coli</b>	R.....	.AWFTP....	.....	....DGRHID	QITTVLNQLK
<i>L. casei</i>	AKFDDRVLHD	DAFAAKYGD	GLVYGSQWRA	WHTSKGDTID	QLGDVIEQIK
phage T4	RDFG.....	.....	.....	.....GVD	QIIEVIDRIK
<i>B. subt.</i> A	HIWDQWKQED	GTIGHAYG..	....FQLGKK	NRSLNGEKVD	QVDYLLHQLK
	121				170
Rat	TNPDDRRIIM	CAWNPKDLPL	MALPPCHALC	QFYVVGELS	CQLYQRSGDM
Human	TNPDDRRIIM	CAWNPRDLPL	MALPPCHALC	QFYVNSELS	CQLYQRSGDM
<b>E. coli</b>	NDPDSRRRIIV	SAWNVGELDK	MALAPCHAFF	QFYVADGKLS	CQLYQRSCDV
<i>L. casei</i>	THPYSRRLIV	SAWNPEDVPT	MALPPCHTLY	QFYVNDGKLS	LQLYQRSADI
phage T4	KLPNDRRQIV	SAWNPALKY	MALPPCHMFY	QFNVRNGYLD	LQWYQRSVDV
<i>B. subt.</i> A	NNPSSRRHIT	MLWNPDELDA	MALTPCVYET	QWYVKHGLKH	LEVRARSNDM
	171				220
Rat	GLGVFPNIAS	YALLTYMIAH	ITGLQPGDFV	HTLGDABIYL	NHIEPLKIQL
Human	GLGVFPNIAS	YALLTYMIAH	ITGLKPGDFI	HTLGDABIYL	NHIEPLKIQL
<b>E. coli</b>	FLGLFPNIAS	YALLVHMAQ	QCDLEVGFV	WTGGDTHLYS	NHMDQTHLQL
<i>L. casei</i>	FLGVFPNIAS	YALLTHLVAH	ECGLEVGFEI	HTFGDAHLYV	NHLDQIKEQL
phage T4	FLGLFPNIAS	YATLVHIVAK	MCNLI PGDLI	FSGGNTHIYM	NHVEQCKEIL
<i>B. subt.</i> A	ALGNPFNVFQ	YNVLQRMIAQ	VTGYELGEYI	FNIGDCHVYT	RHIDNLKIQM
	221				258
Rat	QREPRFPFKL	RI.....	....LRKVET	IDDFKVEDFQ	IEGYNPHPTI
Human	QREPRFPFKL	RI.....	....LRKVEK	IDDFKAEDFQ	IEGYNPHPTI
<b>E. coli</b>	SREPRPLPKL	II.....	....KRKPES	IFDYRFEDFE	IEGYDPHPGI
<i>L. casei</i>	SRTPRPAPT	QL.....	....NPKHD	IFDFDMKDIK	LLNYDPYPAI
phage T4	RREPKECEL	VISGLPYKFR	YLSTKEQLKY	VLKLRPKDFV	LNNYVSHPP
<i>B. subt.</i> A	EREQFEAPEL	WIN.....	....PEVKD	FYDFTIDDFK	LINYKHGDKL
	264				
Rat	KMEMAV				
Human	KMEMAV				
<b>E. coli</b>	KAPVAI				
<i>L. casei</i>	KAPVAV				
phage T4	KGKMAV				
<i>B. subt.</i> A	LFEVAV				

## **CHAPTER 3**

# **CRYSTAL STRUCTURE OF THE WILD-TYPE RAT TS-dUMP-TOMUDEX COMPLEX**

### **3.1. Introduction**

The analyses of the ligand-protein interactions and ligand-induced conformational changes in human TS have been precluded by the poor crystallization properties of the enzyme. To address these problems, I determined the structure of rat TS, a mammalian enzyme 88% identical overall, and 100% identical at the active site with human TS. The complex of rat TS with the substrate dUMP and the anticancer drug Tomudex was obtained in two crystal forms and the structure solved by molecular replacement (see Methods ). These structures are the first views of a mammalian TS active site with ligands bound (Figure 3.1). Unexpectedly, all six active sites from two crystal structures are in the open conformation.



**Figure 3.1. Ribbon diagram of the rat TS-dUMP-Tomudex ternary complex.** The rat amino acid sequence has two insertions compared with the *E. coli* TS. The position of the insertions on the three-dimensional model are shaded in black and labeled "i1" and "i2". Residues 1-20 and 301-306 are disordered and are not shown in the model.

These results reveal important differences between the mechanism of antifolate-based TS inhibition in mammalian and bacterial cells. The 'r' prefix will be used with the rat TS sequence numbering and the 'e' prefix for the *E. coli* enzyme.

## 3.2. Results

### 3.2.1. Structure of rat TS-dUMP-Tomudex complex (crystal form 1)

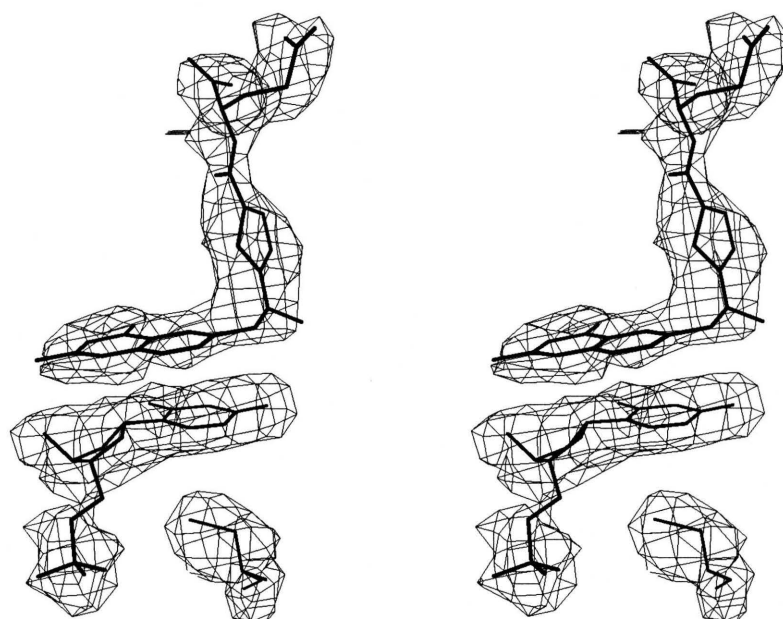
The first crystal form obtained for the rat TS-dUMP-Tomudex complex contained one TS dimer in the asymmetric unit and yielded diffraction data to 3.3-Å resolution. Dr. Larry Hardy of the University of Massachusetts crystallized the protein and measured the diffraction data. The structure was determined by molecular replacement (see Materials and Methods for details and Table 3.1). Despite the low resolution, the ligands were well ordered in the active site and displayed readily discernible conformations in the final electron density map (Figure 3.2).

Surprisingly, the protein was in the open conformation. Main features of the closed conformation are the displacement of the C-terminal loop towards the active site, and multiple smaller movements of the enzyme toward the ligands (Montfort *et al.*, 1990). In the rat TS structure both monomers have a disordered C-terminus, the active site is wider, and the covalent adduct between substrate and the active site cysteine (Cys r189) cannot form.

**Table 3.1 Crystallographic data of the rat TS complexes with dUMP and Tomudex.**

	Crystal form 1	Crystal form 2 (a)	Crystal form 2 (b)
<b>Unit Cell</b>			
Space group	C2	C2	C2
<i>a</i> (Å)	159.6	99.7	99.7
<i>b</i> (Å)	88.5	101.6	101.6
<i>c</i> (Å)	68.8	140.8	140.8
$\beta$ (°)	97.8	101.4	101.4
<b>Data Collection</b>			
Resolution (Å)	15-3.3	20-3.5	8-2.6
Unique reflections	14,261	17,106	35,484
Multiplicity	5.58	2.5	1.8
Completeness (%) <sup>1</sup>	100 / 100	97 / 97	85 / 71
<i>I</i> / $\sigma$ ( <i>I</i> )	18 / 11.3	5.4 / 3.2	10.3 / 2.6
$R_{\text{sym}}^2$	0.01 / 0.14	0.12 / 0.22	0.21 / 0.56
<b>Refinement</b>			
$R_{\text{crys}}^3$ (working set / test set)	0.19 / 0.23		0.16 / 0.22
RMSD, ideal distances (Å)	0.008		0.008
RMSD, ideal bond angles (°)	1.3		1.4
Most favored phi/psi (%)	90		90
<B>, protein atoms (Å <sup>2</sup> )	27		39
<B>, ligand atoms (Å <sup>2</sup> )	35		68
Number solvent mol.	0		161

<sup>1</sup>Total / outer shell.<sup>2</sup> $R_{\text{sym}} = (\sum_h |I_h - \langle I \rangle|) / (\sum_h I_h)$ .<sup>3</sup> $R_{\text{crys}} = (\sum_h |F_{\text{obs}} - F_{\text{calc}}|) / (\sum_h F_{\text{obs}})$ , where the working and test sets are as implemented in X-PLOR (Brunger, 1992a).



**Figure 3.2.** Stereo view of a SA-omit map of the ligands and catalytic cysteine r189, crystal form 1. Monomer "A" of the rat TS-dUMP-Tomudex complex at 3.3 Å.

The rat protein contains three insertions with respect to the *E. coli* TS: an extra 23 amino acids at the N-terminus, and two insertions, one of 12 residues after position r110 and another of 8 residues after residue r139. Only three residues of the N-terminus were visible in the final electron density maps, and both insertions had poorer density compared with the rest of the molecule. Finally, the last seven residues were disordered in all the structures determined in this work (Figure 3.1). To confirm that the rat TS protein used was fully intact, the N-terminus was sequenced in the Biotechnology facility at the University of Arizona, and the weight of the polypeptide was determined by mass spectrometry on the Wadsworth Center (Albany NY). Both techniques indicated that the protein was complete.

Changes in the amino acid sequence between bacterial and mammalian TS are reflected in the structure and function. One interesting example is substitution of the *E. coli* N-terminal methionine by Glu 24 in rat TS. In *E. coli* TS, the N-terminus forms a carbamate with CO<sub>2</sub> in order to make hydrogen bonds to two threonine side chains (Fauman *et al.* , 1994). In the rat TS, the residue corresponding to Met 1 is Glu 24, which side chain forms hydrogen bonds to Thr r69 and Thr r70. Therefore, the bacterial N-terminal carbamate is functionally substituted by a glutamic acid residue.

However, the rat TS ternary complex does not adopt the closed conformation as occurs in the *E. coli* TS-dUMP-Tomudex complex, nor does it form the covalent adduct with dUMP (Rutenber and Stroud, 1996). The more open conformation found in the rat TS complex leads to fewer protein-ligand contacts, as described in a later section.

### 3.2.2. Structure of rat TS-dUMP-Tomudex (crystal form 2)

A second crystal form for rat TS was found using different crystallization conditions, and these crystals diffracted to higher resolution and contained two full TS dimers (~140 kDa) in the asymmetric unit. The second crystal form resulted in a 2.6-Å structure of the rat TS ternary complex with dUMP and Tomudex. The four monomers in crystal form 2 are nearly identical to each other and to the model of crystal form 1. The data were measured in two data sets, however the two crystals were not completely isomorphous, and the data sets were kept separated. The low-resolution data set was used to determine the structure using molecular replacement at 3.5-Å resolution (crystal form 2a, Table 3.1). The second data set was collected from 8 to 2.6 Å, and it was phased by difference Fourier methods to rebuild and refine the fine details of the structure, using the previous low-resolution model as the starting model (crystal form 2b, Table 3.1).

The protein in crystal form 2 is once again in the open conformation, and the N- and C-termini are also disordered. The loop containing Arg r44, which contacts both the dUMP phosphate and the C-terminus in the *E. coli* TS closed conformation structure, is very poorly ordered in crystal form 2, and its position in the model is unreliable. The rat TS insertions are better defined in crystal form 2, although they are still less well ordered than some other portions of the protein, and they display the same conformations as in crystal form 1. The ligands are present in all active sites, but are more poorly ordered than in the first crystal form 1 (Figure 3.3). The electron density for monomers C and D is better than for monomers A and B in the crystal.



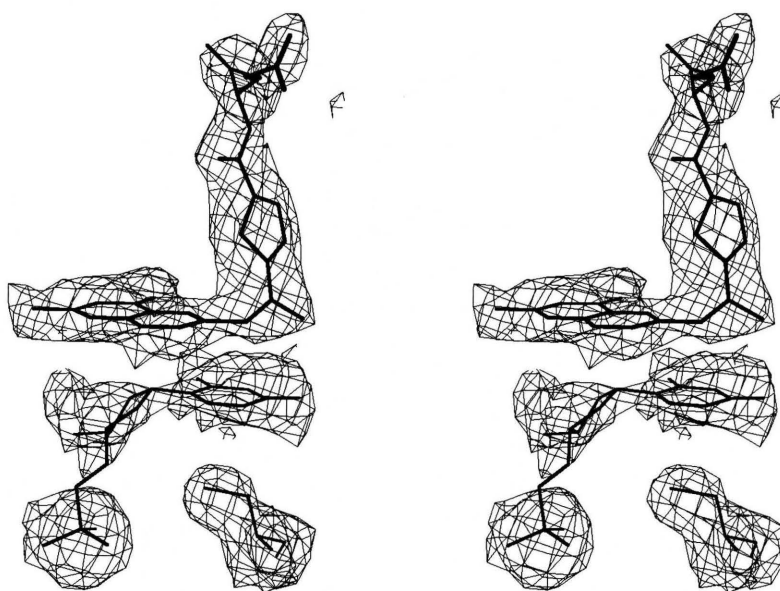
The final models for the four monomers in the second crystal form consist of the same residues as the two monomers in crystal form one: amino acids r21 - r300.

Residues r1 - r20, and r301 - r307 are not visible in any of the six independent monomers resulting from the two crystal structure determinations.

### 3.2.3. Ligand binding in rat TS-dUMP-Tomudex

The ligand arrangements in the *E. coli* TS and rat TS ternary complexes are quite similar except that contacts to the “right” side of the rat active site (as displayed in Figure 3.4), where there are fewer contacts in number, and the hydrogen bonds to the C-terminus, are completely missing. All contacts with dUMP found in the *E. coli* TS complex were found in rat TS except for the covalent adduct with Cys r189, which cannot form when the protein is in the open conformation (Figure 3.4). The Cys r189 sulfur and the pyrimidine ring are in van der Waals contact (3.5 Å in monomer C of crystal form 2), but the residue is poorly aligned for nucleophilic attack. The ligand is unable to shift the ~1.5 Å required for covalent addition without either a shift of the protein to the closed conformation or a loss of several hydrogen bonds to the pyrimidine ring (Arendall and Montfort, unpublished data, and discussed in Montfort and Weichsel, 1997).

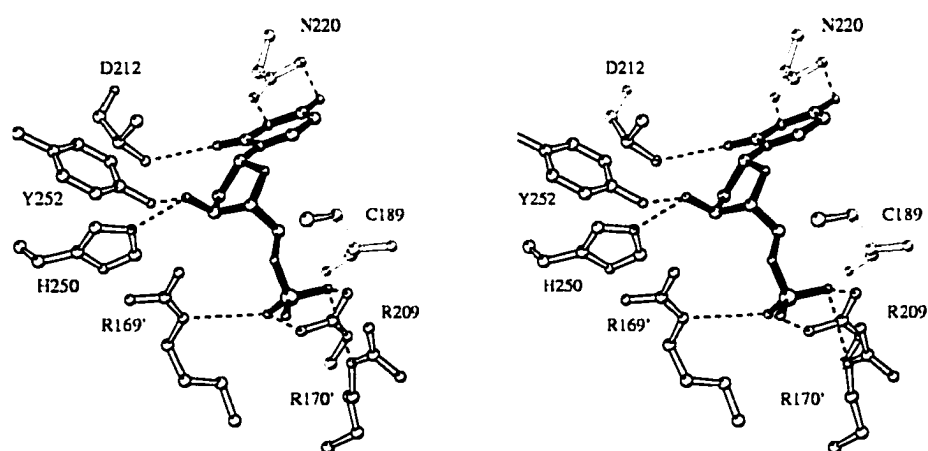
All but one of the nine residues that hydrogen bond to dUMP in the *E. coli* TS closed conformer (Finer-Moore *et al.*, 1990, Montfort *et al.*, 1990) also hydrogen bond in rat TS, except for Arg r44, which is poorly ordered as noted above.



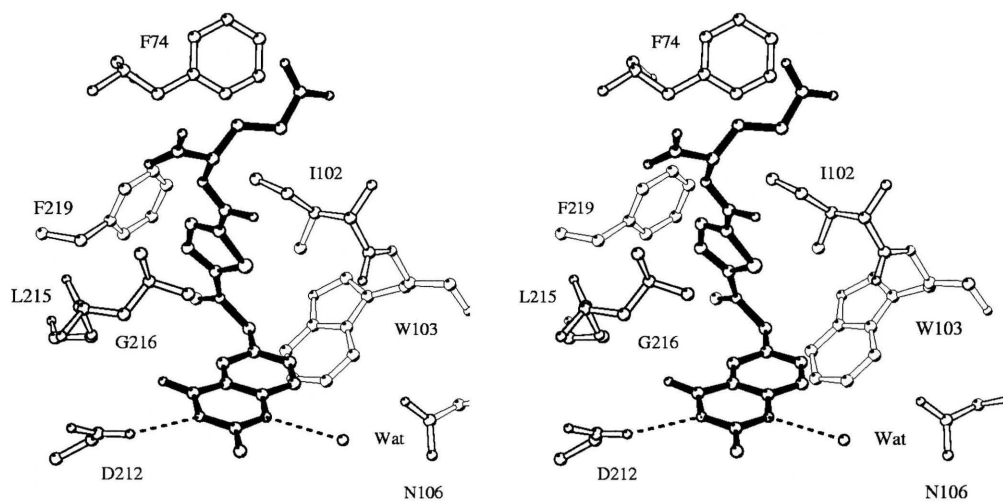
**Figure 3.3. Stereo view of a SA-omit map of the ligands and catalytic cysteine r189, crystal form 2. Monomer "C" of the rat TS-dUMP-Tomudex complex at 2.6 Å.**

However, several of the ordered water molecules found in the *E. coli* TS ternary complexes, including the water linking O4 of the pyrimidine and the carboxylate of Glu 58, were absent in the rat structure. This is probably a consequence of the open conformation, which provides more room for solvent in the active site, allowing for greater solvent disorder.

