

**NANYANG**  
**TECHNOLOGICAL**  
**UNIVERSITY**

**Fragment-based Drug Discovery  
and Mechanistic Studies of Human  
Thymidylate Synthase**

Chen Dan

School of Biological Sciences

A thesis submitted to the Nanyang Technological University in  
fulfillment of the requirement for the degree of  
Doctor of Philosophy

**2015**

## ACKNOWLEDGMENTS

The writing of this thesis has been an amazing journey. It has been a feast of experiences, and also a time of fantastic intellectual growth, fostered by a wonderful learning environment at NTU. This thesis is the result of enormous experiences I have encountered from my supervisor, colleagues, friends and family.

Foremost, I would like to express my sincere gratitude to my supervisor Prof. Pär Nordlund. Pär has been supportive since I began my final year project in his group as an undergraduate. Thanks to Pär, not only for giving me the opportunity to continue my graduate training in drug discovery, but also for providing the freedom to wander intellectually and his invaluable guidance and encouragement when I was dumped about the direction of my research.

In addition, I am especially grateful for Dr. Andreas Larsson and Dr. Anna Jansson, whose steadfast support and intellectual guidance have been invaluable over these years. Our meetings were not only important for the development of my research, they were also essential in keeping me inspired and, at times, helping me regain the confidence that it would all come to a good end (or even “an end”).

The members (the former and present) of the Nordlund group have contributed immensely to my personal and academic time at NTU. They have been a source of friendship as well as good advices. I am grateful for the group who strived in graduate school with me: Kelly, Kit, Rachel, Vidhi and Lekshmy. A deep thank you to our FBDD team members for all the discussions, advices and inspirations during our weekly meetings: Dr. Christofer Björkelid, Dr. Kim Young-Mee, Stephanie and Saranya. I am very thankful to Dr. Teo Hsiang-Ling and Dr. Sue-Li Dahlroth, for sharing of their expertise. My appreciation goes to all other group members as well. It has been a great pleasure and precious experience to work with or alongside of you: Dr. Lim Yan-Ting, Dr. Radoslaw Mikolaj Sobota, Dr. Dai Ling-Yun and Ka-Diam. In addition, a special thank you to Jeslin, for taking care of our administration stuff and processing our invoices.

I thank protein production platform (PPP), which was previously led by Dr. Tobias Carl Cornvik and now by Dr. Anna Jansson. PPP has been playing a critical role in the cloning work and initial scale-up of all the protein constructs. Projects would not be moving forward so fast without the help from PPP. I would like to

express my sincere appreciation to all PPP members (the former and present): Martina, Agnes, Ying-Hua, Lucy, Bee-Ting, Hai-Ting and Daniel.

My acknowledgement also goes to the beamline sources we used for crystallographic studies: Australian Synchrotron and National Synchrotron Radiation Research Center (Taiwan). Thanks to the staff for all the help and professional advice for data collection.

Last but not the least, a special thanks to my family: my parents and my brothers Jun and Kai. My family has been a source of constant and unconditional love throughout my life, which makes all the difference. Thank you for the myriad of ways in which you supported in all my pursuits to find and realize my potential, and to make this contribution to our world.

## **DECLARATION**

The work presented (including data generated and data analysis) was carried out by the author except in the cases outlined below:

- The data in Figure 3.27 was generated by collaborators and published
- The screening data in Figure 3.29 was generated by collaborators and manuscript on it was submitted for review