In contrast to dUMP binding, there were several key differences in the way Tomudex binds to rat and *E. coli* TSs (Figure 3.5). The first differences are due to the open conformation found in rat TS. Van der Waals contacts between Tomudex and Ile r102, Trp r103, Leu r215, Gly r216, and Phe r219 in the rat complex are similar to those found in the *E. coli* complex, although Phe r219 rotates to a new position. The expected hydrogen bond between the side chain of Asp r212 and Tomudex is also present (3.2 Å), but the expected water-mediated hydrogen bond to the C-terminal loop is completely missing, since the loop is disordered, although the water molecule is present in the structure (3.2 Å). Tomudex and dUMP display an extensive van der Waals contact surface in both structures, but the two ligands differ in position by 1 - 1.5 Å in the two structures (Figure 3.6). This difference in position may serve to further dislodge the C-terminus in the rat complex, since the 2-methyl portion of the Tomudex quinazoline overlaps with the expected position for r305 of the C-terminal loop (r301 - r307). Binding may be further altered by two active site differences between the rat and *E. coli* proteins. The first is the change in amino acid sequence of *E. coli* Trp 83 to rat Asn r106 (Figure 3.6).



**Figure 3.4. Stereoview of dUMP binding in rat TS.** Crystal form 1, monomer "A". The ligand is drawn with filled bonds, and residues are labeled with the rat numbering. Ser 210 is omitted for clarity. Arg 169' and 170' are from monomer "B".



**Figure 3.5 Stereoview of Tomudex binding in rat TS.** Crystal form 2, monomer "C". The ligand is drawn with filled bonds, and residues are labeled with the rat numbering.

The tryptophan at this position forms van der Waals contacts with the quinazoline ring of Tomudex in the *E. coli* TS structure, but the asparagine found in the rat protein cannot form these contacts. The loss of this contact may contribute to the inability of Tomudex to induce the closed conformation (Figure 3.6).

The second amino acid change is *E. coli* His 51 to rat Phe r74 at the 'top' of the active site. In the *E. coli* protein, His 51 is on the edge of the active site, but in the rat TS Phe r74 moves further into the active site and rotates to make extensive van der Waals contacts with the inhibitor (Fig. 3.5). Phe r74 also contacts Phe r219, which shifts further into the active site compared to equivalent residue in *E. coli* TS. In this new position, Phe r219 retains extensive van der Waals contacts with the thiophene ring, as noted above, but now contacts Glu r81 on the other side of the active site. For the protein to adopt the closed conformation, Phe r219 would have to move back to a position similar to that found in *E. coli* TS, and Phe r74 would likely have to move out of contact with the inhibitor.

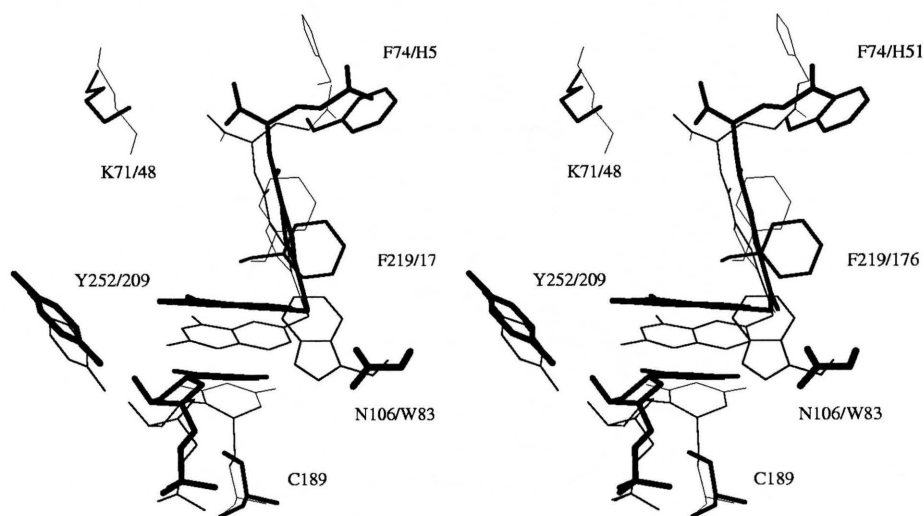
As mentioned before, Tomudex binds less tightly than the parent compound PDDF. Tomudex has a 2-methyl in place of the 2-amino group found in PDDF-quinazoline or  $\text{CH}_2\text{H}_4\text{folate-pterin}$  rings (Figure 1.2), and this change results in the loss of a hydrogen bond to the C-terminus. Also, Tomudex does not have the propargyl group found in PDDF, reducing the hydrophobic contacts of the ligand with the enzyme. Removal of the propargyl group to a quinazoline PDDF-analog reduces the TS inhibitory activity by 200-fold (Jackman *et al.*, 1991). These factors added to the changes in amino

acid sequence discussed below may lead to the observed open conformation in the crystal structures of the rat TS complex with dUMP and Tomudex.

It is important to mention that many folate analog compounds, including PDDF and Tomudex, are substrates for folyl-polyglutamate synthetase (FPGS), and therefore they are polyglutamylated *in vivo*. It has been determined that substitution of the PABA ring by a thiophene group in PDDF analogs increases the affinity of the antifolate for the enzyme folyl-polyglutamate synthetase (reviewed in Jackman *et al.* , 1991). This suggests that although Tomudex has lost affinity at the expense of increased solubility, it may be a better inhibitor by an increased intracellular activation by FPGS. It would be of interest to determine if changes on the chemical structure of the inhibitor such as polyglutamylation may favor the closed conformer. These questions could be addressed by determining the crystal structure of rat TS in ternary complex with dUMP and PDDF, a tighter-binding inhibitor, or the polyglutamylated form of Tomudex.

#### **3.2.4. Comparison of rat TS with the bacterial homologues**

TS is perhaps the most highly conserved enzyme known, with the bacterial and mammalian TSs having ~ 50% sequence identity. It is therefore not surprising that the structures of the bacterial and rat TSs are also nearly identical. The ligand binding arrangement for ligands is nearly identical in rat and *E. coli* TS, as discussed above, and the hydrogen bonding to Thr r69 and Thr r70 near the N-terminus is conserved.



**Figure 3.6. Superposition of the rat and *E. coli* TS-dUMP-Tomudex complexes.** The mammalian model is drawn in thick lines and the bacterial in thin lines. Labeling corresponds to the rat/*E. coli* amino acid sequence.



Two insertions (r111 - r122, and r140 - r147) form extensions of shorter loops found in *E. coli* TS, and have essentially no effect on either the local fold at the point of insertion or the general TS fold.

Insertions after residue r111 occur in both rat and *L. casei* TS, but, interestingly, the rat TS insertion looks nothing like that of *L. casei* TS (Finer-Moore *et al.* , 1994). This suggests that the insertion either does not have a functional role in TS, or that the functional role is species specific. The second rat TS insertion (after residue r140) leads to an alternate conformation for the loop containing the insertion. The residues in this insertion are conserved in mammalian TSs and may play a functional role in higher species, such as providing a surface for binding to other cellular proteins. The conformation of the second insertion is anchored by hydrogen bonds between Tyr r140 and Thr r90, and between Tyr r147 and Asp r142.

The most surprising difference found in the present structures is that the protein conformation of rat TS-dUMP-Tomudex is open, while that of *E. coli* TS-dUMP-Tomudex is closed. The ligand induced conformational change in TS had been previously characterized only in *E. coli* TS, where it was shown to involve the inward shifting of helices and loops on both sides of the active site (Matthews *et al.* , 1990b, Montfort *et al.* , 1990, Hyatt *et al.* , 1997).

The largest shift was in the C-terminal loop, which moves 5-6 Å on forming a ternary complex. Additional inward shifts of 0.3 - 1 Å occur for about half of the remaining amino acids, resulting in an active site volume reduced by nearly 50% (Roberts and Montfort, unpublished data). Superpositioning of the rat ternary complex (monomer

C of the second crystal form) with the first monomer of the *E. coli* TS-dUMP-Tomudex ternary complex yielded an RMS deviation in C $\alpha$  positions of 2.1 Å. This value is greater in magnitude than that for most of the atoms involved in the ligand induced conformational change. Similar values are obtained for superpositioning with an *E. coli* TS-dUMP binary complex (RMSD = 2.0 Å, Roberts and Montfort, unpublished data) and with a *L. casei* TS-phosphate binary complex (RMSD = 1.9 Å, (Finer-Moore *et al.*, 1993)). Thus, the analyses of the conformational change are hampered by species-specific differences among the structures.

Despite this complication, it is clear that the same amino acids that shift on ligand binding to *E. coli* TS also shift in the rat TS ternary complex when compared with the *E. coli* TS ternary complex. For example, the C-terminus is completely disordered in the rTS structure, and residues on both sides of the active site shift outward by ~ 1 Å (after superpositioning of the two structures), including the K-helix and preceding residues (r252 - r265), and the loop containing Tomudex contact residues Ile r102 and Trp r103. Another indication is that the distance across the active site is about 1 Å further in the rat TS structure than in the *E. coli* TS ternary complex structure, but is about the same as that for the *E. coli* and *L. casei* TS binary complex structures. For example, the distance between the C $\alpha$  atoms for residues Tyr r252 and Cys r189, which contact dUMP from opposite sides of the active site (Fig 4a), is 16.5, 16.4, and 16.2 Å for the ‘open’ complexes rat TS-dUMP-Tomudex, *E. coli* TS-dUMP, and *L. casei* TS-phosphate, but is only 15.4 Å for the ‘closed’ complex *E. coli* TS-dUMP-Tomudex.

One difference in the structure determinations of rat and *E. coli* TS is that the former was crystallized in a solution of moderate ionic strength (0.2M salt, 20% PEG), while all of the *E. coli* TS structures have been crystallized in solutions of more than 2 M of salt. Evidence that this difference in ionic strength is not the reason for the difference in the rat TS and *E. coli* TS structures comes from the crystal structure of TS from *Leishmania major*, which is a bifunctional enzyme containing both TS and DHFR activities in a single polypeptide (Knighton *et al.* , 1994). This protein was crystallized in 0.12 M ammonium sulfate and 19% PEG in a ternary complex with FdUMP and PDDF, and displayed a closed conformation nearly identical to that of the high-salt *E. coli* TS-FdUMP-PDDF structure (Knighton *et al.* , 1994). Interestingly, the *Leishmania* protein also has an Asn substituted for Trp at position r106, as described above for rat TS, yet the tighter-binding PDDF induces the closed conformation in this protein despite the loss of contacts with this residue.

### 3.2.5. Comparison of human and rat TS

The only other mammalian TS structure so far reported is that for the unliganded human TS at 3.0 Å resolution (Schiffer *et al.* , 1995). In this structure, none of the mammalian insertions were visible, and the active site cysteine, Cys h180, was found to lie in a non-functional conformation with the catalytic thiol buried, possibly due to disulfide bond formation with Cys h195 during purification. In the rat TS structures of

the present study, two of the three-mammalian insertions are visible, and the active site thiol is in the functional position.

### 3.3. Discussion

The two crystal structures of rat TS in complex with dUMP and Tomudex determined in this study provide the first look at a mammalian TS with ligands bound, and provide the first well-resolved models for understanding the structural basis for mammalian-specific TS functions. The most surprising revelation in these structures is that binding of Tomudex to the rat and *E. coli* proteins differs. Rat TS monomers were determined in six different crystalline environments (two in crystal form 1 and four in crystal form 2), and all six were essentially identical and in the 'open' conformation. In contrast, the ternary complex formed with *E. coli* TS (Rutenber and Stroud, 1996) is in the same closed conformation that all *E. coli* TS inhibitory complexes have exhibited (Montfort and Weichsel, 1997).

While it is possible that the rat TS-dUMP-Tomudex complex may also adopt the closed conformation, I only found the open conformer in six active sites from two independent crystal structures obtained from different crystallization conditions. Further evidence in support of this conclusion comes from kinetic studies: Tomudex inhibition is competitive with CH<sub>2</sub>H<sub>4</sub>folate for the rat protein, indicating free-exchange of both inhibitor and cofactor occurs, whereas inhibitors such as PDDF (Fig. 1.2) exhibit a slow-binding phase where inhibition changes from competitive to a form with an extremely

long off rate (26.5 h for the *L. casei* protein (Pogolotti *et al.*, 1986)). Consistent with the stability of this complex is the *E. coli* TS-dUMP-PDDF structure, which has the closed conformation and a covalent bond between dUMP and Cys 146 (Montfort *et al.*, 1990). Interestingly, the rat TS complexes with more tightly binding inhibitors such as PDDF do not crystallize in same manner as the Tomudex complexes, suggesting such ligands induce the closed conformation in the protein (Sotelo-Mundo and Montfort, unpublished observations).

The open conformation in the present structures may result from the accumulation of several destabilizing factors. First, the inhibitor has fewer contacts with the protein than other inhibitors such as PDDF. The latter compound can hydrogen bond to the C-terminal loop through its 2-amino group, whereas Tomudex has a 2-methyl in this position. PDDF also has a bulkier N10 substituent than Tomudex, which makes several contacts with the protein. Second, the rat protein has lost an amino acid contact to the inhibitor through the change of Trp 83 for Asn r106. Third, the energetic barrier to active site closure may be larger in rat TS than in *E. coli* TS through the species-specific contacts of Phe r74, which contacts Phe r219 and alters its position such that it now sterically opposes active site closure.

It is interesting that, despite the forgoing discussion, binding of Tomudex to rat TS is at least 10-fold tighter than to *E. coli* TS. One possible reason for this is that Phe r74 becomes buried on Tomudex binding to rat TS, but in *E. coli* TS, this residue is a histidine and not involved in binding. A second possible reason is that binding energy is apparently not expended to induce the closed conformation in rat TS, whereas in *E. coli*

TS, some portion of the binding energy must be used to overcome the energetic barrier to closure. A possible design strategy for the improvement of Tomudex would therefore be to further stabilize binding of the inhibitor to the open conformer of rat TS. This could be accomplished by, for example, replacing the N10-methyl with a more bulky substituent, or by adding substituents to the quinazoline rings.

## **CHAPTER 4**

### **CRYSTAL STRUCTURES OF THE K48Q AND R166Q *ESCHERICHIA COLI* TS MUTANTS**

#### **4.1. Introduction**

The amino acid sequence of TS is perhaps one of the most conserved through evolution. The role of individual residues in ligand binding, catalysis and conformational change during catalysis by TS have been studied using single-residue mutations. These studies have helped to uncover many details of the TS catalytic mechanism (Carreras and Santi, 1995, Montfort and Weichsel, 1997), however several key aspects of TS function remain unclear. In this work, I have studied two residues that interact with the TS substrates. Arginine 166 (R166) is one of the four arginines that contact the phosphate group of dUMP (Montfort *et al.* , 1990). Mutagenesis studies have shown that R166 cannot be mutated to lysine or any other residue without nearly complete loss of enzymatic activity (Climie *et al.* , 1990). Lysine 48 (K48) is another invariant residue that contacts the folate glutamate-tail *via* a hydrogen bond with a water molecule (Finer-Moore *et al.* , 1990, Matthews *et al.* , 1990b, Montfort *et al.* , 1990). The amino group of

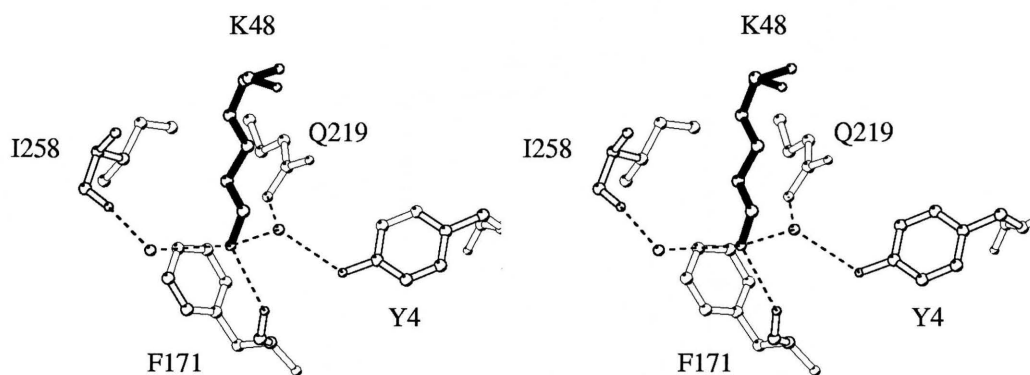
K48 also participates in a hydrogen-bond network with N219, Y4 and the backbone of residues 170 and 171, and contacts the glutamyl-tail of CH<sub>2</sub>H<sub>4</sub>folate via a water molecule, shown in Fig. 4.1 (Hyatt, 1997). The mutation of K48 to glutamine produces a 420-fold decrease in  $k_{cat}$  in the *E. coli* TS (F. Maley, unpublished), although the K48R mutant of the T4-phage TS retains 7% of enzymatic activity (LaPat-Polasko *et al.*, 1990). Thus, two residues that contact substrates at the furthest possible points from the catalytic center are apparently vital for catalysis. The studies described below were designed to uncover the molecular basis for these findings.

#### 4.2. Structure of *E. coli* TS K48Q mutant complexes

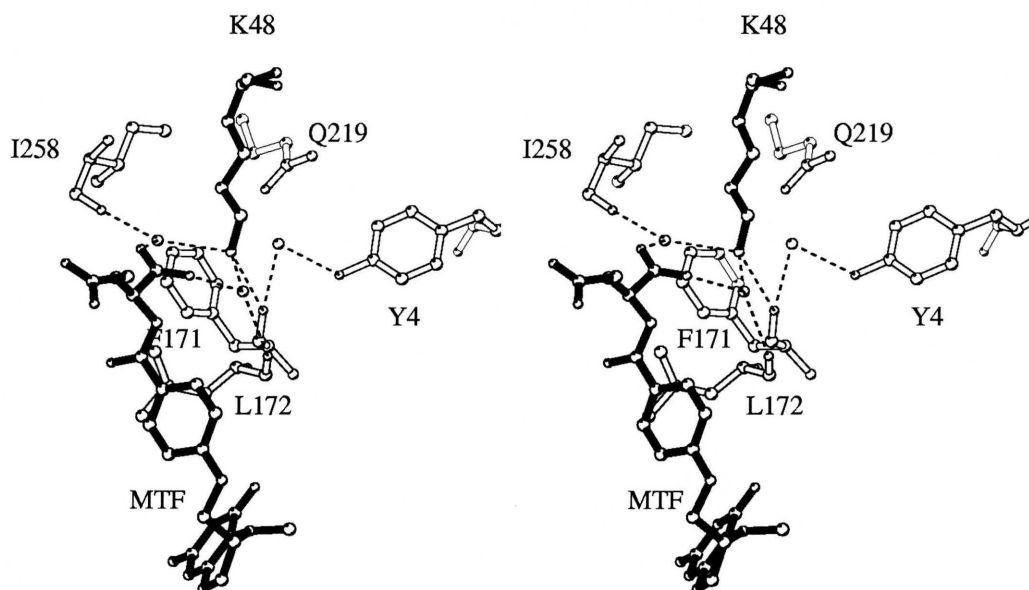
As mentioned above, lysine 48 is a conserved residue that contacts the folyl-glutamate *via* a water molecule. Structurally, K48 is located on the loop that connects  $\beta$ -strand i and  $\alpha$ -helix B (Figure 1.3). In the crystal structure of the *E. coli* TS binary complex with dUMP (Roberts and Montfort, in preparation), K48 makes a hydrogen bond with a water molecule, that as well contacts the hydroxyl group of Y4 and the amide side-chain of Q219 (Figure 4.1.A). The K48 side-chain amino group also makes hydrogen bonds with the main-chain carbonyl of F171 and with I258 *via* a water molecule. Upon formation of a ternary complex, either with CH<sub>2</sub>H<sub>4</sub>folate or PDDF, K48 now contacts the carbonyl group of the glutamate-end of the folate *via* a hydrogen bond with a water molecule (Figure 4.1.B).



A)



B)



**Figure 4.1. Contacts of K48 in the wild type *E. coli* TS binary and ternary complexes.** A) Stereo view of the contacts of K48 in the binary complex with dUMP (from Roberts and Montfort, in preparation). B) Stereo view of the contacts of K48 with the glutamate tail of the folate, in the ternary complex with FdUMP and CH<sub>2</sub>H<sub>4</sub>folate (labeled as MTF). Trigonal crystal form, from Hyatt *et al.*(1997).

This water molecule is present in the binary complex, although the hydrogen bond with K48 is strengthened upon ternary complex formation. Taken together, these contacts form a hydrogen-bond network that involves Y4 from helix A, Q219 from helix K, F171 from the loop between strand iii and helix J and I258 from the C-terminus (Figure 4.1).

One of the possible roles of K48 may be to provide an electrostatic attraction to the negatively charged glutamate-tail of the folate. In support of this hypothesis is the restoration of enzymatic activity of the K48R mutant in the phage T4 TS (LaPat-Polasko *et al.* , 1990), otherwise inactive by substitution of other residues. To gain further understanding of the role of K48 in binding and catalysis, I have determined the crystal structures of several complexes of K48Q.

#### **4.2.1. K48Q-dUMP Binary Complex**

TS undergoes an extensive ligand-induced conformational change, and has been more thoroughly studied in the *E. coli* enzyme than in other species. I approached this issue by determining mutant structures in both the open and closed conformations. The goal of the binary complex structure determination was to determine the effect of the mutation on nucleotide binding and conformation. Pavel Strop, a former member of our group, obtained cubic crystals of this complex.

**Table 4.1 Crystallographic data for the *E. coli* TS K48Q mutant structures.**

	K48Q- dUMP*	K48Q- 5NO <sub>2</sub> dUMP	Apo-K48Q with dUMP- PDDF	Apo-K48Q with FdUMP- CH <sub>2</sub> H <sub>4</sub> folate*
Unit Cell				
Space group	I2 <sub>1</sub> 3	P6 <sub>3</sub>	P6 <sub>3</sub>	P6 <sub>3</sub>
<i>a</i> (Å)	132.9	127.4	127.4	127.7
<i>c</i> (Å)		67.8	68.3	68.4
Data Collection				
Resolution (Å)	15-2.2	35-2.6	35-2.80	36-2.75
Unique reflections	18,173	17,435	13,874	15,795
Multiplicity	4.4	4.8	3.2	4.7
Completeness (%) <sup>1</sup>	91.2	89.5	87.1	84.4
<i>I</i> / $\sigma$ ( <i>I</i> )	-	7.2/1.7	5.6/1.2	8.8/1.9
<i>R</i> <sub>sym</sub> <sup>2</sup>	0.09	0.09/0.40	0.12/0.51	0.08/0.39
Refinement				
<i>R</i> <sub>crys</sub> <sup>3</sup> (working set / test set)	0.18/0.21	0.18/0.22	0.18/0.24	0.18/0.22
RMSD, ideal distances (Å)	0.005	0.007	0.008	0.009
RMSD, ideal bond angles (°)	1.3	1.3	1.4	1.4
Most favored phi/psi (%)	92	90	88	89
< <i>B</i> >, protein atoms (Å <sup>2</sup> )	32	30	30	33
< <i>B</i> >, ligand atoms (Å <sup>2</sup> )	49	23	-	-

<sup>1</sup>Total / outer shell.<sup>2</sup> $R_{\text{sym}} = (\sum_h |I_h - \langle I \rangle|) / (\sum_h I_h)$ .<sup>3</sup> $R_{\text{crys}} = (\sum_h |F_{\text{obs}} - F_{\text{calc}}|) / (\sum_h F_{\text{obs}})$ , where the working and test sets are as implemented in X-PLOR (Brunger, 1992a).

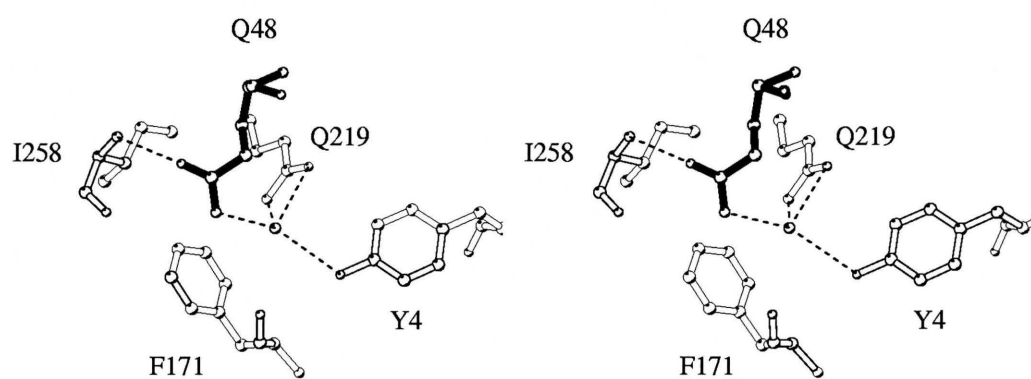
\* Crystallization and data collection by Pavel Strop.

He collected crystallographic data to 2.2 Å resolution, which I used for the structure determination by difference Fourier methods. Simulated annealing and SA-omit maps were used during the molecular rebuilding to remove the phase bias from the initial model.

The crystals of the mutant binary complex belong to the cubic spacegroup  $I2_13$ , which contains one monomer per asymmetric unit. Omit maps showed unambiguous density for the mutated residue and for the ligand dUMP, allowing glutamine at position 48 and the nucleotide to be placed in the structural model early on during refinement. As mentioned, SA-refinement and SA-omitted maps were used during rebuilding to remove phase-bias, and the thermal B-factors were refined individually.

The final model has good stereochemistry, with 92% of the residues in the core region of the Ramachandran plot. The R-factors for this model are  $R_{\text{free}} = 0.21$  and  $R_{\text{cryst}} = 0.17$ . After a three-dimensional superposition with the binary complex of wild type *E. coli* TS with dUMP, the RMS deviation of the C $\alpha$  backbone between the models is 0.15 Å. This RMSD value is similar to the observed for other non-active site mutants in *E. coli* TS when compared with the corresponding wild-type structure (Hyatt, 1997), and indicates that the structures are nearly identical.

The direct effect of the mutation is on the hydrogen bond network between K48 and Y4, F171, Q219 and I258. In the wild type binary complex K48 contacts the main chain of Ile 258 through a water molecule, while Y4 and Q219 are also hydrogen-bonded through a water molecule (Figure 4.1).

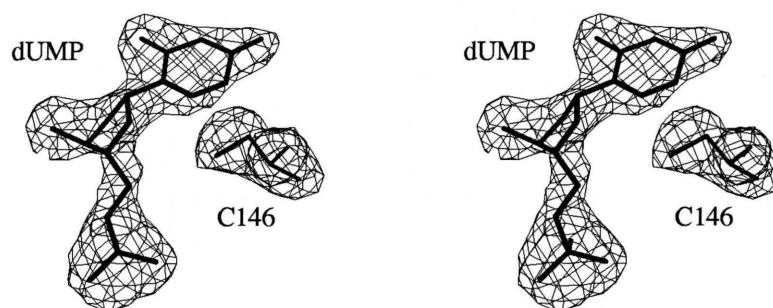


**Figure 4.2. Contacts of Q48 on the mutant binary complex.** From this work.

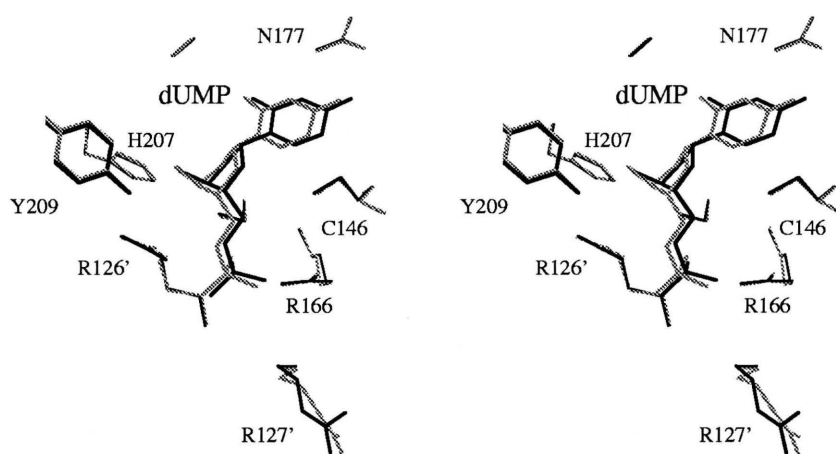
In the K48Q mutant binary complex, the hydrogen bonding in the K48Q mutant is similar to wild type, although contact with the F171 carbonyl is lost (Figure 4.2). The side-chain carbonyl of Q48 now contacts the main chain amide of I258, and Y4 and Q219 are hydrogen-bonded by a water molecule in contact with the amino group of the Q48 side-chain.

The loss of the contact with F171 does not produce any significant shifts in the connecting secondary structure elements. An RMSD calculation of the corresponding C $\alpha$ 's of individual helices and strands of the wild-type and mutant binary structures shows that helix J and strand iii have an RMS deviation of 0.12 Å, the same as for the whole protein..

The electron density for dUMP was similar to the wild type (Figure 4.3) and refined to an average B-factor of 49 Å<sup>2</sup>, which is identical to the average thermal factor of dUMP in the wild type structure. However, dUMP is displaced by approximately 0.5 Å compared to its position in the wild type structure (Figure 4.4). The interaction with N177, which determines the identity of the nucleotide bound to TS (Liu and Santi, 1992) is maintained by the hydrogen bonds to the pyrimidine ring atoms O4 and N3 (Roberts and Montfort, in preparation). The ribose hydroxyl also makes contacts with two conserved residues, H207 and Y209. Although the hydrogen bonds are conserved, the distance between the dUMP-3'OH and the Y209-hydroxyl is increased by 0.2 Å. It is not clear if these shifts are significant, due to the higher mobility of dUMP in the wild type binary complex.



**Figure 4.3. SA-omit map of the active site in the K48Q-dUMP mutant complex.** The substrate dUMP and the catalytic Cys 146 were omitted during the electron density map calculation. The map was contoured at  $3\sigma$ .



**Figure 4.4. Superposition of the wild type and K48Q mutant binary complexes.** The mutant complex is drawn with solid lines and the wild type with grey



which makes three hydrogen bonds with the nucleotide phosphate. Mutation of R126' to glutamine greatly reduces nucleotide binding, and upon tertiary complex formation the mutant binds dUMP in a different conformation (Strop *et al.* , 1997).

In the wild type TS-dUMP binary complex, R166 contacts the nucleotide-phosphate by a weak hydrogen bond, but in the K48Q-dUMP binary complex R166 now makes two strong hydrogen bonds to the nucleotide-phosphate, and now R126' makes only one instead of two hydrogen bonds to the nucleotide'. Apparently the changes in hydrogen bonding on the phosphate binding site accounts for the displacement in the ribose and pyrimidine rings of dUMP on the mutant binary complex. However, it is not clear how or if these differences are due to the mutation to Q48, which is located in the folate-binding site.

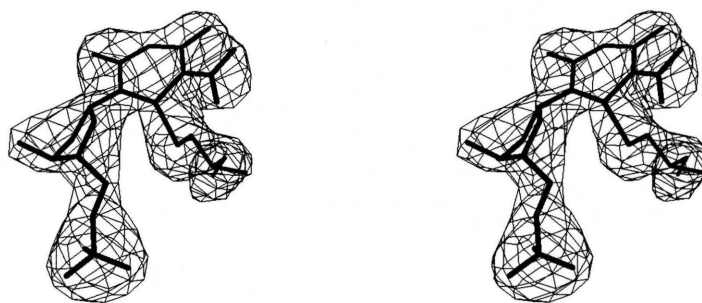
#### **4.2.2. K48Q-5NO<sub>2</sub>dUMP Binary Complex**

I determined the crystal structure of K48Q TS mutant in binary complex with 5-NO<sub>2</sub>dUMP, to determine whether the mutation increased the barrier to active site closure and covalent adduct formation in the absence of folate. The substrate-analog 5-NO<sub>2</sub>dUMP is a competitive inhibitor of TS with a  $K_i = 27$  nM (Wataya *et al.*, 1980), that forms a covalent adduct with TS. The recently determined crystal structure of the binary complex of *E. coli* TS with 5-NO<sub>2</sub>dUMP represents the first closed complex obtained in the absence of folates (Arendall and Montfort, in preparation).

The crystals obtained for the *E. coli* TS K48Q-5-NO<sub>2</sub>dUMP binary complex belonged to the hexagonal space group P6<sub>3</sub>. The asymmetric unit contained a dimer, in contrast to the wild-type equivalent structure, which crystallized in the trigonal space group P321, with one monomer per asymmetric unit (Arendall and Montfort, in preparation). Therefore the hexagonal crystal form allowed the observation of two independent active sites.

The structure was determined by difference Fourier methods. The starting model was the wild-type complex of *E. coli* TS-dUMP-PDDF in the P6<sub>3</sub> space group, without ligands. The difference electron density maps showed the positive and negative electron density corresponding to the change in side chain at position 48, and strong positive electron density in both active sites corresponding to the omitted ligands. Furthermore, the electron density was continuous between the catalytic residue C146 and the density corresponding to the nucleotide (Figure 4.5). Additional minor peaks were found and the model was rebuilt accordingly.

It is important to mention that this complex contains a mixed-disulfide modification at C50 with  $\beta$ -mercaptoethanol. The diffraction experiment was repeated twice using crystals grown with enzyme treated with 50 mM DTT for one hour at room temperature and then extensively dialyzed against fresh buffer containing 50 mM DTT. Nonetheless some residual density remained at C50, although not as strong as in the first crystal form. It is possible that K48 reduces the accessibility of C50 in the wild type enzyme, and upon mutation the sulfhydryl becomes more exposed and reactive.

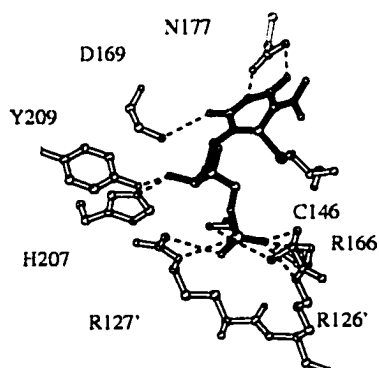
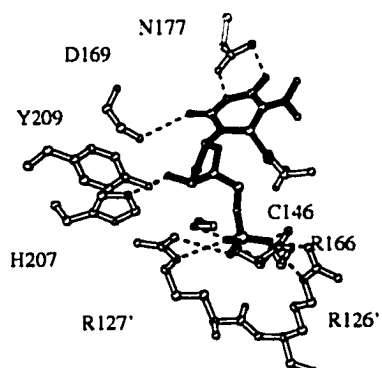


**Figure 4.5. SA-omit map for the K48Q mutant binary complex with 5-NO<sub>2</sub>-dUMP.** Both nucleotide and C146 were omitted in the map calculation. The substrate analog is covalently linked to Cys 146.

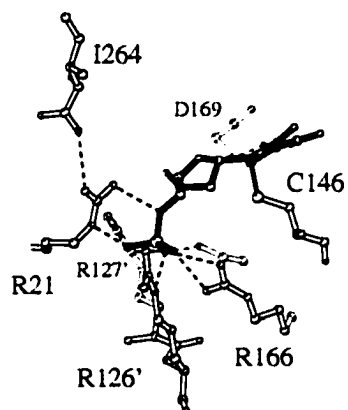
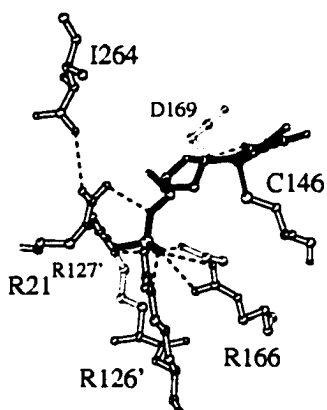
Further rebuilding, and positional and thermal B-factor refinement produced a model similar to the corresponding wild-type complex. The covalent bond between the nucleotide and the catalytic residue C146 was restrained and refined to a distance of 1.9 Å in both mutant and wild type structures.