## TABLE OF CONTENTS

<b>ACKNOWLEDGMENTS.....</b>	<b>2</b>
<b>DECLARATION .....</b>	<b>4</b>
<b>ABBREVIATIONS .....</b>	<b>7</b>
<b>LIST OF FIGURES .....</b>	<b>10</b>
<b>LIST OF TABLES .....</b>	<b>12</b>
<b>ABSTRACT .....</b>	<b>13</b>
<b>CHAPTER 1 INTRODUCTION.....</b>	<b>14</b>
<b>1.1 Introduction to fragment-based drug discovery .....</b>	<b>14</b>
1.1.1 Evolution of drug discovery .....	14
1.1.2 Fragment-based drug discovery.....	15
<b>1.2 Pharmacology of cancer .....</b>	<b>19</b>
1.2.1 Cancer: an introductory overview.....	20
1.2.2 Cancer chemotherapy.....	23
<b>1.3 Chemotherapy targeting nucleotide synthesis.....</b>	<b>28</b>
1.3.1 The roles of nucleotides .....	28
1.3.2 Nucleotide <i>de novo</i> biosynthesis .....	30
1.3.3 Regulation of nucleotide biosynthesis.....	34
1.3.4 Nucleotide degradation and salvage.....	38
1.3.5 Pharmacological inhibition of nucleotide synthesis .....	41
<b>1.4 Human thymidylate synthase: a critical target in cancer chemotherapy.....</b>	<b>43</b>
1.4.1 Overview of thymidylate synthase .....	43
1.4.2 Structure and catalytic mechanism of TS.....	46
1.4.3 Targeting hTS in cancer chemotherapy .....	52
<b>1.5 Aims of the project .....</b>	<b>59</b>
<b>CHAPTER 2 MATERIALS AND METHODS .....</b>	<b>60</b>
<b>2.1 Fragment library.....</b>	<b>60</b>
<b>2.2 Molecular cloning and expression.....</b>	<b>60</b>
<b>2.3 Mutagenesis.....</b>	<b>60</b>
<b>2.4 Protein purification .....</b>	<b>61</b>
<b>2.5 Size-exclusion chromatography (SEC) coupled with multi-angle light scattering (MALS).....</b>	<b>62</b>
<b>2.6 Enzymatic assay .....</b>	<b>62</b>
<b>2.7 Differential scanning fluorimetry (DSF).....</b>	<b>62</b>
<b>2.8 Differential static light scattering (DSLS).....</b>	<b>63</b>
<b>2.9 Isothermal titration calorimetry (ITC).....</b>	<b>63</b>
<b>2.10 Crystallization .....</b>	<b>63</b>
<b>2.11 Crystal soaking and freezing.....</b>	<b>64</b>
<b>2.12 Data collection and structure determination.....</b>	<b>64</b>
<b>2.13 Ribonucleic acids (RNAs) .....</b>	<b>65</b>
<b>2.14 Native PAGE of protein-RNA complexes.....</b>	<b>65</b>
<b>2.15 Panning of the phage display library .....</b>	<b>65</b>
<b>2.16 Phage enzyme-linked immunosorbent assay (ELISA) .....</b>	<b>66</b>
<b>2.17 Co-Immunoprecipitation (CoIP) of VH-hTS complexes.....</b>	<b>67</b>
<b>2.18 Pull-down assay of VH-hTS complexes .....</b>	<b>67</b>
<b>2.19 Cellular thermal shift assay (CETSA) .....</b>	<b>68</b>
<b>2.20 Compound screening on hTS using CETSA.....</b>	<b>69</b>
<b>2.21 Surface plasmon resonance (SPR) .....</b>	<b>69</b>
<b>2.22 Computational tools .....</b>	<b>70</b>

2.23 Compound synthesis .....	70
<b>CHAPTER 3 RESULTS .....</b>	<b>71</b>
3.1 Characterization of hTS protein .....	71
3.1.1 Protein purification of hTS.....	71
3.1.2 Biophysical characterization of hTS .....	71
3.2 Crystal structures of hTS.....	77
3.2.1 Structure of full-length hTS.....	77
3.2.2 Structure of N-terminal truncated hTS.....	78
3.2.3 Conformational changes upon dUMP and antifolate binding .....	81
3.3 Fragment-based drug discovery targeting hTS .....	89
3.3.1 Fragment screening using DSF.....	89
3.3.2 Crystallographic screening of fragment cocktails .....	90
3.3.4 Structure-based drug design.....	94
3.4 Characterization of inactive conformers of hTS.....	109
3.4.1 Crystal structures of mutant hTS in the inactive conformation .....	109
3.4.2 Biophysical characterization of mutant hTS .....	115
3.4.3 Evaluation of type IV hits in mutant hTS.....	117
3.4.4 Investigation of RNA binding to hTS.....	118
3.4.5 Identification of hTS binders from VH phage-display library .....	121
3.5 Collaborative work.....	123
3.5.1 Cellular thermal shift assay (CETSA).....	123
3.5.2 Characterization of hits from CETSA based screening .....	124
<b>CHAPTER 4 DISCUSSION.....</b>	<b>127</b>
4.1 DSF as an informative characterization technique for hTS.....	127
4.2 Implications of hTS structures.....	129
4.3 Fragment screening and hit translation into crystal structure.....	130
4.4 Novel starting points for hTS inhibitor development .....	132
4.5 Fragment merging strategy .....	134
4.6 Implications of the binding mode of A37.....	134
4.7 Insights into the transition between active and inactive states of hTS.....	136
4.8 Potential allosteric inhibitors for hTS .....	138
<b>CHAPTER 5 CONCLUSION AND FUTURE PERSPECTIVE .....</b>	<b>139</b>
<b>REFERENCES .....</b>	<b>142</b>
<b>APPENDIX .....</b>	<b>157</b>

## ABBREVIATIONS

5-FU	5-fluorouracil
5,10-CH <sub>2</sub> -THF	5,10-methylenetetrahydrofolate
ALL	Acute lymphoblastic leukemia
AMP	Adenosine monophosphate
SASA	Solvent accessible surface area
ASU	Asymmetric unit
ATCase	Aspartate transcarbamoylase
CDR	Complementarity determining regions
CETSA	Cellular thermal shift assay
CML	Chronic myeloid leukemia
CoIP	Co-immunoprecipitation
CPSaseII	Carbamoylphosphate synthetase II
CV	Column volume
DHF	Dihydrofolate
DHFR	Dihydrofolate reductase
DHODH	Dihydroorotate dehydrogenase
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleoside triphosphate
DSF	Differential scanning fluorimetry
DSL	Differential static light scattering
dTMP	Deoxythymidine monophosphate
dUMP	Deoxyuridine monophosphate
EDTA	Ethylenediaminetetraacetic acid
ELISA	Enzyme-linked immunosorbent assay
FBDD	Fragment-based drug discovery
FdUMP	Fluorodeoxyuridine monophosphate
FPGS	Folypolyglutamate synthetase
GMP	Guanosine monophosphate
HHV3	Human herpes virus 3
HRP	Horseradish peroxidase
HTS	High throughput screening