Comparisons of the refined model with the wild-type binary complex show both structures to be quite similar. The nucleotide is well ordered in both active sites, with an average B-factor of 21 and 25 for the two active sites of the K48Q binary complex, compared to 25 on the trigonal wild type structure (Arendall and Montfort, in preparation). In the K48Q mutant structure, the C-terminal loop has an RMSD of 0.48Å compared with the wild-type 5-NO<sub>2</sub>dUMP binary complex, however in both structures the protein adopts a closed conformation. In the K48Q-5-NO<sub>2</sub>dUMP complex, arginine 21 makes hydrogen bond with the C-terminal carbonyl and the nucleotide-phosphate in both active sites (Figure 4.6.B). As for the hydrogen bonding of the ligand, the mutant structure is missing two water molecules that form hydrogen bonds between the 5-nitro group and to E58 and H147. All the other contacts between 5-NO<sub>2</sub>dUMP and the active site are present (Figure 4.6). In conclusion, K48Q does not interfere with the nucleophilic attack of the catalytic cysteine-146.

A)



B)



**Figure 4.6. Stereo views of the nucleotide binding site in the binary complex of the K48Q TS mutant with 5-NO<sub>2</sub>dUMP.** A) The nucleotide is covalently bound to the enzyme through C146. B) The binary complex is in the closed conformation, where R21 contacts the nucleotide phosphate and anchors the C-terminus (I264) in the closed position. N177, H207 and Y209 were omitted in this view for clarity.

Neither does it prevent the ligand-induced conformational change from occurring. It is not clear whether the disturbance of the hydrogen-bond network with Y4H, Q219 and I258, is responsible for the marginal structural differences between the wild type and mutant structures. However, the results indicate that mutation of K48 does not preclude nucleotide binding or the formation of the covalent adduct between C146 and the substrate.

#### **4.2.3. Apo-K48Q crystallized in the presence of dUMP-PDDF and FdUMP-CH<sub>2</sub>H<sub>4</sub>folate**

The structure determination of ternary complexes with the K48Q mutant were undertaken to investigate the effect of loss of the positive charge in the vicinity of the folate-binding site. Two molecules were used, the enzyme cofactor CH<sub>2</sub>H<sub>4</sub>folate with the substrate-analog FdUMP and the folyl-based inhibitor PDDF with dUMP.

In wild type TS, the ternary complex with CH<sub>2</sub>H<sub>4</sub>folate and FdUMP is thought to be closely analogous to the covalently bound intermediate **III** of the catalytic mechanism (see Figure 1.6) (Matthews *et al.* , 1990a, Hyatt *et al.* , 1997). PDDF is a TS-inhibitor with  $K_d = 10$  nM that induces the closed conformation and the covalent adduct with C146. However, since it lacks the imidazolidine ring it cannot react with the nucleotide (Montfort *et al.* , 1990).

Therefore the crystal structure of both ternary complexes should provide information about the effect of the mutation on folate binding and conformational change.

Crystallization and data collection of the K48Q mutant with FdUMP-CH<sub>2</sub>H<sub>4</sub>folate was done by Pavel Strop. I obtained crystals of the mutant in the presence of dUMP and PDDF, and measured crystallographic data. In both cases, the difference Fourier maps contained only fragmented electron density for the ligands. After SA-refinement and calculation of SA-omit maps it was clear that the crystals contain either no ligands or badly disordered ligand in the active site, except for the position of the nucleotide phosphate, which contained a well ordered sulfate ion from the crystallization buffer. Further rebuilding and crystallographic refinement, which included simulated-annealing, and positional and thermal B-factor refinement with a maximum likelihood target, reduced the noise level to a few peaks modeled as ordered water molecules.

The backbone conformations of both apo-TS were open. The C-terminal loop, which moves towards the ligands in the closed conformation, and the loop between helix A and strand i were poorly ordered. Arginine 21, which makes a hydrogen bond with the C-terminus carbonyl and anchors the closed conformation was also disordered in both K48Q unliganded structures. Y209, which contacts the dUMP-3'-hydroxyl in the wild type binary and ternary complexes was also further away from the active site.

#### **4.2.4. Discussion**

The structural analyses of the K48Q mutant complexes confirm the critical role of K48 for folate binding. The crystallographic results presented here show that the early steps in the reaction mechanism, namely nucleotide binding, and covalent adduct

formation are not impaired by the mutation. Also, the K48Q mutation does not interfere with the folate-independent active site closure induced by 5-NO<sub>2</sub>dUMP. However, the mutation does increase  $K_m$  for dUMP to 32  $\mu$ M, (6 fold increase compared to wild type) and  $K_m$  for CH<sub>2</sub>H<sub>4</sub>folate to 310  $\mu$ M, (12 fold increase). Preliminary results on measuring the interaction between folates and nucleotides using CD spectroscopy (Galivan *et al.*, 1975) indicate that 5-NO<sub>2</sub>dUMP produces changes in ellipticity when bound to K48Q similar to those produced when bound to the wild type enzyme. The interaction between BW1843U89 and 5-NO<sub>2</sub>dUMP with K48Q is nearly identical to wild type, reflected in an ellipticity minimum at 285 nm and a maximum at 324 nm (Dr. Frank Maley, personal communication). However, the determination of the dissociation constant of CH<sub>2</sub>H<sub>4</sub>folate or folate analogs is still needed to test the importance of K48 for TS binding of glutamylated folates.

The fact that BW1843U89 binds to K48Q similarly as to wildtype supports the hypothesis that K48 is required only for binding the cofactor glutamate. BW1843U89 is unique among the folyl-based TS inhibitors in that its glutamate tail lies in an alternate position away from K48, as observed in a crystal structure, due to a ligand-induced distortion of the active site (Weichsel and Montfort, 1995). The crystallization of the complex between K48Q, a nucleotide and BW1843U89 is underway to directly test this question.

The importance of water mediated contacts for ligand binding is not unique for TS. Other enzymes, such as tyrosine hydroxylase, also use water-mediated contacts to contact the ligands. Tyrosine hydroxylase catalyzes the hydroxylation of tyrosine to form



L-DOPA, using tetrahydrobiopterin, oxygen and a non-heme iron as substrates (reviewed in Fitzpatrick, 1998). Mutation of E332, a conserved residue in tyrosine and phenylalanine hydroxylases, to alanine, increases  $V_{\max}$  for L-DOPA 125-fold, with marginal increases in  $K_m$  for substrates (Colette Daubner and Fitzpatrick, 1999).

The crystal structure of human phenylalanine hydroxylase shows that the equivalent residue to E332 makes a hydrogen bond to iron through a water molecule (Erlandsen *et al.*, 1997). Interestingly, in the crystal structure of tyrosine hydroxylase complexed with the cofactor-analog dihydrobiopterin, E332 is close and almost in hydrogen-bond distance to the active site iron (Goodwill *et al.*, 1998). The authors suggest that E332 does not participate in the first step of the reaction, which is tetrahydropterin oxidation, but is important for the productive binding of the intermediate peroxypterin, just before L-DOPA formation. The E332A mutant has a marginal increase in  $K_m$  but has incremented 125-fold the  $V_{\max}$  for L-DOPA (Goodwill *et al.*, 1998). It is remarkable that both E332 in tyrosine hydroxylase and K48 reach the ligands through water molecules, which suggests that contacts with ligands through water molecules allows for conformational change during catalysis.

Returning to the analysis of the K48Q mutant, the apparent loss of folate binding seems to reduce the formation of the closed conformer in the presence of nucleotide and cofactor, leading to a severe drop in the catalytic activity.

One way to study ternary complex formation in K48Q would be to determine the crystal structure of a ternary complex where the positive charge contributed by K48 would not be required for binding, as it occurs with folate-analog BW1843U89.

Determination of such a structure is critical for determining the role of K48 in complex formation. Crystallization experiments using BW1843U89 and dUMP and 5-NO<sub>2</sub>dUMP are underway to directly address these questions.

### 4.3. Structures of the *E. coli* TS R166Q mutant complexes

R166 was the second residue investigated. R166 is one of four arginines that contact the phosphate group of the nucleotide in the active site, and its mutation to glutamine reduces  $k_{\text{cat}}$  3,400-fold (F. Maley, unpublished). The role of the other three arginines at the active site has been studied by kinetic and structural characterization of single-site mutants. R126', which comes from the opposite TS monomer, makes only one hydrogen bond to the nucleotide phosphate, however, its mutation to glutamate reduces  $K_m$  for dUMP 600 fold, due to steric conflict between the phosphate group and the new position of E126' (Strop *et al.* , 1997). In contrast, Arg 127' (R179' in *L. casei* TS) has been mutagenized to Thr, Ala, Lys and Glu without significant loss of enzymatic activity (Climie *et al.* , 1990, Michaels *et al.* , 1990, Santi *et al.* , 1990). Arg 21, located in a flexible loop between helix A and  $\beta$ -sheet i, is also sensitive to mutation (Michaels *et al.* , 1990).

R166 is in van der Waals contact with the thiol group of Cys 146, and has been postulated to have a catalytic role by stabilizing the thiolate anion, which may be required for the nucleophilic attack of dUMP to initiate catalysis, and for stabilizing certain intermediates during catalysis. For example, R166 may be responsible for instability in

the covalent adduct between C146 and dUMP, thereby preventing the enzyme from becoming trapped in a 'thermodynamic hole' (Hyatt *et al.* , 1997). It also forms several hydrogen bonds within the protein, and with the nucleotide phosphate. To investigate the role of R166 in ligand binding and catalysis, I have determined the crystal structures of the R166Q *E. coli* TS mutant from crystals grown with dUMP and in the presence of dUMP-PDDF and FdUMP-CH<sub>2</sub>H<sub>4</sub>folate.

**Table 4.2 Crystallographic data for the *E. coli* TS R166Q mutant structures.**

	Apo-R166Q with dUMP	R166Q-PDDF- SO <sub>4</sub>	R166Q- CH <sub>2</sub> H <sub>4</sub> folate- SO <sub>4</sub> *
Unit Cell			
Space group	I2 <sub>1</sub> 3	P6 <sub>3</sub>	P6 <sub>3</sub>
<i>a</i> (Å)	132.9	127.1	127.4
<i>c</i> (Å)		67.7	67.9
Data Collection			
Resolution (Å)	34-2.94	18-2.2	35-2.6
Unique reflections	8,138	31,445	19,519
Multiplicity	7.0	3.1	4.9
Completeness (%) <sup>1</sup>	96.1	99.2	100
I/σ(I)	7.4	6.7	7.2
R <sub>sym</sub> <sup>2</sup>	0.09/0.38	0.09/0.38	0.09/0.39
Refinement			
R <sub>crys</sub> <sup>3</sup> (working set / test set)	0.17/0.21	0.19/0.23	0.19/0.25
RMSD, ideal distances (Å)	0.008	0.007	0.012
RMSD, ideal bond angles (°)	1.4	1.4	1.8
Most favored phi/psi (%)	88	92	88
<B>, protein atoms (Å <sup>2</sup> )	39	26	20
<B>, ligand atoms (Å <sup>2</sup> )	-	68	n.d.

<sup>1</sup>Total / outer shell.

<sup>2</sup>R<sub>sym</sub> = (Σ<sub>h</sub>|I<sub>h</sub> - <I>|) / (Σ<sub>h</sub>I<sub>h</sub>).

<sup>3</sup>R<sub>crys</sub> = (Σ<sub>h</sub>|F<sub>obs</sub> - F<sub>calc</sub>|) / (Σ<sub>h</sub>F<sub>obs</sub>), where the working and test sets are as implemented in X-PLOR (Brunger, 1992a).

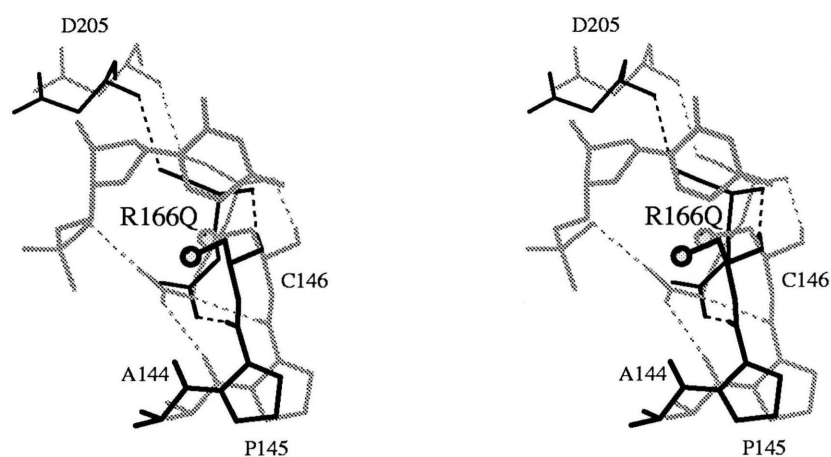
\* The ligands in this structure were not refined due to poor electron density.

#### 4.3.1. Apo-R166Q crystallized with dUMP

I undertook the co-crystallization of R166 with dUMP to determine the effect of the mutation on nucleotide binding. I obtained cubic crystals that diffracted to 2.94 Å which were isomorphous to the wild-type complex, and I determined the structure directly from the difference electron density maps.

The main feature of this structure is the lack of ligands, not even sulfate bound by the remaining three active-site arginines (R21, R127' and R166). This result is similar to the structure of the *E. coli* TS R126' mutant crystallized with dUMP, in which the nucleotide was absent from the active site (Strop *et al.* , 1997). In contrast, the electron density for the mutation was clear and rebuilt accordingly. The shortening of the side chain, from arginine to glutamine, does not permit any interaction with the phosphate-binding region, neither with C146. However, the glutamine side chain orients in a different direction and the hydrogen bond made by the  $\epsilon$ -nitrogen of the guanidinium group to the main chain carbonyl in the wild type complex is conserved in the R166Q-dUMP binary complex, made now by the side chain amino group of glutamine 166. Also, the contacts made by Q166 mainchain with D205 and C146 are maintained as in the wild type structure (Figure 4.7.).

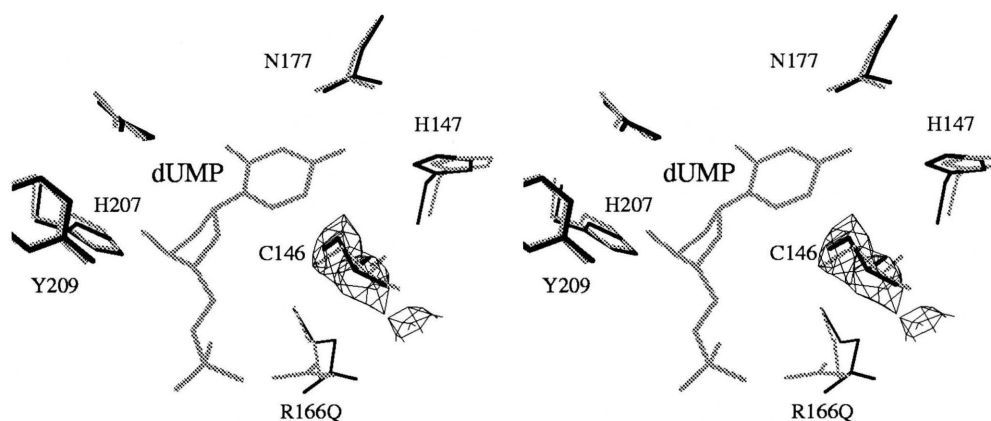
The main-chain backbone of the R166-binary complex is similar to the wild-type protein, and the RMS deviation between the C $\alpha$  backbones of the wild type binary complex and the apo-R166Q is 0.23 Å.



**Figure 4.7. Stereo view of the mutation site on the apo-R166Q TS.** The apo-R166Q mutant is drawn in black and the wild type binary complex with dUMP (Roberts and Montfort, in preparation) was superimposed and drawn in gray. Hydrogen bonds are drawn as dashed lines, and the cysteine sulfur atom is drawn as a sphere.

A more detailed analysis reveals several shifts in the position of the side-chains of active-site residues which would preclude nucleotide binding. First, the active site Cys 146 has moved 1.4 Å towards the active site such that the thiol group would clash with the C5 and C6 atoms of the nucleotide pyrimidine ring if it were present (Figures 4.7. and 4.8.). After superimposing the wild type binary and R166Q unliganded structures, C5 and C6 were 2.3 and 2.6 Å from the C146 thiol. The imidazole ring of His 147 also moves 0.53 Å towards the active site, and the side chain ring CD2 was in a short contact of 3.35 Å to the O4 atom of the pyrimidine ring in the wild type complex (Figure 4.8). Thus, it seems that R166 not only serves to bind the phosphate group of dUMP but is also necessary for the correct positioning of C146. This role will be more evident on the mutant structures obtained in the presence of folates and nucleotide. From all the above, it seems that R166 has several roles, in nucleotide binding by contacting the nucleotide phosphate, and in positioning the catalytic cysteine ready for nucleophilic attack of the nucleotide pyrimidine ring.

In contrast to R166Q, in the R126'E mutant nucleotide binding is precluded by a shift of the glutamate towards the phosphate binding site (Strop *et al.* , 1997). Upon addition of folates, the R126'Q mutant binds nucleotide although in a shifted conformation, where only one active site forms the covalent adduct with dUMP (Strop *et al.* , 1997). As for R166Q, addition of folate does not help nucleotide binding, precisely because the new position of the folate partially occupies the nucleotide binding site, as described below.



**Figure 4.8. SA-omit map of the apo-R166Q TS mutant.** The catalytic C146 was omitted from the map calculation, and the map was drawn around the wild type nucleotide-binding site. The apo-R166Q mutant is drawn in black and the wild type binary complex with dUMP was superimposed and drawn in gray. The map was contoured at  $3\sigma$ . H147 and C146 of the mutant make unfavorable contacts with dUMP of the wild type structure.



### 4.3.2 R166Q crystallized in the presence of dUMP-PDDF and R166Q-FdUMP-CH<sub>2</sub>H<sub>4</sub>folate

The *E. coli* R166Q TS mutant was crystallized in the presence of dUMP-PDDF or FdUMP-CH<sub>2</sub>H<sub>4</sub>folate. The resulting crystals belong to the P6<sub>3</sub> space group and diffracted to 2.2 and 2.3 Å, respectively. Both complexes were isomorphous with the wild-type corresponding complexes, which were used as initial models for difference electron density maps.

In both structures, the main feature is the apparent binding of the cofactor or antifolate to the active site in an extended conformation that overlaps with the nucleotide-binding site (Figure 4.9 and 4.10). Careful model building and multiple calculations of SA-omit maps revealed electron density for the folate and for an ion modeled as a sulfate molecule, bound in the phosphate-binding site. The electron density corresponding for the inhibitor PDDF is poor but is best modeled in an extended conformation. This probably results from the absence of nucleotide in the active site, as was also observed in the crystal structure of the binary complex of wild type *E. coli* TS and PDDF (Kamb *et al.*, 1992).