IMAC	Immobilized metal affinity chromatography
IMP	Inosine monophosphate
IPTG	Isopropyl- $\beta$ -D-1-thiogalactopyranoside
ITC	Isothermal titration calorimetry
ITDRF	Isothermal dose response fingerprint
LC-MS	Liquid chromatography–mass spectrometry
MALS	Multi-angle light scattering
ODCase	Orotidine decarboxylase
OPRTase	Orotate phosphoribosyltransferase
PABA	<i>para</i> -aminobenzoic acid
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate-buffered saline
PCFT	Proton-coupled folate transporter
PCR	Polymerase chain reaction
PDB	Protein Data Bank
PEG	Polyethylene glycol
P <sub>i</sub>	Phosphate ion
PRPP	Ribose-5-phosphoribosepyrophosphate
RFC	Reduced folate carrier
RMSD	Root-mean-square deviation
RNA	Ribonucleic acid
RNR	Ribonucleotide reductase
SAR	Structure-activity relationship
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
SHMT	Serine hydroxymethyl transferase
SPR	Surface plasmon resonance
T <sub>agg</sub>	Aggregation temperature
TBS	Tris-buffered saline
TBSM	TBS containing 4% skim milk
TBST	TBS with 0.05% Tween 20
TCEP	Tris-(2-carboxyethyl)phosphine

THF Tetrahydrofolate  
T<sub>m</sub> Melting temperature  
TMB 3,3',5,5'-Tetramethylbenzidine  
TS Thymidylate synthase  
TSA Thermal shift assay  
VH Single-heavy chain variable-fragment antibody

## LIST OF FIGURES

Figure 1.1 Fragment-based drug discovery .....	18
Figure 1.2 Cancer incidence and mortality in 2012 .....	21
Figure 1.3 Nucleotide chemical structures.....	29
Figure 1.4 <i>De novo</i> synthesis of purine nucleotides .....	31
Figure 1.5 <i>De novo</i> synthesis of pyrimidine nucleotides .....	33
Figure 1.6 Biosynthesis, interconversion and degradation of pyrimidine nucleotides .....	40
Figure 1.7 Function of TS and regeneration of folates .....	44
Figure 1.8 Protein sequence alignment of TS from different organisms.....	45
Figure 1.9 Crystal structure of <i>E. coli</i> TS complexed with dUMP .....	47
Figure 1.10 Crystal structure of <i>E. coli</i> TS complexed with dUMP and THF .....	49
Figure 1.11 Chemical mechanism of the reaction catalyzed by TS .....	51
Figure 1.12 Historical development in targeting hTS in cancer chemotherapy .....	54
Figure 1.13 Mechanism of action of fluoropyrimidines and antifolates .....	56
Figure 3.1 Different hTS protein constructs and SEC-MALS data .....	73
Figure 3.2 Evaluation of hTS constructs using DSF.....	74
Figure 3.3 Binding of dUMP and raltitrexed .....	76
Figure 3.4 Crystal structure of ligand-free full-length hTS (vc1) .....	77
Figure 3.5 Crystal structures of N-terminal truncated hTS (vc5) .....	79
Figure 3.6 Ligand-free and dUMP-bound crystal structures of vc5 in crystal form 3 .....	82
Figure 3.7 Alignment of the binary and ternary complex structures .....	83
Figure 3.8 Binding of antifolates in the active site .....	84
Figure 3.9 Water molecules in the back pocket near Glu87 .....	86
Figure 3.10 DSF-based fragment screening data .....	90
Figure 3.11 Fragment cocktail 1 and difference Fourier peak in the crystal structure .....	91
Figure 3.12 Four types of fragment hits identified in the crystal structures .....	93
Figure 3.13 Binding of type I fragments .....	94
Figure 3.14 Crystal structures of hTS complexed with compound <b>A5</b> and <b>A8</b> .....	97
Figure 3.15 Crystal structures of hTS in complex with <b>A11</b> , <b>A15</b> and <b>A27</b> .....	100
Figure 3.16 Crystal structure of hTS in complex with <b>A33</b> .....	101
Figure 3.17 Crystal structures of hTS in complex with <b>A36</b> and <b>A37</b> .....	102