Overall, in both mutant structures, the enzyme has a closed conformation. The RMS deviations between the C $\alpha$  backbones of the wild type and R166Q complexes are listed in Table 4.3. From a comparison of RMS deviations and a superposition of the models, it seems clear that although the R166 complexes have a slightly wider conformation, they are closer to the closed conformation. Residues that make van der

Waals contact with the folates such Trp 80 and 83 make van der Waals contacts with the extended folate, but in a different manner due to the extended conformation of the ligand. Although residues like Y209 which are sensitive to active site opening superimpose with the position of the same residue in the wild type complex, disorder in the R21-loop and the C-terminus is probably due to lack of nucleotide phosphate and to missing contacts between the active site and the folate.

PDDF is better ordered at the glutamate tail, which is in contact with K48 *via* a water molecule as in the wild-type ternary complex. The PABA ring and the propargyl groups are still ordered, making van der Waals contacts with the active site. However, the position of the quinazoline ring is less reliable, likely due to a free rotation of the quinazoline of PDDF (Figure 4.9.).

The quality of the electron density at the active site of the R166 mutant crystallized in the presence of dUMP and CH<sub>2</sub>H<sub>4</sub>folate was hard to interpret. Important features of the structure are that the nucleotide is not bound, and that the electron density corresponding to the glutamate and PABA ring of the cofactor are readily interpreted. However is not clear whether the five-membered ring is open, and apparently the catalytic cysteine is oxidized in a manner not readily interpretable. Therefore, the ligands have not been refined and that structure will not be further analyzed.

It is not clear why the addition of folate seems to stabilize binding of sulfate on the phosphate-binding site. It can be thought that the quinazoline ring of PDDF contributes with a hydrogen bond to the arginine cluster and that is enough to stabilize the binding of the sulfate. Assuming that such is the case, a hydrogen bond has been drawn in

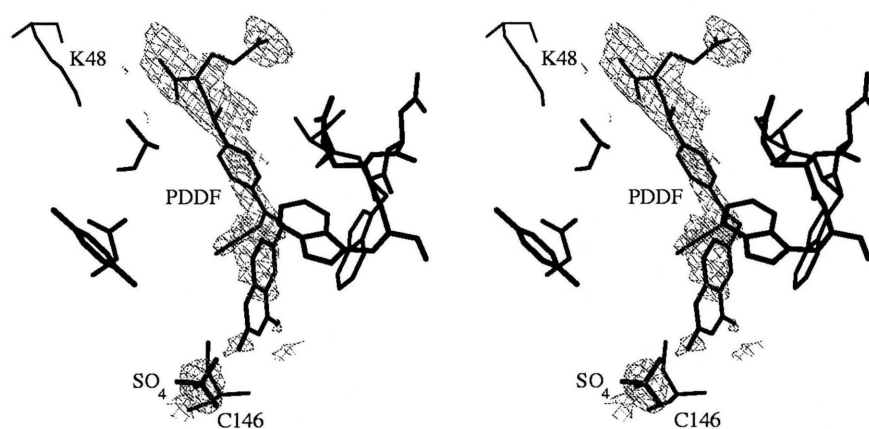
Figure 4.10.B, between the sulfate ion and the 2-amino group of PDDF. The elevated thermal factors of that portion of the antifolate suggest that the ring is poorly ordered.

A comparison of the R166Q structures with wild type complexes suggest that the folate-bound R166 is more similar to the ternary complex compared to the binary or apo TSs (Table 4.3). The absence of nucleotide in the R166Q active site despite including a final 3 mM concentration of ligand, suggests that the mutant has completely lost the affinity for the nucleotide.

The second role of R166 seems to be the alignment of Cys 146 towards the active site cavity. The change in orientation of C146 due to loss of the contact with R166 becomes more evident in the R166Q-PDDF complex (see Figure 4.7. and 4.10.B). Another consequence of the extended conformation of the folate is that contacts with the TS C-terminus are lost. These results suggest that R166 helps to keep the active site cysteine in place, and that loss of that contact also contributes to loss of nucleotide binding. As for a role in the chemical catalysis of the enzyme, namely the destabilization of the covalent adduct between Cys 146 and C6 from the nucleotide (Hyatt *et al.* , 1997), it is not possible to determine its effect since nucleotide binding is lost in this mutant.

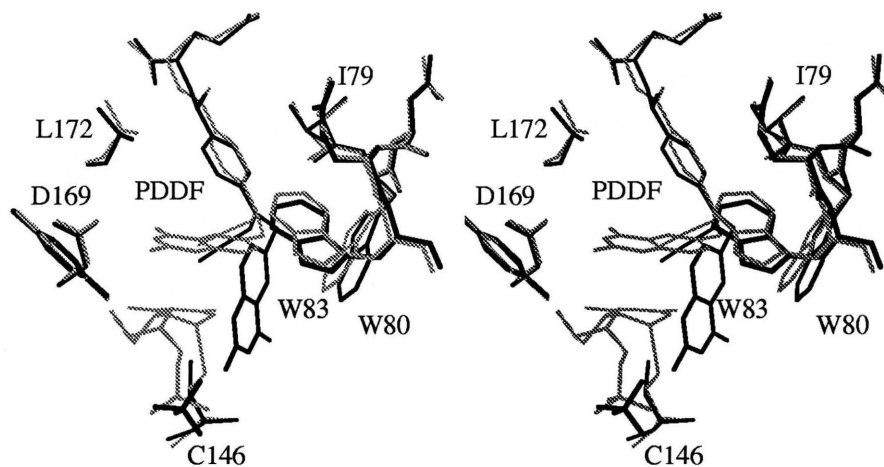
As concluded for K48Q, R166Q seems to be inactive even when it can induce the closed conformation in the presence of folate and nucleotide, since the first step of the reaction, namely nucleophilic attack of C146 to the pyrimidine ring is impaired due to distortion of the catalytic cysteine 146. At the same time, R166 contributes with two hydrogen bonds to binding of the nucleotide phosphate in the wild type complex, which are lost by mutation to the shorter and neutral glutamine residue. Combination of these

two factors may account for the 3,400-fold drop in  $k_{\text{cat}}$  on the R166Q mutant. In other words, R166Q is unable to bind the reaction intermediates and therefore has lost its catalytic power.

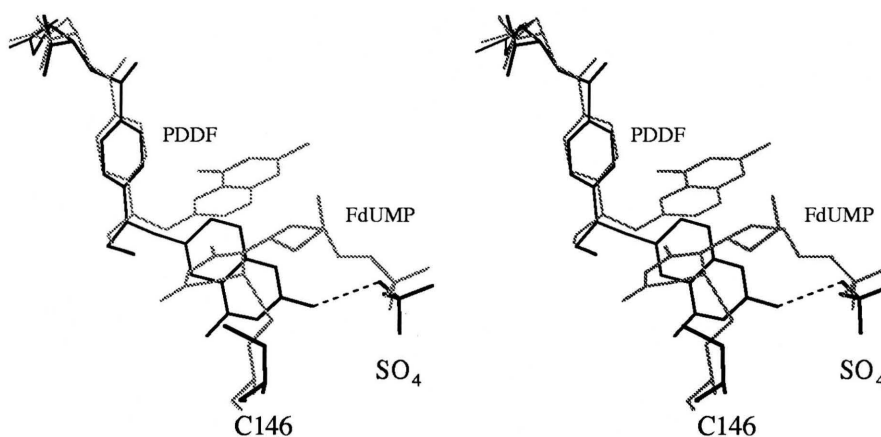


**Figure 4.9. SA-omit map of the complex of the R166Q TS mutant with PDDF and SO<sub>4</sub>.** The map was contoured at 2  $\sigma$  for the ligands.

A)



B)



**Figure 4.10. Superpositioning of the R166Q-PDDF and the wild type TS-dUMP-PDDF complexes.** The wild type ternary complex was superimposed and drawn in gray, and the mutant structure is drawn in black. **A)** The antifolate in the R166Q mutant has an extended conformation, but in a closed conformation. PDDF contacts W80 and W83 as in the wild type ternary complex. **B)** Active site view to show the rotation of the C146 side chain conformation in the R166Q mutant complex. The quinazoline group of PDDF makes a hydrogen bond with the sulfate ion, drawn as a dotted line. Other active site residues were omitted for clarity.

**Table 4.3 Comparison of wild type ternary complexes and R166 mutant *E. coli* TS structures**

	WT-dUMP	WT-dUMP- PDDF	WT- FdUMP- CH <sub>2</sub> H <sub>4</sub> folate	WT-5NO <sub>2</sub> - dUMP
R166-PDDF-SO <sub>4</sub>	0.72	0.26	0.28	0.37
R166-CH <sub>2</sub> H <sub>4</sub> folate-SO <sub>4</sub>	0.65	0.44	0.44	0.47

RMS deviations are in Angstroms, calculated with the program LSQMAN (Kleywegt and Jones, 1994, Kleywegt and Jones, 1997a)

## CHAPTER 5

### SUMMARY AND CONCLUSIONS

Thymidylate synthase is an ideal system to address structure-function questions regarding catalysis and ligand-induced conformational change in non-allosteric enzymes. TS is readily overexpressed and purified in milligram quantities, and is amenable for crystallographic studies. Although there are more than seventy TS crystal structures, and the focus of much of the TS structural work has been towards the design of inhibitors with anticancer activity, most of them are of TS from *E. coli* and *L. casei*. There is only one determination of human TS, that of the apo protein at low resolution (Schiffer *et al.* , 1995).

In Chapter 3, I report the crystal structure of rat TS complexed with substrate dUMP and the anticancer drug Tomudex, in two crystal forms (Sotelo-Mundo *et al.* , 1999). This is the first structural determination of a mammalian TS with ligands bound. Unlike the *E. coli* TS, the rat enzyme adopts an open conformation, with the C-terminus



disordered. Several factors may together produce this result, including changes in the chemical structure of the antifolate and differences in amino acid sequence between the bacterial and mammalian TS.

Tomudex differs from the parent compound PDDF in three ways: the quinazoline 2-amino has been substituted by a methyl group, the PABA group has been changed to the isosteric thiophene ring, and the propargyl group has been removed all together. (Figure 1.2). Loss of the 2-amino group precludes formation of a hydrogen bond with the penultimate residue of the C-terminal loop, and loss of the propargyl group removes extensive hydrophobic contacts with the TS active site. That the mammalian TS ternary complex has an open conformation is an unexpected result, but of great importance since this is the type of information that can only be obtained through structural studies, and may initiate a new round of structure-based drug design (Verlinde and Hol, 1994). Ligands such as BW1843U89 that were thought to bind in a similar manner to other folates have been found to produce significant distortion into the *E. coli* TS active site (Weichsel and Montfort, 1995, Stout and Stroud, 1996).

In Chapter 4, I report studies on the role of individual amino acids in ligand binding and conformational change. Two single-site mutants were studied, K48Q and R166Q. In both mutants, ligand binding was severely impaired as revealed in the crystal structures of the mutants crystallized in the presence of substrate, cofactor or inhibitor. Loss of the positive charge of lysine 48 in the folate binding site did not prevent formation of the covalent adduct with the catalytic cysteine or interfere with active site closure. However K48Q formed binary complexes with dUMP or 5-NO<sub>2</sub>dUMP.

Unexpectedly, the K48Q mutant was unable to bind dUMP-PDDF or FdUMP-CH<sub>2</sub>H<sub>4</sub>folate from the crystallization solution, resulting in crystals of the apo-K48Q mutant. The binding site of folate is formed mainly from van der Waals contacts with hydrophobic residues and the nucleotide, three hydrogen bonds to the enzyme and two to ordered water molecules. Loss of the electrostatic interaction with the  $\epsilon$ -amino of K48 seems to prevent productive binding of CH<sub>2</sub>H<sub>4</sub>folate or PDDF to the TS active site.

However, there are folate-based inhibitors that do not seem to interact with K48, such as BW1843U89, a quinazoline-based antifolate that distorts the TS active site. The crystal structure of the ternary complex of *E. coli* TS with dUMP and BW1843U89 shows that although the compound has an  $\alpha$ -glutamate-tail like CH<sub>2</sub>H<sub>4</sub>folate and PDDF, it does not contact K48 (Weichsel and Montfort, 1995). The binding of BW1843U89 has been assessed in the *E. coli* K48Q TS mutant using CD spectroscopy (F. Maley, personal communication), and determined that BW1843U89 binds to the mutant in a similar manner as to wild type TS (F. Maley, personal communication).

R166 is a critical residue both for nucleotide binding and for orientation of the active site cysteine. Unlike K48Q, R166Q prevents nucleotide or ion binding to TS. Crystallization of R166 with either PDDF or CH<sub>2</sub>H<sub>4</sub>folate and nucleotides, produces a binary complex where the folate adopts an extended conformation similar to that observed for the non-productive binary complex of *E. coli* TS and PDDF (Kamb *et al.* , 1992). In the R166-PDDF structure, the quinazoline ring points toward the phosphate-binding site, apparently providing an extra amino group that helps to bind a sulfate ion from the crystallization buffer. R166 also maintains the catalytic cysteine in an optimal

position for nucleophilic attack of the nucleotide, since the thiol group is distorted in all three mutant structures. Whether R166 reduces the  $pK_a$  of C146 is not known, and further studies of the R166Q mutant would probably be difficult to perform due to the loss of nucleotide binding.

### 5.1 Future Studies

The work presented in this dissertation generates several hypotheses regarding structure, function and mechanism of TS. The unexpected open conformation of the rat TS in the presence of Tomudex and dUMP suggest that tighter binding inhibitors are required to induce the closed conformation. The crystallization of the rat enzyme with PDDF-dUMP,  $CH_2H_4$ folate-FdUMP and BW1843U89-dUMP should be pursued, as all of these ternary complexes should favor the closed conformation. Due to the ease of crystallization of rat TS, it can be used for more specific structure-function studies using single or multiple residue mutants, exploiting now the differences in amino acid sequence between the bacterial and mammalian TSs on ligand binding.

Several antifolates where position 2 of the antifolate-quinazoline ring is substituted with bulky groups such as a pyrimidine or piperazine ring have sub-micromolar inhibition constants for *E. coli* TS (Hennequin *et al.*, 1996). It would be of great interest to investigate the effect of large quinazolinyl-substituents on the protein conformation of the mammalian TS. Spectroscopic techniques would also be useful to study the nature of the inhibition mechanism of these ligands, where stabilization of the

open conformation could operate as the main mechanism of enzyme inactivation.

However, since Tomudex is a better substrate for FPGS than PDDF or other folyl-analogs containing the PABA group (Jackmann *et al.*, 1996), it will be also of interest to determine the structure of the rat TS in complex with polyglutamylated-Tomodex and dUMP to address the effect of such process on TS inhibition.

It also remains to be determined how other novel antifolates such as BW1843U89 bind to the mammalian enzyme. In particular, it would be interesting to determine if BW1843U89 induces the closed conformation, if it distorts the active site as occurs in the bacterial enzyme, and if it stabilizes the covalent adduct with nucleotide.

For the *E. coli* TS mutants, several ligand-bound mutant structures should be determined to confirm the role of K48 and R166 in ligand binding and catalysis. Preliminary data indicate that the K48Q mutant binds BW1843U89 in a very similar manner as the wild type (F. Maley, personal communication). It may be possible that isoindolyl-based antifolates can bind in the absence of the electrostatic attraction of K48. Other ligands that could be used to study the contribution of K48 on folate binding are those that do not contain a terminal glutamyl-group such as Thymitaq (Webber *et al.*, 1996).

For the R166Q mutant, it is of great interest to determine whether the  $pK_a$  of the active site cysteine is altered due to the mutation. The determination of pH profiles of cysteine modification or nucleotide dehalogenation should determine whether R166 influences the reactivity and stability of the thiolate. However, since nucleotide binding to R166Q is very poor or inexistent, it is possible that such approach may not work. Another

complex to be determined is the binary complex with 5-NO<sub>2</sub>dUMP, to see if the nucleotide analog induces active site closure. One mutant that may be able to partially overcome the loss of nucleotide binding would be the R166K mutant. Although it is known that such mutant is unable to rescue a thymidine auxotroph, it may be useful for study of the early steps of the reaction. It is unlikely that lysine 166 would be able to contact the active site cysteine, therefore R166K may be a better mutant for addressing the reactivity of the thiolate and the stability of the covalent adduct in later stages of the TS reaction.

## **APPENDIX A**

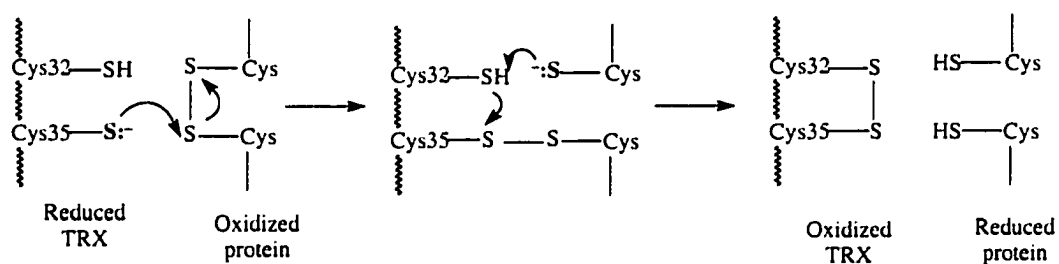
### **CRYSTAL STRUCTURE OF HUMAN THIOREDOXIN OXIDIZED BY A POTENTIAL ANTI-CANCER DRUG**

#### **A.1. Introduction**

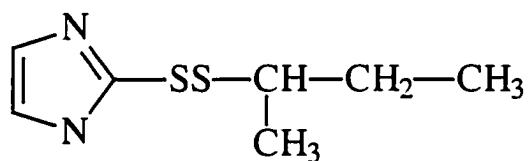
The following describes experiments concerning the crystal structure determination of human thioredoxin (TRX) oxidized by a biologically active disulfide with anticancer activity. The project was aimed at understanding the mechanism of action of the disulfide inhibition.

Thioredoxin (TRX) is a disulfide oxidoreductase protein found in every organism, from archaeobacteria to humans (Holmgren, 1985, Holmgren, 1995). The TRX active site contains the conserved sequence Trp31-Cys32-Gly33-Pro34-Cys35, where the two cysteines perform the disulfide exchange reaction, leading to an oxidized TRX active site (covalent bond between Cys 32 and Cys 35) and a reduced substrate (Figure A.1.A.

A)



B)



**Figure A.1. Mechanism of disulfide exchange by TRX, and chemical structure of the disulfide TRX-inhibitor IV-2.** A) After Holmgren (1985). B) Structure of 1-methylpropyl 2-imidazolyl disulfide (IV-2), from Kirkpatrick *et al*, (1992)

The redox exchange partners of TRX are proteins with important biological roles such as ribonucleotide reductase, an enzyme critical for deoxyribonucleotide synthesis (Holmgren, 1985), transcription factor NF- $\kappa$ B (Matthews *et al.*, 1992, Qin *et al.*, 1996), glucocorticoid receptor (Grippe *et al.*, 1983) and interleukin 2 (Wakasugi *et al.*, 1990), among others. In plants, TRX regulates malate dehydrogenase and fructose-1,6-bisphosphatase, key photosynthetic enzymes, *via* a light-driven redox mechanism involving the ferredoxin-thioredoxin reductase complex (Holmgren, 1985, Buchanan *et al.*, 1994).

A remarkable feature of human TRX is that in addition to its role in the cytoplasm, it also functions as an autocrine growth factor in the extracellular space. Human TRX was found to be the growth factor secreted by virus-transformed leukemic cells (ADF) (Wakasugi *et al.*, 1990). Also, solid tumors overexpress and secrete TRX via a leaderless pathway (Rubartelli *et al.*, 1992), and mammalian cells overexpressing TRX show a transformed phenotype (Powis *et al.*, 1994). That the growth factor activity of TRX is linked to its oxidoreductase function is shown by the fact that wild type human TRX stimulates 3T3 fibroblast cell growth in culture, but not when active site cysteines 32, 35 or both are mutated to serines (Oblong *et al.*, 1994).

Hence, TRX is an attractive target for anticancer drug design by compounds that inhibit its interaction with the TRX reductase system. Since the growth factor stimuli occurs in the extracellular space, drugs targeted to TRX do not need to enter the cell and therefore would be less likely to interfere with other cellular pathways. Those



considerations are important for the chemical nature of the chemotherapeutics that can target TRX.

One TRX inhibitor and potential anticancer drug compound is 1-methylpropyl 2-imidazolyl disulfide (IV-2), shown in Figure A.1.B (Kirkpatrick *et al.*, 1992, Oblong *et al.*, 1995). IV-2 is an heterodisulfide that stops TRX-dependent growth of transformed cells *in vitro* (Kirkpatrick *et al.*, 1998). The crystal structure of human TRX in different redox states has been previously determined (Weichsel *et al.*, 1996), however the mechanism of inhibition by disulfides remains to be determined. Therefore, I pursued the structural characterization of the human TRX/IV-2 complex. This appendix reports the crystal structure determination of human TRX oxidized by the disulfide IV-2. I will show that IV-2 functions as a disulfide oxidant, which attacks the active site. Preliminary results suggest that IV-2 can also modify Cys 73. Future work may explain in more detail the mechanism of action of disulfides on TRX.

## A.2. Materials and methods

### A.2.1. Protein overexpression and purification

Recombinant human TRX was overexpressed in *E. coli* (Oblong *et al.* , 1994). Transformed *E. coli* BL21( $\lambda$ DE3)/pLysS with a pET-3a expression vector (Novagen, Milwaukee, WI) containing the coding region of the wild-type human TRX gene were provided by Dr. Garth Powis from the Arizona Cancer Center. The transformed cells were grown at 37 °C in Luria broth media with 100 mg/L of ampicillin, 25 mg/L of chloramphenicol, 1% glucose and 1 mM MgSO<sub>4</sub> until an optical density ( $A_{600}$ ) of 0.6 was reached. At that point the cells were induced by addition of 0.4 mM IPTG and the culture was incubated for three hours at 37 °C. The media was centrifuged and the pellet resuspended and washed with 0.9 % (w/v) NaCl and stored at -20 °C.

The protein was purified as previously described (Andersen *et al.*, 1997). The frozen cell pellet was resuspended in 100 mM Tris-HCl pH 8.0, 1 mM EDTA, 1 mM PMSF, 5 mM DTT, to a ratio of 8 gr. of cells per mL of buffer. The mixture was sonicated and the lysate was spun at 215,000x g for 1 hr. The supernatant was loaded on a Q Sepharose column (Pharmacia, Piscataway NJ), washed with 10 mM Tris-HCl pH 8.0, 1 mM EDTA, 5 mM DTT and eluted with the same buffer with a salt gradient from 0 to 500 mM NaCl. The fractions obtained were analyzed by SDS-PAGE and pooled. Ammonium sulfate was added to the pooled fractions to a final concentration of 85%, and incubated in ice for at least one hour. The ammonium sulfate precipitate was recovered by

centrifugation at 10,000 x g and resuspended in 10-15 mL of 10 mM Tris-HCl pH 7.5, 1 mM EDTA, 5 mM DTT. The volume of the solution was reduced by ultrafiltration using an Amicon ultrafiltration cell. The concentrated solution was loaded into a Sephacryl S-100 HR gel filtration FPLC column, equilibrated with 10 mM Tris-HCl pH 7.5, 1 mM EDTA, 200 mM NaCl and 5 mM DTT. The column was eluted with the same buffer at a flow rate of 1 mL/min. The fractions were analyzed by SDS-PAGE and if needed, they were pooled, concentrated and ran over the gel filtration column until no impurities were observable in an overloaded SDS-PAGE gel.

#### **A.2.2. Crystallization and data collection**

Non-reducing crystallization conditions were found from initial leads obtained by sparse-matrix analysis (Jancarik and Kim, 1991). Crystals were obtained by the hanging drop vapor diffusion method using recombinant wild type human TRX. TRX was kept reduced by maintaining DTT at 5 mM during protein purification and storage, but it was removed just before reaction with the disulfide IV-2 by gel filtration using a desalting column (BioRad). TRX was reacted with IV-2 to a final concentration of 0.4 mM TRX dimer and 2 mM disulfide in 10 mM HEPES pH 7.5. The solution was passed through a 0.2  $\mu$  membrane filter, and hanging drops were set up by mixing the protein solution (2-6  $\mu$ l), with equal volumes of the reservoir liquid. The drops were equilibrated against 200 mM sodium acetate pH 5.0, PEG 4000 14%, at 25°C. Monoclinic crystals appeared overnight and continued to grow for a few days more.

The crystals diffracted at 1.9 Å and a single crystal was used for the entire data collection. X-ray diffraction data were measured at room temperature with a FAST area detector and an Enraf-Nonius FR571 rotating anode as the X-ray source. Intensities were estimated using the programs MADNES (Messerschmidt and Pflugrath, 1987) and PROCOR (Kabsch, 1988). Data were processed with ABSURD, ROTAVATA and AGROVATA from the CCP4 suite (CCP4, 1994). Data collection statistics are summarized in Table A.1.

### **A.2.3 Phase Determination, Model Building and Structure Refinement**

Phases were estimated by difference Fourier maps, using the X-ray crystal structure of reduced human TRX, PDB entry 1ert (Weichsel *et al.*, 1996) as a starting model. The model was included directly in X-PLOR, ver. 3.851 (Brunger, 1992b). A bulk-solvent model (Jiang and Brunger, 1994) was calculated to use all low-resolution reflections, and  $R_{\text{free}}$  was used to monitor all refinement steps. General crystallographic calculations were made with CCP4 (CCP4, 1994). Water molecules were built using a  $|F_o - F_c|/\sigma_c$  map contoured to a  $3\sigma$  level or higher and positioned to have reasonable hydrogen bonding with the model. Several rounds of simulated annealing, positional and individual B-factor refinement, plus manual model rebuilding using O ver. 5.1 (Jones *et al.*, 1991), were needed to optimize the model. Once no further drop in  $R_{\text{free}}$  was obtained, positional and individual B-factor refinement was performed for the entire

dataset. The stereochemical quality of the model was assessed with X-PLOR and PROCHECK (Laskowski *et al.*, 1993).

### **A.3. Results**

#### **A.3.1. Crystallization of IV-2/TRX**

Prior to this work, DTT was required to obtain TRX crystals, resulting in reduced protein in the crystals (Weichsel *et al.* , 1996). Since I was interested in investigating the effect of IV-2 on TRX, and DTT would reduce the ligand, being itself a disulfide, I had to start from conditions free of reducing agent. This was accomplished by first reacting reduced TRX free of reducing agent with the disulfide IV-2, and using that protein solution to search crystallization conditions.

The sparse-matrix search provided initial leads with the precipitating agent PEG 4000, and at higher pH's compared to the previously determined conditions. Concentration of precipitant and pH were optimized until single-crystals were obtained spontaneously with sizes of up to 1 mm in the longest dimension. Nonetheless, the resulting crystals were isomorphous with the previous crystal forms (monoclinic C2), allowing me to use the Fourier difference map method to determine the structure.

**Table A.1. Crystallographic results for IV-2 oxidized wild type human thioredoxin.**

Complex	Disulfide Oxidized TRX
Space group	C2
Cell Constants (Å)	$a=67.59$
	$b=26.43$
	$c=51.80$
	$\beta=95.34$
Resolution (Å) (°)	1.9
Unique reflections	6359
Total reflections	12479
Completeness	85
Multiplicity	2.1
$R_{\text{sym}}^a$	7.5
$R_{\text{cryst}}^b$	18.9
$R_{\text{free}}^c$	24.1
RMS Deviations	
Bond distances (Å)	0.008
Bond angles (°)	1.1
No. of water molecules	39
PDB accession number	2eru

<sup>a</sup> $R_{\text{sym}} = 100(\sum_h |I_h - \langle I \rangle|) / (\sum_h I_h)$ , where  $\langle I \rangle$  is the mean intensity of all symmetry related reflections  $I_h$ .

<sup>b</sup> $R_{\text{cryst}} = 100(\sum |F_{\text{obs}} - F_{\text{calc}}|) / (\sum F_{\text{obs}})$ , for all non-zero reflections.

<sup>c</sup> $R_{\text{free}}$  as for  $R_{\text{cryst}}$  using a subset of the data (5%) not included in the refinement, as implemented in X-PLOR (Brunger, 1987).

### A.3.2. Determination of the structure of IV-2 oxidized TRX

The initial electron density difference map calculated using the structure factors from TRX-IV-2 and phases from the reduced human TRX structure showed that the active site was oxidized (Cys32-Cys35) and that Cys73 was forming a covalent bond across the two-fold axis with a symmetry related molecule. The remaining cysteines 62 and 69 were found reduced, therefore ruling out the possibility of a covalent modification by the disulfide IV-2. Thus IV-2 worked only as an oxidizing agent under the crystallization conditions used.

Positive and negative electron density between the S $\gamma$  atoms of Cys 32 and 35 indicated the formation of covalent disulfide bond (Figure A.3). They were rebuilt accordingly and the C32-C35 disulfide bond was refined as such. In this structure, the asymmetric or unique portion of the crystal unit cell contains one TRX monomer, however TRX makes a disulfide bond with a symmetry-related monomer produced by a two-fold rotation (C73-C73'). This disulfide bond has been shown not to be a crystallographic artifact, but to be critical for the function and a unique feature of the mammalian TRXs (Weichsel *et al.* , 1996, Andersen *et al.* , 1997).

The final model consists of 105 amino acid residues and 39 water molecules, with one in a special position in the crystallographic two-fold axis. Pro 75 is in a *cis*-conformation, as in the reduced TRX model (Weichsel *et al.* , 1996). This oxidized TRX structure had better refinement parameters ( $R_{\text{free}} = 0.24$ ) compared to the previously reported air-oxidized TRX model (PDB entry 1eru), which had a  $R_{\text{free}} = 0.33$ . The RMS

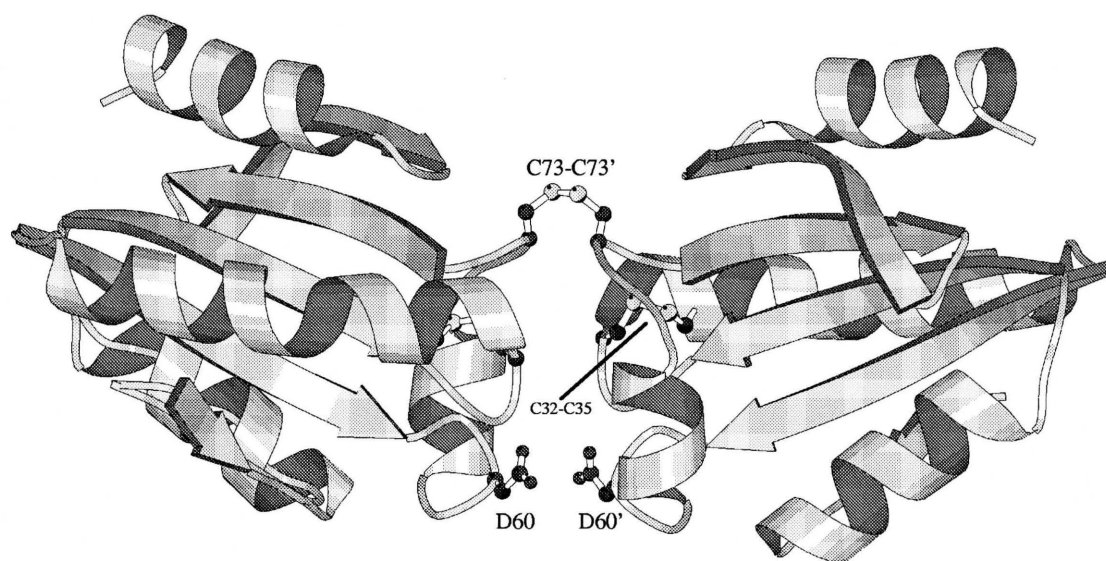
deviation between the two C $\alpha$  backbones is 0.22 Å and for the whole molecule is 0.78 Å. The model has good stereochemistry (Table A.1), and 93% of the residues were in the core region of the Ramachandran plot (Laskowski *et al.* , 1993). Met 1 and the side chains of residues Glu 13 and Lys 39 were disordered, therefore their B-factors were set to 100 Å<sup>2</sup>.

### A.3.3. Conformational change in TRX produced by oxidation

Active site oxidation leads to a local conformational change in the vicinity of C32 and C35, which does not affect the general conformation of TRX. The C $\alpha$  backbone of disulfide-oxidized TRX superimpose with the reduced TRX structure with an RMS deviation of 0.17 Å, within the upper limit of coordinate error as estimated with a Luzzati plot (0.21 Å). During air oxidation, the formation of the C32-C35 disulfide leads to a tightening of the  $\alpha$ 2 helix (Weichsel *et al.* , 1996). This did not occur in the present structure, as the expected hydrogen bonds were missing. The distance between C32 carbonyl and K36 amide was 3.46 Å, and between G33 carbonyl and M37 amide was 3.3 Å. However, the hydrogen bond between the Cys 32 sulfhydryl and Ala amide 29 is lost, as expected for the change in redox state of that residue and a new hydrogen bond between the C32 sulfur and C35 amide is found in both structures.

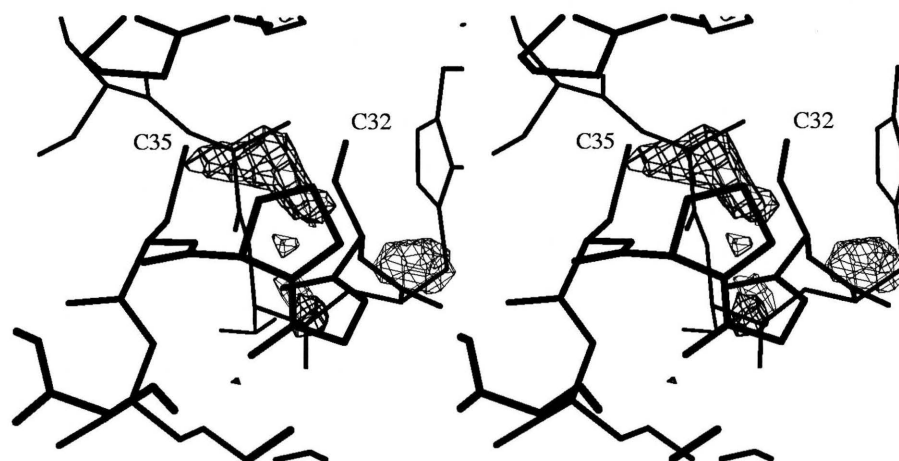
The hydrogen bond across the dimer interface found in the reduced structure at pH 3.8 between the symmetry related D60-D60' side chains is lost in the present structure, as expected at pH 5.0.



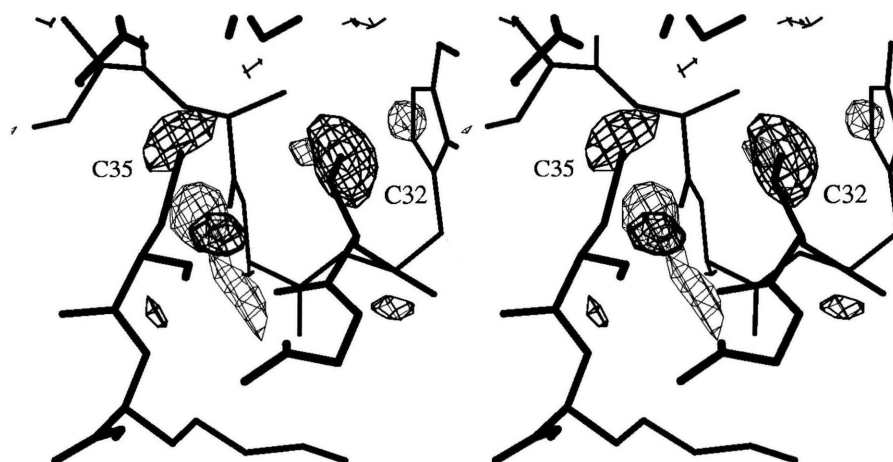


**Figure A.2. Ribbon diagram of the oxidized human thioredoxin dimer.** The intermolecular disulfide C73-C73', the oxidized active site C32-C35 and Asp 60 across the dimer interface are drawn as ball-and-sticks. From this work, PDB entry 2eru.