Figure 3.18 Enzymatic activity inhibited by follow-up compounds <b>A33</b> , <b>A34</b> and <b>A36</b> .....	103
Figure 3.19 Crystal structures hTS in complex with compounds of type II binding mode .....	107
Figure 3.20 Crystal structures of A191K-sf and A191K-s .....	110
Figure 3.21 Structural comparison of the active and inactive conformations of hTS .....	113
Figure 3.22 Biophysical characterization of M190K and A191K .....	116
Figure 3.23 Crystal structures of the dimer interface binders in the active and inactive states of hTS .....	118
Figure 3.24 DSF data of F13 on M190K and A191K .....	119
Figure 3.25 Interaction of RNA with the active and inactive states of hTS .....	120
Figure 3.26 Interaction assays and crystal structures of VH binders with hTS .....	122
Figure 3.27 CETSA experiments on hTS .....	123
Figure 3.28 CETSA data of compound <b>A27</b> on hTS .....	124
Figure 3.29 Pilot screen using CETSA-HTS to measure target engagement of hTS .....	125
Figure 3.30 Biophysical data for CBK115334 .....	126
Figure 4.1 Model 1 of the biphasic melting profile of hTS on DSF .....	128
Figure 4.2 Good correlation observed between DSF data and $K_D$ values from ITC .. .....	131
Figure 4.3 Potential vectors of expansion on <b>A36</b> .....	133
Figure 4.4 Structural differences between hTS and <i>P. falciparum</i> DHFR-TS .....	135
Figure 4.5 Electrostatic potential surface of the active and inactive states of hTS ..	137

## LIST OF TABLES

Table 1.1 Selected clinical-stage compounds originated from FBDD .....	19
Table 1.2 Examples of chemotherapy agents targeting nucleotide synthesis .....	42
Table 3.1 Data collection and refinement statistics of different crystal forms .....	80
Table 3.2 Data collection and refinement statistics of binary and ternary complex structures .....	88
Table 3.3 Biophysical data of compounds <b>A1-A8</b> .....	96
Table 3.4 Biophysical data of compounds <b>A9-A33</b> .....	98
Table 3.5 Biophysical data of compounds <b>A34-A37</b> .....	104
Table 3.6 Biophysical data of compounds <b>B1-B10</b> .....	108
Table 3.7 Data collection and refinement statistics of structures of A191K and M190K .....	114

## **ABSTRACT**

Human thymidylate synthase (hTS) is the sole enzyme responsible for *de novo* biosynthesis of dTMP and is vital to cell proliferation and survival. Inhibition of hTS has been explored and found to be effective in cancer chemotherapy, while the efficacy is limited by drug resistance. Herein fragment screening is applied to hTS to identify novel start point for drug design. Four types of hits are found, of which two are progressed to structure- and biophysics-driven drug design due to novel binding modes in the active site. The elaborated compounds show significant improvements in affinities and open new avenues for development of hTS inhibitors to overcome drug resistance. Meanwhile, another type of hit reveals a potential allosteric site of hTS possibly involved in regulating overexpression-induced drug resistance to hTS inhibitors.

## **CHAPTER 1 INTRODUCTION**

### **1.1 Introduction to fragment-based drug discovery**

#### **1.1.1 Evolution of drug discovery**

In pharmacology, drugs are chemical substances that interact with specific molecular components of an organism to mediate biochemical and physiological effects on that organism. Drug discovery and development dates back to thousands of years in the history of human civilization. In these ancient times, medicines were derived mainly from plants and supplemented with animal materials and minerals (Castiglioni et al., 1947). In this first era of drug discovery, the medicines were likely discovered by trial and error without prior knowledge of the ingredients. It was in the renaissance period that many advances were achieved in medicinal preparation and medical treatment which laid the foundation of today's drug discovery and development (Raviña Rubira, 2011). Drug research started to follow scientific principles in the late 1800s. The development of organic chemistry and chemical analysis opened the door toward the isolation and characterization of bioactive chemicals for plant materials. This led to the introduction of concepts such as receptor-ligand interactions and paved the way for rational drug design (Rang, 2006). After World War I, the modern pharmaceutical industry came into existence. The advent of synthetic organic chemistry significantly displaced the plant-based medicinal chemistry and ushered in a new era of drug discovery and development. Synthetic drugs using chemical methods started to attract widespread interest. From 1930s, drug discovery concentrated on screening of natural products, which was the origin of today's high throughput screening (HTS) of millions of synthetic compounds (Pereira and Williams, 2007). The development of cell models contributed immensely to understanding the interaction between drugs and drug targets. In parallel with the continued development of screening techniques and medicinal chemistry, the advancements made in genomics and proteomics facilitated drug targets identification for specific diseases. From 1980s, HTS techniques made it possible to screen a large number of compounds. With improved techniques for protein expression and purification, protein nuclear magnetic resonance (NMR) and X-ray crystallography became increasingly accessible. The maturation of these

techniques allowed researchers at Abbott to include structural information already in the 1990s and use it as a main lead optimization metric (Shuker et al., 1996). Structure-activity relationship (SAR) explores the relationship between a molecule's biological activity and its chemical structure. The concept of SAR is in many ways the seed to fragment-based drug discovery (FBDD) extensively used today in drug development. Since then, there have been many examples of using NMR and/or X-ray crystallography to determine the binding mode of small-molecule ligands and expand them into potent and specific compounds using rational drug design strategies (Klebe, 2000; Rees et al., 2004). Nowadays, FBDD has become an alternative approach to HTS in many pharmaceutical companies and academic groups (Hajduk and Greer, 2007). The processes for HTS and FBDD based on purified proteins can be quite similar, including the use of protein-ligand structures to understand SAR and to guide optimization of the binding affinity, selectivity and drug-like properties of a compound for a particular drug target site. However, the start point of the FBDD process with small and physiochemical suitable fragment molecules potentially makes this process more efficient.

### **1.1.2 Fragment-based drug discovery**

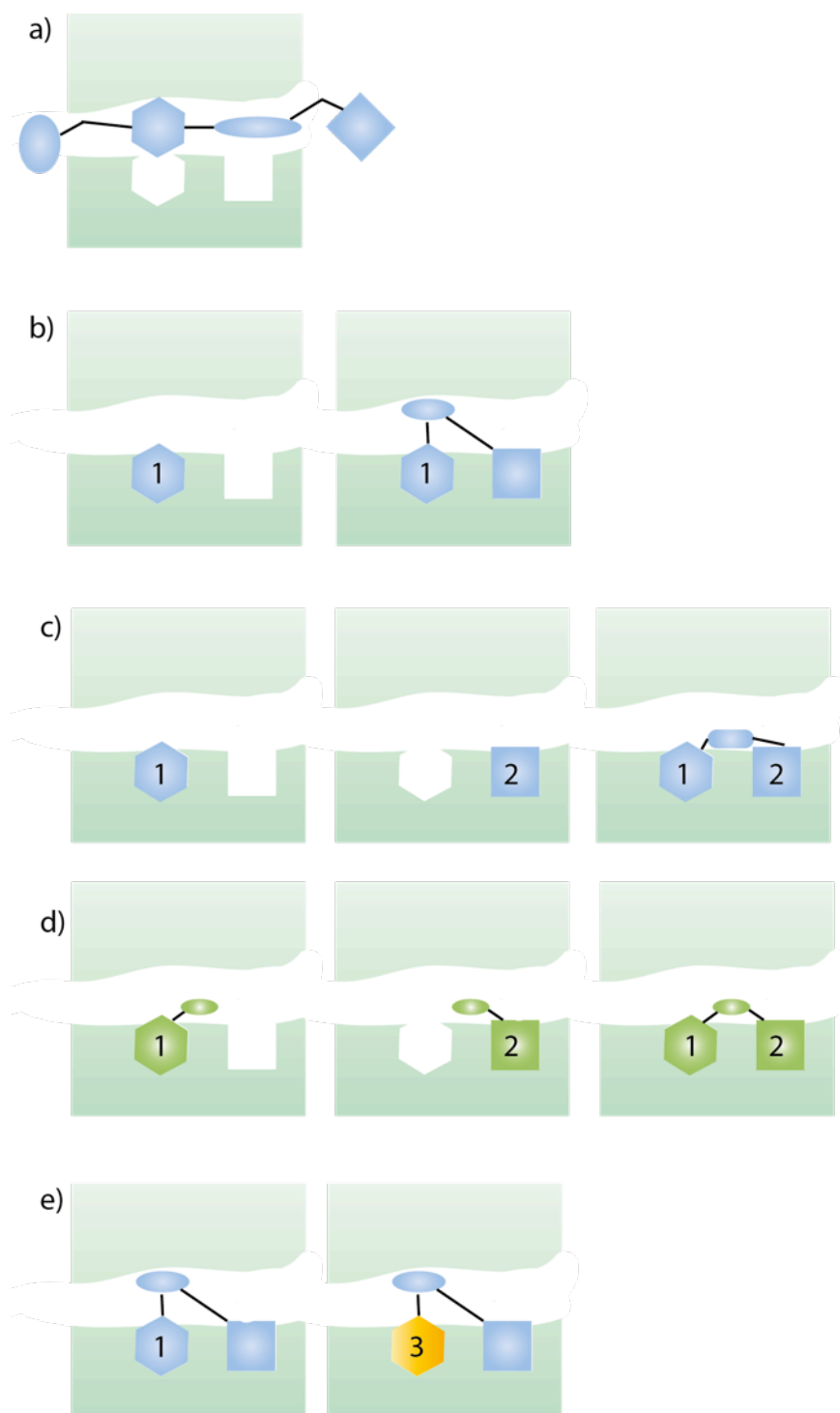
The central theme of FBDD is to screen a library of low molecular weight (LMW) compounds and subsequently apply an iterative cycle of structure-based drug design to obtain desired drug-like properties including potency, selectivity and appropriate physiochemical profile. The concept of using fragments was developed by William Jencks in the early 1980s, who proposed that the affinity of a molecule could be understood as a function of the affinities of separate components (Jencks, 1981). This fragment approach is gaining wide acceptance and one drug originated from FBDD has been approved after six years' efforts at Plexxikon (Tsai et al., 2008). Supporting the efficiency of FBDD, there are also many drug candidates in different stages of clinical trials that have been developed with this strategy (Table. 1.1). The approach of starting with small has become big. Given the substantial advances achieved in the development of combinatorial chemistry and high throughput capabilities, one question is why one instead would pursue FBDD to start with small and weakly binding molecules. The answer is that FBDD provides an avenue to help in compensating for our incomplete understanding of biological interactions and

offers an alternative, and partly complementary, route to identifying chemical starting points for drug development:

- Fragment libraries can be much smaller than HTS libraries because screening small molecules can sample a greater portion of chemical universe. It has been demonstrated that the number of possible chemical structures increases exponentially with the molecular size. The estimated number of possible drug-sized molecules (containing 30 non-hydrogen atoms) is up to  $10^{63}$  (Bohacek et al., 1996), making even a library of  $10^6$  compounds (a fairly large HTS library) infinitesimally small. While the molecular size or number of heavy atoms decreases in the starting compound/fragment, the number of theoretical chemical structures that can be accessed plummets. Fragments are typically defined as compounds with a molecular weight of 250 Da or less, although criteria vary. Based on the analysis, the sampling ability of fragments is considerably better than HTS drug-sized compounds. The typical size of fragment library used in FBDD is 500 to 2000 compounds. This small size of the library has advantages associated with library construction, handling, storage, as well as for screening.
- The hit rates of FBDD are generally higher than HTS due to higher probability of forming a match with the targeted protein as well as the sampling advantages as discussed above. It has been theoretically formulated (Hann et al., 2001) and experimentally validated (Schuffenhauer et al., 2005) that molecules with larger size and higher complexity have a greater chance of forming mismatches with the target protein (Figure 1.1a). The smaller, less complex nature of fragments reduces the probability of forming suboptimal interactions or clashes in the binding site. Better sampling of the chemical space coupled with higher hit rates of less complex molecules makes fragment screening a powerful strategy for identifying novel starting points for drug development.
- Fragment hits provide a greater scope for chemical evolution when following medicinal chemistry and also an opportunity to tackle the attrition problem. Chemical evolution in FBDD is performed using the strategy of structure-based drug design, an iterative cycle of structure-guided optimization, compound synthesis and evaluation (Bohacek et al., 1996). The process starts

with a detailed analysis of three-dimensional structures of protein-fragment complexes. The structural information is used to evolve the fragments into high affinity leads by incorporating unique interactions in the binding site, either by expanding with additional binding elements (fragment growing, Figure 1.1b) or by joining fragments together (fragment merging or linking, Figure 1.1c and d). Detailed structural insights of a series of gradually optimized compounds in the binding site facilitate the establishment of SAR within the series and also the elucidation of some modulating factors for drug-like properties as well as potency. The chemical traceability during fragment evolution makes it possible to trace the attrition back to the controlling factors and tackle the problem (Figure 1.1e).

- Fragment screening requires more sensitive detection techniques, as fragments are often of low affinity. However, at the same time these techniques provide invaluable information to grow fragment hits to drug candidates. The development of FBDD has prompted continued efforts made to push the limit of sensitivity and automation of various methods so that they are applicable to high throughput screening of fragments. Even though fragment screening can be driven by biochemical, biophysical or computational methods, there is no case of a successful fragment screening campaign without a biophysical screening component. A wide range of biophysical methods are used in fragment screening, including X-ray crystallography, NMR, surface plasmon resonance (SPR), thermal shift assays (TSA) and mass spectrometry (MS) (Scott et al., 2012). As every method has its advantages and limitations, it is the judicious combination of orthogonal methods that can lead to the most successful strategy in FBDD programs.



**Figure 1.1. Fragment-based drug discovery.** a) Schematic representation of a HTS hit with mismatch to the binding pocket. b) Fragment growing. Fragment 1 binds at one site of the target. Fragment 1 is built to make contact with the top surface and expanded to the second binding site. c) Fragment linking. Fragment 1 and fragment 2 bind to adjacent sites of the target, which can be joined by a linking group to yield compounds spanning both sites. d) Fragment merging. Fragment 1 and 2 overlap in their binding poses and can be merged together. e) Fragment optimization. Lead molecules can be re-engineered to optimize a particular property in a more independent and predictable manner compared to HTS.

**Table 1.1.1. Selected clinical-stage compounds originated from FBDD**

<b>Drug; status</b>	<b>Targets (indications)</b>	<b>Originator</b>	<b>Detection and development</b>	<b>MW (Da) (Initial; final)</b>	<b>cLogP</b>	<b>References</b>
Vemurafenib (PLX4032); Approved	Mutant BRAF kinase (metastatic melanoma)	Plexxikon	Functional assays of 20,000 scaffolds plus crystallography-informed design	209; 490	4.2	Proceedings of the National Academy of Sciences 105, 3041–3046 (2008)
MK-8931; Phase III	$\beta$ -secretase 1 (Alzheimer's disease)	Merck	NMR screening of 10,000 compounds followed by crystallography-informed design	186; ND	ND	Journal of Medicinal Chemistry 53, 942–950 (2010)
Navitoclax (ABT-263); Phase II	BCL-2/BCL-XL (Leukemia)	Abbott	Protein-detected NMR screen of 17,000 compounds optimized with SAR by NMR	216*, 170*; 975	>5	Journal of Medicinal Chemistry 51, 6902–6915 (2008)
AT13387; Phase II	HSP90 (gastrointestinal stromal tumors)	Astex	Ligand-detected NMR screen of 1,600 compounds followed by 125 crystallography screens and optimization	223; 409	3.5	Journal of Medicinal Chemistry 53, 5956–5969 (2010)
AT7519; Phase II	Cyclin-dependent kinases (multiple myeloma)	Astex	Crystallography screening of 500 fragments followed by optimization	118; 382	0.3	Journal of Medicinal Chemistry 51, 4986–4999 (2008)
AZD5363; Phase II	Protein kinase B (multiple solid tumors)	AstraZeneca/ Astex/CR-UK	Virtual screen of about 300,000 compounds optimized with crystallography-informed design	118; 428	1.31	Journal of Medicinal Chemistry 56, 2059-2073 (2013)