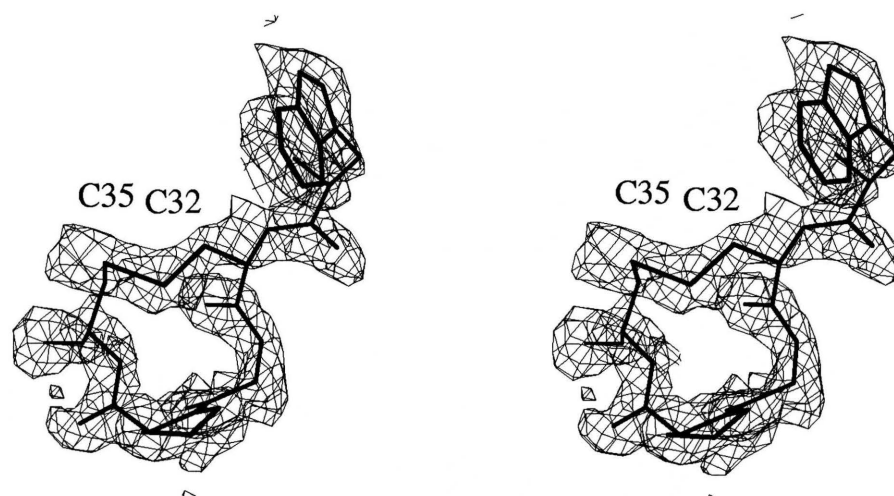
A)



B)



**Figure A.3. Difference electron density maps between reduced and oxidized human TRX.** A)  $|F_{\text{oxi}} - F_{\text{red}}| \alpha_{\text{red}}$  electron density map contoured at  $+1\sigma$ . The model displayed is the reduced human TRX active site. B) Same difference map contoured at  $-1\sigma$ .



**Figure A.4. Electron density map of the oxidized human TRX active site.** The  $2|F_o - F_c| \alpha_c$  map was contoured at  $+1 \sigma$ . The model corresponds to residues 31 to 35 of the refined oxidized human TRX.

The Trp 31 indole ring moves toward the dimer interface in both oxidized structures, with a W31-W31' CZ2 distance across the dimer interface of 4.6 Å in the disulfide-oxidized and 5.6 Å in the reduced TRX. This movement plus the change in orientation of Asp 60, leads to the positioning of an ordered water molecule along the two-fold axis, in hydrogen bond contact with W31 and D60.

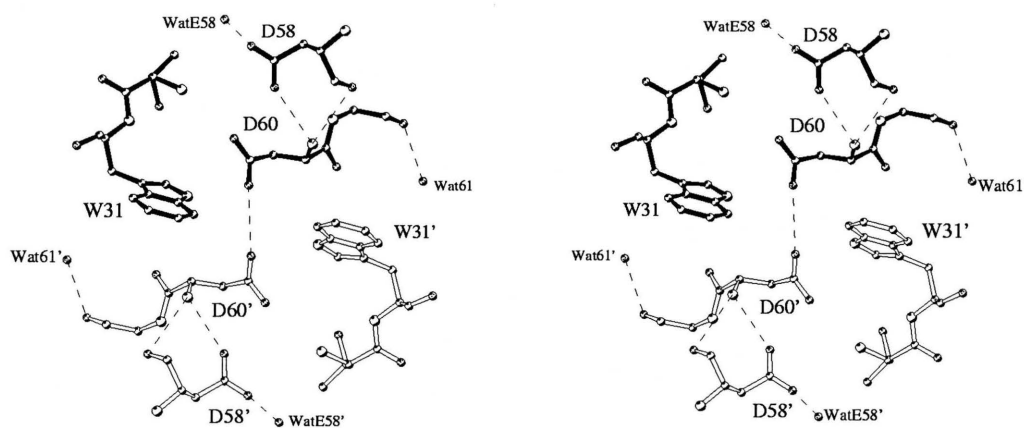
#### A.4. Discussion

The major conclusions of this work are that although IV-2 was designed to irreversibly add to C32 (Kirkpatrick *et al.* , 1992), the present study provided the first evidence that IV-2 is in fact a substrate for the TRX-TRX reductase system. Also this structure confirmed and improved the previously determined oxidized TRX structure, under different crystallization conditions and pH.

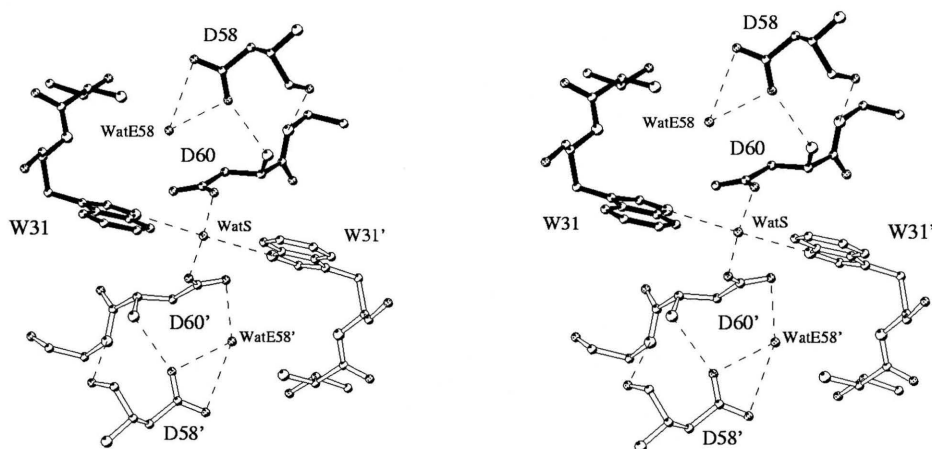
The active site oxidation in TRX produces a local conformational change due to the shift of C32 and C35 to form the disulfide bond. The major movement occurring at the active site corresponds to a rotation of 29° in the C32  $\chi_1$  angle to form the disulfide bond with C35. The nearby residues accommodate this torsional change through small adjustments in  $\phi$  and  $\psi$ , but in a different way compared with the air-oxidized TRX. .

Active site oxidation with IV-2 or the increase in pH produced changes in the dimer interface reflected in a shift in W31.

A)



B)



**Figure A.5. Effects on hydrogen bonding at the human TRX dimer interface by changes in pH and oxidation.** Stereo views of the TRX dimer interfaces. **A)** Reduced human TRX structure determined at pH 3.8 (Weichsel et al, 1996). **B)** Oxidized human TRX structure determined at pH 5.0, this work.

The symmetry-related indole rings move closer together, and the movements of the D60 side chains allow an ordered water molecule (WatS) to position in the dimer interface.

None of the structural cysteines (62 or 69) were found modified by the disulfides, since they are buried, and the covalent dimer between C73-C73' was found as in all the previous work. Further determination of the dissociation constant of the TRX dimer (Andersen *et al.* , 1997) showed that in fact I had oxidized reduced TRX dimers when reacted with IV-2, due to the high protein concentration used.

This work lead to further experiments done in our group and the Powis lab that suggest that IV-2 functions through modification of C73 rather than through active site oxidation. I did not find C73 modified by the alkyl portion of IV-2 on the crystals grown in the presence of IV-2 since C73 was not available due to formation of TRX homodimers. However, it remains to be investigated if the mechanism of IV-2 biological activity depends on depletion of the extracellular pool of reduced TRX, or if it occurs *via* thioalkylation of Cys 73 and further modification of the TRX reductase active site.

## REFERENCES

- Agarwalla, S., S. LaPorte, L. Liu, J. Finer-Moore, R. M. Stroud and D. V. Santi (1997). A novel dCMP methylase by engineering thymidylate synthase. Biochemistry **36**(50): 15909-17.
- Andersen, J. F., D. A. R. Sanders, J. R. Gaskaska, A. Weichsel, G. Powis and W. R. Montfort (1997). Human thioredoxin homodimers: regulation by pH, role of Asp 60, and crystal structure of the Asp 60 -> Asn mutant. Biochemistry **36**: 13979-88.
- Austin, J. C., A. Fitzhugh, J. E. Villafranca and T. G. Spiro (1995). Stereoelectronic activation of methylenetetrahydrofolate by thymidylate synthase: resonance Raman spectroscopic evidence. Biochemistry **34**: 7678-7685.
- Barrett, J. E., D. A. Maltby, D. V. Santi and P. G. Schultz (1998). Trapping of the C5 Methylene Intermediate in Thymidylate Synthase. J. Am. Chem. Soc. **120**: 449-450.
- Belfort, M., G. F. Maley and F. Maley (1983). Characterization of the Escherichia coli thyA gene and its amplified thymidylate synthetase product. Proc Natl Acad Sci U S A **80**(7): 1858-61.
- Bellisario, R. L., G. F. Maley, J. H. Galivan and F. Maley (1976). Amino acid sequence at the FdUMP binding site of thymidylate synthase. Proc. Natl. Acad. Sci. USA **73**: 1848-1852.
- Bertino, J. R. (1997). Chemotherapy of colorectal cancer: history and new themes. Semin Oncol **24**(5 Suppl 18): S18-3-S18-7.
- Bosch, L., E. Habers and C. Heidelberger (1958). Studies on fluorinated pyrimidines B. Effects on nucleic acid metabolism in vitro. Cancer Res. **18**: 335-343.
- Branden, C. I. and T. A. Jones (1990). Between objectivity and subjectivity. Nature **343**: 687.
- Brunger, A. T., Kuriyan, J, Karplus, M (1987). Crystallographic R-factor refinement by molecular dynamics. Science **235**: 458-460.
- Brunger, A. T. (1992a). Free *R* value: a novel statistical quantity for assessing the accuracy of crystal structures. Nature **355**: 472-475.
- Brunger, A. T. (1992b). A System for X-ray Crystallography and NMR., Yale University Press.

**REFERENCES - Continued**

- Brunger, A. T. (1993). X-PLOR: A System for X-ray Crystallography and NMR, Yale Univ. Press, New Haven, ed. 3.1.
- Brunger, A. T., A. Krukowski and S.-C. P. J. Erickson (1990). Slow-Cooling Protocols for Crystallographic Refinement by Simulated Annealing. Acta Cryst. **A46**: 585-593.
- Brunger, A. T. and L. M. Rice (1997). Crystallographic refinement by simulated annealing: methods and applications. Methods in Enzymology **277**: 243-269.
- Buchanan, B. B., P. Schurmann, P. Decottignies and R. M. Lozano (1994). Thioredoxin: a multifunctional regulatory protein with a bright future in technology and medicine. Arch Biochem Biophys **314**(2): 257-60.
- Byrd, R. A., W. H. Dawson, P. D. Ellis and R. B. Dunlap (1978). <sup>19</sup>F nuclear magnetic resonance investigation of the ternary complex formed between native thymidylate synthetase, 5-fluoro-2'-deoxyuridylate, and 5,10-methylenetetrahydrofolate. J. Am. Chem. Soc. **100**: 7478-7486.
- Carreras, C. W. and D. V. Santi (1995). The Catalytic Mechanism and Structure of Thymidylate Synthase. Ann. Rev. Biochem. **64**: 721-762.
- CCP4 (1994). The CCP4 Suite: Programs for Protein Crystallography. Acta Crystallogr. **D50**: 760-763.
- Chambers, J. L. and R. M. Stroud (1977). Difference Fourier refinement of the structure of DIP-trypsin at 1.5 Å with a minicomputer technique. Acta Crystallogr. **B33**: 1824-1837.
- Chu, E. and C. J. Allegra (1996). The role of thymidylate synthase as an RNA binding protein. Bioessays **18**: 191-198.
- Chu, E., S. M. Copur, J. Ju, T. Chen, S. Khleif, D. M. Voeller, N. Mizunuma, M. Patel, G. F. Maley, F. Maley and C. J. Allegra (1999). Thymidylate synthase protein and p53 mRNA form an In vivo ribonucleoprotein complex. Mol Cell Biol **19**(2): 1582-94.
- Ciesla, J., K. Weiner, R. Weiner, J. Reston, G. Maley and F. Maley (1995). Isolation and expression of rat thymidylate synthase cDNA: phylogenetic comparison with human and mouse thymidylate synthases. Biochem. Biophys. Acta **1261**: 233-242.



### REFERENCES - Continued

- Climie, S., L. Ruiz-Perez, D. Gonzalez-Pacanowska, P. Prapunwattana, S. W. Cho, R. Stroud and D. V. Santi (1990). Saturation Site-directed Mutagenesis of Thymidylate Synthase. J. Biol. Chem. **265**(31): 18776-18779.
- Colette Daubner, S. and P. F. Fitzpatrick (1999). Site-directed mutants of charged residues in the active site of tyrosine hydroxylase. Biochemistry **38**: 4448-4454.
- Danenberg, P. V., R. J. Langenbach and C. Heidelberger (1974). Structures of reversible and irreversible complexes of thymidylate synthetase and fluorinated pyrimidine nucleotides. Biochemistry **13**(5): 926-33.
- Dunlap, R. B., N. G. Harding and F. M. Huennekens (1971). Thymidylate synthetase and its relationship to dihydrofolate reductase. Ann N Y Acad Sci **186**: 153-65.
- Eldin, S. and W. P. Jencks (1995). Lifetimes of Iminium Ions in Aqueous Solution. J. Am. Chem. Soc. **117**: 4851-4857.
- Engh, R. A. and R. Huber (1991). Accurate bond and angle parameters for X-ray protein structure refinement. Acta Crystallogr. **A47**: 392-400.
- Erlandsen, H., F. Fusetti, A. Martinez, E. Hough, T. Flatmark and R. C. Stevens (1997). Nature Struct. Biol. **4**: 995-1000.
- Fantz, C. R., D. Shaw, J. G. Moore and H. T. Spencer (1998). Retroviral coexpression of thymidylate synthase and dihydrofolate reductase confers fluoropyrimidine and antifolate resistance. Biochem Biophys Res Commun **243**(1): 6-12.
- Fauman, E. B., E. E. Rutenber, G. F. Maley, F. Maley and R. M. Stroud (1994). Water-Mediated Substrate/Product Discrimination: The Product Complex of Thymidylate Synthase at 1.83 Å. Biochemistry **33**: 1502-1511.
- Finer-Moore, J. S., E. B. Fauman, P. G. Foster, K. M. Perry, D. V. Santi and R. M. Stroud (1993). Refined Structures of Substrate-bound and Phosphate-bound Thymidylate Synthase from *Lactobacillus casei*. J. Mol. Biol. **232**: 1101-1116.
- Finer-Moore, J. S., G. F. Maley, F. Maley, W. R. Montfort and R. M. Stroud (1994). Crystal Structure of Thymidylate Synthase from T4 Phage: Component of a Deoxynucleoside Triphosphate-Synthesizing Complex. Biochemistry **33**: 15459-15468.
- Finer-Moore, J. S., W. R. Montfort and R. M. Stroud (1990). Pairwise Specificity and Sequential Binding in Enzyme Catalysis: Thymidylate Synthase. Biochemistry **29**(30): 6977-6986.

### REFERENCES - Continued

- Fitzpatrick, P. (1998). Comprehensive Biological Catalysis. M. Sinnott, Ed. New York, Academic Press. **III**: 181-194.
- Fox, K. M., F. Maley, A. Garibian, L. M. Changchien and P. Van Roey (1999). Crystal structure of thymidylate synthase A from *Bacillus subtilis*. Protein Sci **8**(3): 538-44.
- Friedkin, M. (1959). . The Kinetics of Cellular Proliferation. J. F. Stohlman, Ed.. New York, Grune & Stratton: 97.
- Galivan, J. H., F. Maley and C. M. Baugh (1977). Protective effect of the pteroylpolyglutamates and phosphate on the proteolytic inactivation of thymidylate synthetase. Arch. Biochem. Biophys. **184**: 346-354.
- Galivan, J. H., G. F. Maley and F. Maley (1975). The effect of substrate analogs on the circular dichroic spectra of thymidylate synthase from *Lactobacillus casei*. Biochemistry **14**: 3338-3344.
- Galivan, J. H., G. F. Maley and F. Maley (1976). Factors Affecting Substrate Binding in *Lactobacillus casei* Thymidylate Synthetase as Studied by Equilibrium Dialysis. Biochemistry **15**(2): 356-362.
- Giegé, R. and A. Ducruix (1992). Crystallization of nucleic acids and proteins : a practical approach. Oxford [England] ; New York, IRL Press.
- Goodwill, K. E., C. Sabatier and R. C. Stevens (1998). Crystal structure of tyrosine hydroxylase with bound cofactor analogue and iron at 2.3 Å resolution: self-hydroxylation of Phe 200 and the pterin-binding site. Biochemistry **39**: 13437-13445.
- Grippo, J. F., W. Tienrungrroj, M. K. Dahmer, P. Housley and W. B. Pratt (1983). Evidence that the endogenous heat-stable glucocorticoid receptor-activating factor is thioredoxin. J. Biol. Chem. **258**: 13658-13664.
- Hardy, L. W., J. S. Finer-Moore, W. R. Montfort, M. O. Jones, D. V. Santi and R. M. Stroud (1987). Atomic Structure of Thymidylate Synthase: Target for Rational Drug Design. Science **235**: 448-455.
- Heidelberger, C., N. K. Chanakari and P. V. Danenberg (1957). Fluorinated pyrimidines, a new class of tumor inhibitory compounds. Nature **179**: 663-666.
- Hennequin, L. F., F. T. Boyle, J. M. Wardleworth, P. R. Marsham, R. Kimbell and A. L. Jackman (1996). Quinazoline antifolates thymidylate synthase inhibitors: lipophilic analogues with modification to the C2-methyl substituent. J Med Chem **39**(3): 695-704.