Table is adapted from (Baker, 2013). \*Molecular mass of two initial hits that were optimized and linked. Abbreviations: BCL-2, B cell lymphoma 2; cLogP, calculated logarithm of the 1-octanol–water partition coefficient of the non-ionized molecule; HSP90, heat shock protein 90; ND, not disclosed; SAR, structure–activity relationship.

## **1.2 Pharmacology of cancer**

### **1.2.1 Cancer: an introductory overview**

#### **1.2.1.1 Cancer incidence, prevalence and mortality**

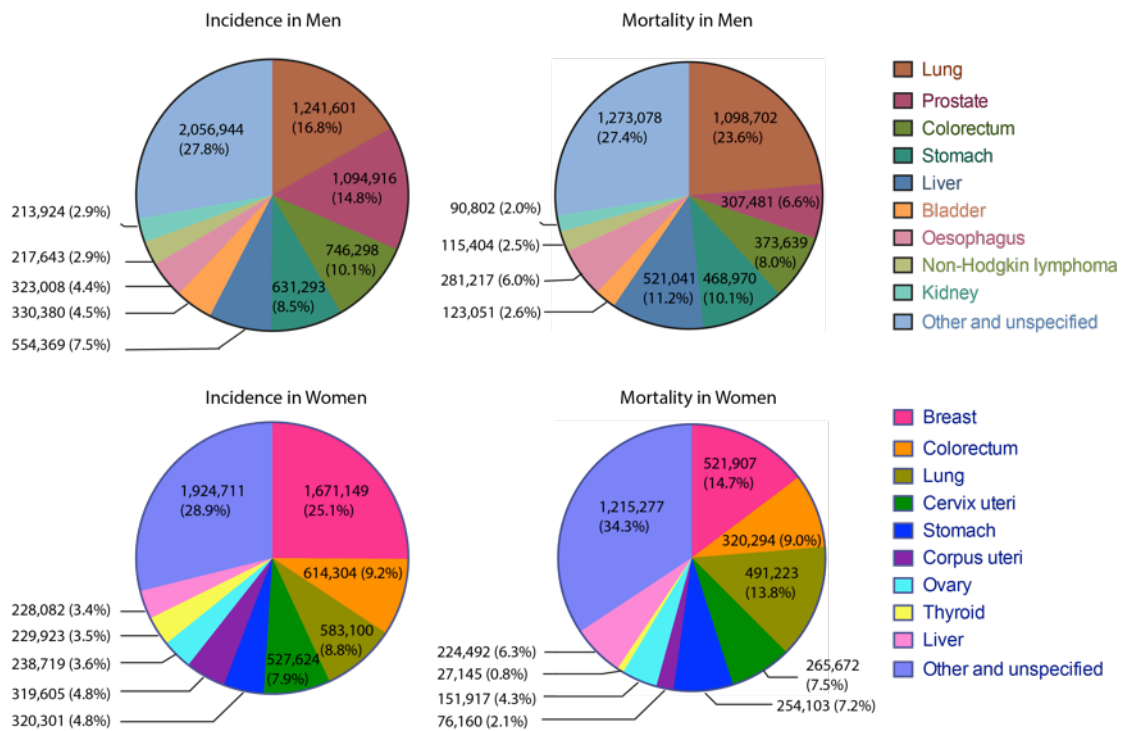
Cancer is a leading cause of morbidity and mortality worldwide, with approximately 14 million new cases and 8.2 million cancer related deaths in 2012 (Ferlay et al., 2015). According to the World Health Organization report, the annual cancer case is expected to increase to 22 million over the next two decades. The global cancer burden continues to increase as a consequence of population aging and growth and also an increasing adoption of cancer-causing lifestyles for example smoking, unhealthy diet and lack of physical activity. Although data vary between more developed and less developed countries, the most common cancers in terms of incidence worldwide are cancers of lung, breast and colorectal; the most common cancers causing deaths are lung cancer, liver cancer and stomach cancer (Figure 1.2).