### REFERENCES - Continued

- Hentze, M. W., S. W. Caughman, T. A. Rouault, J. G. Barriocanal, A. Dancis, J. B. Harford and R. D. Klausner (1987). Identification of the iron-responsive element for the translational regulation of human ferritin mRNA. Science **238**(4833): 1570-3.
- Holmgren, A. (1985). Thioredoxin. Annu. Rev. Biochem. **54**: 237-271.
- Holmgren, A. (1995). Thioredoxin structure and mechanism: conformational changes on oxidation of the active-site sulfhydryls to a disulfide. Structure **3**: 239-243.
- Huang, W. and D. Santi (1997). Active Site General Catalysts Are Not Necessary for Some Proton Transfer Reactions of Thymidylate Synthase. Biochemistry **36**: 1869-1873.
- Huang, W. and D. V. Santi (1994). Isolation of a covalent steady-state intermediate in glutamate 60 mutants of thymidylate synthase. Journal of Biological Chemistry **269**(50): 31327 -31329.
- Hyatt, D. C. (1997). Crystallographic studies of thymidylate synthase: exploring the catalytic mechanism, conformational change, and the role of conserved residues. Department of Biochemistry, The University of Arizona. pp. 175.
- Hyatt, D. C., F. Maley and W. R. Montfort (1997). Use of Strain in a Stereospecific Catalytic Mechanism: Crystal Structures of *Escherichia coli* Thymidylate Synthase Bound to FdUMP and Methylenetetrahydrofolate. Biochemistry **36**: 4585-4594.
- Jackman, A. L., F. T. Boyle and K. R. Harrap (1996). Tomudex (ZD1694): from concept to care, a programme in rational drug discovery. Invest New Drugs **14**(3): 305-16.
- Jackman, A. L. and A. H. Calvert (1995). Folate-based thymidylate synthase inhibitors as anticancer drugs. Annals of Oncology **6**: 871-881.
- Jackman, A. L., P. R. Marsham, R. G. Moran, R. Kimbell, B. M. O'Connor, L. R. Hughes and A. H. Calvert (1991). Thymidylate synthase inhibitors: the in vitro activity of a series of heterocyclic benzoyl ring modified 2-desamino-2-methyl-N10-substituted- 5,8-dideazafolates. Adv Enzyme Regul **31**: 13-27.
- Jancarik, J. and S.-H. Kim (1991). Sparse matrix sampling: a screening method for crystallization of proteins. J. Appl. Cryst. **24**: 409-411.
- Jiang, J. S. and A. T. Brunger (1994). Protein hydration observed by x-ray diffraction: solvation properties of penicillopepsin and neuraminidase crystal structures. J. Mol. Bio. **243**: 100-115.

### REFERENCES - *Continued*

- Jones, A. (1978). A Graphics Model Building and Refinement System for Macromolecules. J. Appl. Crystallogr. **11**: 268-272.
- Jones, T. A., J. Y. Zou, S. W. Cowan and M. Kjeldgaard (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr. **A47**: 110-119.
- Jones, T. R., A. H. Calvert, A. L. Jackman, S. J. Brown, M. Jones and K. R. Harrap (1981). A potent antitumor quinazoline inhibitor of thymidylate synthetase: synthesis, biological properties, and therapeutic results in mice. Eur. J. Cancer **17**: 11-19.
- Kabsch, W. (1976). Acta Cryst. **A32**: 922-923.
- Kabsch, W. (1988). Evaluation of Single-Crystal X-ray Diffraction Data from a Position-Sensitive Detector. J. Appl. Crystallogr. **21**: 916-934.
- Kallen, R. G. and W. P. Jencks (1966). The mechanism of the condensation of formaldehyde with tetrahydrofolic acid. J. Biol. Chem. **241**: 5851-5863.
- Kamb, A., J. S. Finer-Moore and R. M. Stroud (1992). Cofactor Triggers the Conformational Change in Thymidylate Synthase: Implications for an Ordered Binding Mechanism. Biochemistry **31**: 12876-12884.
- Kamen, B. (1997). Folate and antifolate pharmacology. Semin Oncol **24**(5 Suppl 18): S18-30-S18-39.
- Kirkpatrick, D. L., M. L. Jimale, K. M. King and T. Chen (1992). Synthesis and evaluation of imidazolyl disulfides for selective cytotoxicity to hypoxic EMT6 tumor cells *in vitro*. Eur. J. Med. Chem. **27**: 33-37.
- Kirkpatrick, D. L., M. Kuperus, M. Dowdeswell, N. Potier, L. J. Donald, M. Kunkel, M. Berggren, M. Angulo and G. Powis (1998). Mechanisms of Inhibition of the thioredoxin growth factor system by antitumor 2-imidazolyl disulfides. Bioch. Pharm. **55**: 987-994.
- Kitchens, M. E., A. M. Forsthoefel, Z. Rafique, H. T. Spencer and F. G. Berger (1999). Ligand-mediated induction of thymidylate synthase occurs by enzyme stabilization. Implications for autoregulation of translation. J. Biol. Chem. **274**(18): 12544-7.
- Kleywegt, G. J. and T. A. Jones (1994). A super position. ESF/CCP4 Newsletter **31**: 9-14.

### REFERENCES - Continued

- Kleywegt, G. J. and T. A. Jones (1997a). Detecting folding motifs and similarities in protein structures. Meth. Enzymol. **277**: 525-545.
- Kleywegt, G. J. and T. A. Jones (1997b). Model building and refinement practice. Methods in Enzymology. C. W. C. J. a. R. M. Sweet. San Diego, Academic Press. **277**: 208-230.
- Knighton, D. R., C.-C. Kan, E. Howland, C. A. Janson, Z. Hostomska, K. M. Welsh and D. A. Matthews (1994). Structure of and kinetic channelling in bifunctional dihydrofolate reductase-thymidylate synthase. Nature Struct. Biol. **1**(3): 186-194.
- Knowles, J. R. (1987). Tinkering with enzymes: what are we learning? Science **236**: 1252-1258.
- LaPat-Polasko, L., G. F. Maley and F. Maley (1990). Properties of bacteriophage T4 thymidylate synthase following mutagenic changes in the active site and folate binding region. Biochemistry **29**: 9561-9572.
- Laskowski, R. A., M. W. MacArthur, D. S. Moss and J. M. Thornton (1993). PROCHECK: A program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. **26**: 283-291.
- Liu, L. and D. V. Santi (1992). Mutation of asparagine 229 to aspartate in thymidylate synthase converts the enzyme to a deoxycytidylate methylase. Biochemistry **31**(22): 5100-4.
- Lockshin, A. and P. V. Danenberg (1980). Hydrodynamic Behavior of Human and Bacterial Thymidylate Synthetases and Thymidylate Synthetase-5-Fluoro-2'deoxyuridylate-5,10-Methylenetetrahydrofolate Complexes. Evidence for Large Conformational Changes during Catalysis. Biochemistry **19**(118): 4244-4251.
- Lorenson, M. Y., G. F. Maley and F. Maley (1967). The purification and properties of thymidylate synthetase from chick embryo extracts. J. Biol. Chem. **242**: 3332-3344.
- Maley, F., J. Pedersen-Lane and L. Changchien (1995). Complete Restoration of Activity to Inactive Mutants of *Escherichia coli* Thymidylate Synthase: Evidence that *E. coli* Thymidylate Synthase is a Half-the-Sites Activity Enzyme. Biochemistry **34**: 1469-1474.
- Maley, G. F. and F. Maley (1988). Properties of a Defined Mutant of *Escherichia coli* Thymidylate Synthase. J. Biol. Chem. **263**(16): 7620-7627.

### REFERENCES - Continued

- Mathews, C. K. and K. E. Van Holde (1996). Biochemistry. Menlo Park, The Benjamins/Cummings Publishing Co., Inc.
- Matthews, B. W. (1968). Solvent content of protein crystals. J. Mol. Biol. **33**: 491.
- Matthews, D. A., K. Appelt, S. J. Oatley and N. H. Xuong (1990a). Crystal Structure of *Escherichia coli* Thymidylate Synthase Containing Bound 5-fluoro-2'-deoxyuridylate and 10-Propargyl-5,8-dideazafolate. J. Mol. Biol. **214**: 923-936.
- Matthews, D. A., J. E. Villafranca, C. A. Janson, W. W. Smith, K. Welsh and S. Freer (1990b). Stereochemical Mechanism of Action for Thymidylate Synthase Based on the X-ray Structure of the Covalent Inhibitor Ternary Complex with 5-Fluoro-2'-deoxyuridylate and 5,10-Methylenetetrahydrofolate. J. Mol. Biol. **214**: 937-948.
- Matthews, J. R., N. Wakasugi, J.-L. Virelizier, J. Yodoi and R. T. Hay (1992). Thioredoxin regulates the DNA binding activity of NF- $\kappa$ B by reduction of a disulphide bond involving cysteine 62. Nucleic Acids Res. **20**(15): 3821-3830.
- McGuire, J. J. and J. K. Coward (1984). Pteroylpolyglutamates: biosynthesis, degradation, and function. New York, John Wiley & Sons. **1**: 135-190.
- Messerschmidt, A. and J. W. Pflugrath (1987). Crystal Orientation and X-ray Pattern Prediction Routines for Area-Detector Diffractometer Systems in Macromolecular Crystallography. J. Appl. Cryst. **20**: 306-315.
- Michaels, M. L., C. W. Kim, D. A. Matthews and J. H. Miller (1990). *Escherichia coli* thymidylate synthase: Amino acid substitutions by suppression of amber nonsense mutations. Proc. Natl. Acad. Sci. USA **87**: 3957-3961.
- Montfort, W. R., K. M. Perry, E. B. Fauman, J. S. Finer-Moore, G. F. Maley, L. Hardy, F. Maley and R. M. Stroud (1990). Structure, Multiple Site Binding, and Segmental Accommodation in Thymidylate Synthase on Binding dUMP and an Anti-Folate. Biochemistry **29**(30): 6964-6977.
- Montfort, W. R. and A. Weichsel (1997). Thymidylate Synthase: Structure, Inhibition, and Strained Conformations During Catalysis. Pharmacology & Therapeutics **76**(Nos. 1-3): 29-43.
- Oblong, J. E., M. Berggren, P. Y. Gasdaska, S. R. Hill and G. Powis (1995). Site-directed mutagenesis of Lys36 in human thioredoxin: the highly conserved residue affects reduction rates and growth stimulation but is not essential for the redox protein's biochemical or biological properties. Biochemistry **34**(10): 3319-24.

### REFERENCES - Continued

- Oblong, J. E., M. Berggren, P. Y. Gasdaska and G. Powis (1994). Site-directed mutagenesis of active site cysteines in human thioredoxin produces competitive inhibitors of human thioredoxin reductase and elimination of mitogenic properties of thioredoxin. J. Biol. Chem. **269**: 11714-11720.
- Orengo, C. A., A. D. Michie, S. Jones, D. T. Jones, M. B. Swindells and J. M. Thornton (1997). CATH- A Hierarchic Classification of Protein Domain Structures. Structure **5**(8): 1093-1108.
- Otwinowski, Z. (1993). Oscillation data reduction program. Proceedings of the CCP4 Study weekend: Data collection and processing, SERC Daresbury laboratory, England.
- Poe, M., L. M. Jackman and S. J. Benkovic (1979). 5,10-Methylene-5,6,7,8-tetrahydrofolate. Conformation of the Tetrahydropyrazine and Imidazolidine Rings. Biochemistry **18**: 5527-5530.
- Pogolotti, A. L., Jr., P. V. Danenberg and D. V. Santi (1986). Kinetics and Mechanism of Interaction of 10-Propargyl-5,8-dideazafolate with Thymidylate Synthase. J. Med. Chem. **29**: 478-482.
- Pogolotti, A. L., Jr., C. Weill and D. V. Santi (1979). Thymidylate Synthetase Catalyzed Exchange of Tritium from [5-<sup>3</sup>H]-2'-Deoxyuridylate for Protons of Water. Biochemistry **18**(13): 2794-2798.
- Powis, G., J. E. Oblong, P. Y. Gasdaska, M. Berggren, S. Hill and D. L. Kirkpatrick (1994). The thioredoxin/thioredoxin reductase redox system and control of cell growth. Oncol. Res. **6**: 539-544.
- Qin, J., G. M. Core and A. M. Gronenborn (1996). Ionization equilibria for side-chain carboxyl groups in oxidized and reduced human thioredoxin and in the complex with its target peptide from the transcription factor NFkB. Biochemistry **35**(1): 7-13.
- Rubartelli, A., A. Bajetto, G. Allavena, E. Wollman and R. Sitia (1992). Secretion of thioredoxin by normal and neoplastic cells through a leaderless secretory pathway. J. Biol. Chem. **267**(34): 24161-24164.
- Rutenber, E. and R. Stroud (1996). Binding of the anticancer drug ZD1694 to E. coli thymidylate synthase: assessing specificity and affinity. Structure **4**: 1317-1324.
- Samsonoff, W., J. Reston, M. McKee, B. O'Connor, J. Galivan, G. Maley and F. Maley (1997). Intracellular location of thymidylate synthase and its state of phosphorylation. J. Biol. Chem. **272**: 13281-13285.

### REFERENCES - *Continued*

- Santi, D. V., C. S. McHenry and H. Sommer (1974). Mechanism of interaction of thymidylate synthase with 5-fluorodexoyuridylate. Biochemistry **13**: 471-481.
- Santi, D. V., K. Pinter, J. Kealey and V. J. Davisson (1990). Site-directed Mutagenesis of Arginine 179 of Thymidylate Synthase. A Nonessential Substrate-Binding Residue. J. Biol. Chem. **265**: 6770-6775.
- Schiffer, C. A., I. J. Clifton, V. J. Davisson, D. V. Santi and R. M. Stroud (1995). Crystal Structure of Human Thymidylate Synthase: A Structural Mechanism for Guiding Substrates into the Active Site. Biochemistry **34**: 16279-16287.
- Shoichet, B. K., R. M. Stroud, D. V. Santi, I. D. Kuntz and K. M. Perry (1993). Structure-Based Discovery of Inhibitors of Thymidylate Synthase. Science **259**: 1445-1450.
- Smith, G. K., H. Amyx, C. M. Boytos, D. S. Duch, R. Ferone and H. R. Wilson (1995). Enhanced Antitumor Activity for the Thymidylate Synthase Inhibitor 1843U89 through Decreased Host Toxicity with Oral Folic Acid. Cancer Research **55**: 6117-6125.
- Smith, I., A. Jones, M. Spielmann, M. Namer, M. D. Green, J. Bonnetterre, H. E. Wander, T. Hatschek, N. Wilking, J. Zalcborg, J. Spiers and L. Seymour (1996). A phase II study in advanced breast cancer: ZD1694 (Tomudex) a novel direct and specific thymidylate synthase inhibitor. British Journal of Cancer **74**(3): 479-481.
- Sotelo-Mundo, R. R., J. Ciesla, J. M. Dzik, W. Rode, F. Maley, G. F. Maley, L. W. Hardy and W. R. Montfort (1999). Crystal Structures of Rat Thymidylate Synthase Inhibited by Tomudex, a Potent Anticancer Drug. Biochemistry **38**: 1087-1094.
- Stout, T. J., U. Schellenberger, D. V. Santi and R. M. Stroud (1998). Crystal structure of a unique thermal-stable thymidylate synthase from *Bacillus subtilis*. Biochemistry **37**: 14736-14747.
- Stout, T. J. and R. M. Stroud (1996). The complex of the anti-cancer therapeutic, BW1843U89, with thymidylate synthase at 2.0 Å resolution: implications for a new mode of inhibition. Structure **4**: 67-77.
- Stout, T. J., D. Tondi, M. Rinaldi, D. Barlocco, P. Pecorari, D. V. Santi, I. D. Kuntz, R. M. Stroud, B. K. Shoichet and M. P. Costi (1999). Structure-based design of inhibitors specific for bacterial thymidylate synthase. Biochemistry **38**(5): 1607-17.



### REFERENCES - Continued

- Strop, P., L. Chanchien, F. Maley and W. R. Montfort (1997). Crystal structures of a marginally active thymidylate synthase mutant, Arg 126 -> Glu. Protein Science **6**: 2504-11.
- Takimoto, C. H. (1997). Antifolates in clinical development. Seminars in Oncology **24 Supp 8(5)**: S18-40 - S18-51.
- Tong, Y., X. Liu-Chen, E. A. Ercikan-Abali, G. M. Capiiaux, S. C. Zhao, D. Banerjee and J. R. Bertino (1998a). Isolation and characterization of thymitaq (AG337) and 5-fluoro-2-deoxyuridylate-resistant mutants of human thymidylate synthase from ethyl methanesulfonate-exposed human sarcoma HT1080 cells. J Biol Chem **273**(19): 11611-8.
- Tong, Y., X. Liu-Chen, E. A. Ercikan-Abali, S. C. Zhao, D. Banerjee, F. Maley and J. R. Bertino (1998b). Probing the folate-binding site of human thymidylate synthase by site-directed mutagenesis. Generation of mutants that confer resistance to raltitrexed, Thymitaq, and BW1843U89. J Biol Chem **273**(47): 31209-14.
- veer Reddy, G. and A. Pardee (1983). Inhibitor evidence for allosteric interaction in the replitase multienzyme complex. Nature **304**: 86-88.
- Verlinde, C. L. and W. G. Hol (1994). Structure-based drug design: progress, results and challenges. Structure **2**: 577-587.
- Wakasugi, N., Y. Tagaya, A. Wakasugi, M. Mitsui, M. Maeda, J. Yodoi and T. T. (1990). Adult T-cell leukemia-derived factor/Thioredoxin produced by both human T-lymphotropic virus type 1 and Epstein-Barr virus-transformed lymphocytes, acts as an autocrine growth factor and synergized with interleukin-1 and interleukin-2. Proc. Natl. Acad. Sci. USA **87**: 8282-8286.
- Wataya, Y., A. Matsuda and D. V. Santi (1980). Interaction of Thymidylate Synthetase with 5-Nitro-2'-deoxyuridylate. J. Biol. Chem. **255**: 5538-5544.
- Wataya, Y., K. Negishi and H. Hayatsu (1973). Debromination of 5-Bromo-2'-deoxyuridine by Cysteine. Formation of Deoxyuridine and S-[5-(2'-Deoxyuridyl)]cysteine. Biochemistry **12**(20): 3992-3998.
- Webber, S., C. A. Bartlett, T. J. Boritzki, J. A. Hillard, E. F. Howland, A. L. Johnston, M. Kosa, S. A. Margosiak, C. A. Morse and B. V. Shetty (1996). AG337, a novel lipophilic thymidylate synthase inhibitor: in vitro and in vivo preclinical studies. Cancer Chemotherapy & Pharmacology **37**(6): 509-517.

**REFERENCES - *Continued***

Weichsel, A., J. R. Gasdaska, G. Powis and W. R. Montfort (1996). Crystal structures of reduced, oxidized, and mutated human thioredoxins: evidence for a regulatory homodimer. Structure **4**(6): 735-751.

Weichsel, A. and W. R. Montfort (1995). Ligand-Induced Distortion of an Active Site in Thymidylate Synthase Upon Binding Anticancer Drug 1843U89. Nature Structural Biology **2**(12): 1095-1101.