#### **1.2.1.2 Molecular biology of cancer**

Cancer is a generic term of a collection of diseases characterized by the common feature of abnormal cell growth through which cells may acquire the potential to invade its surroundings and spread to other organs of the body (Hanahan and Weinberg, 2000). The origin of the word cancer dates back to about 400 BC. It was recognized by the Greek physician Hippocrates, the "Father of Medicine" (Sudhakar, 2009). The process by which cancers are constituted is a multistep event resulting from the accumulation of alterations in critical regulatory pathways. Generally, cells are continuously exposed to DNA damaging events, but because of our efficient DNA repair machinery only a small amount of lesions are retained in the genome and become somatically acquired mutations. The accumulation of specific mutations that confer a growth advantage on the cancer cell over normal cells is necessary to drive the development of most cancers.

Cancers result from loss of normal regulatory mechanisms, which arises from the alterations in two major classes of "cancer genes": oncogenes and tumor suppressor genes (Vogelstein and Kinzler, 2004). Mutations in oncogenes generally lead to an increased function of the proteins that are often regulators of cell

proliferation and apoptosis. Tumor suppressor genes generally encode proteins functioning to restrain cell growth or regulate cellular adhesion and motility, and they are normally inactivated with apparent loss-of-function in cancer cells. A key feature of cancer cells is that they are able to re-wire their internal signaling network and therefore become, to varying degrees, independent of internal and external inhibitory signals (Martin, 2003).



**Figure 1.2. Cancer incidence and mortality in 2012.** Estimated numbers of new cancer cases and death in men and women in 2012. Data retrieved from (Ferlay et al., 2015).

### **1.2.1.3 Cancer Therapy**

Human beings have been fighting cancer since it was first recognized. Surgery was the first treatment for cancer and effective cancer surgery began in the late 1800s. Surgery improves the probability of cancer survival as it physically removes malignant tissue and minimizes the likelihood for spread of the cancer. Nowadays, surgery still remains an effective treatment for localized cancers allowing minimal invasion with tremendous improvements in organ conservation. Radiotherapy for cancer treatment was established in the 1930s, which kills proliferating cells with ionizing radiation (Delaney et al., 2005). Radiotherapy is useful in damaging localized tumor when surgical resection is challenging. The introduction of chemicals for cancer treatment and the foundation of chemotherapy took place in the mid-twentieth century when nitrogen mustard was used to treat Hodgkin's disease (Gilman, 1963). With remarkable advances in our understanding of cancer biology and the human genome, chemotherapy is going through a massive transformation. Chemotherapy drugs, the use of single agents or drug cocktails, are often administered as an adjunct to surgery and/or radiotherapy to achieve remission.

In the past 20 years, we have witnessed a remarkable progress in cancer therapy made largely through technological advances and extensive synergies between developments in the academic, pharmaceutical and clinical communities. The access to a wide range of investigative techniques has brought about earlier and more accurate diagnosis, increased sophistication and specificity, enabling the determination of more appropriate and specifically tailored treatments. With the advent of translational medicine, the development of anticancer drugs has started an era of collaboration and sharing of information (Szalma et al., 2010). The information obtained from quantitative and functional analysis of transcriptomes, microRNAomes and methylomes of various cancer cell-line models and clinical specimens continues to disclose the mystery of cancer and plays a pivotal role in the development of anticancer drugs. Apart from the classical surgery, radiotherapy and chemotherapy that are without doubt, the current mainstay of cancer therapeutics, new anticancer approaches are emerging, including cancer immunotherapy (Pardoll, 2012), gene therapy (Izquierdo, 2005) and cell therapy (June, 2007). Now, cancer research is not focused only on the malignant cells itself. Numerous other processes involved in

cancer biology are investigated to improve the lifespan and life quality of cancer patients, for example, the process of invasion (Steeg and Theodorescu, 2008) and angiogenesis (Zhao and Adjei, 2015).

## **1.2.2 Cancer chemotherapy**

### **1.2.2.1 Principles of cancer chemotherapy**

The term "chemotherapy" was coined by the famous German chemist Paul Ehrlich during his search for a "magic bullet" against a major infectious disease of his day, syphilis. Cancer chemotherapy refers to the treatment of cancers using anticancer drugs, and more specifically, using cell killing drugs. The key thread of this pharmacological strategy is to target the selective differences between cancer cells and normal cells. However, for the most parts, cancer cells use the same biochemical machinery as normal cells, which forces investigators to explore rather minute differences. Cancer cells have been considered to differ from normal cells essentially in their ability to survive and proliferate. In the late 1960s, Skipper demonstrated that this difference made chemotherapeutic cures possible in mouse tumors (Skipper et al., 1964). This pioneer work has had a continuing impact and diverts development to cancer chemotherapy (DeVita and Chu, 2008).

Many chemotherapeutic agents exert their cytotoxic effects by interfering with cell proliferation and/or promoting cell apoptosis, to achieve a larger proportion of cell death in cancer cells than normal cells (Makin, 2014). With the dramatic improvements in our understanding of cancer biology, new classes of drug targets potentially more specific for the dysregulated growth of certain malignant cells, have been discovered such as signal transduction inhibitors, grow factor receptor antagonists, apoptosis regulators and antiangiogenic agents.

### **1.2.2.2 Classical cancer chemotherapy agents**

Chemotherapy drugs have been classified into several groups according to their mechanism of action and chemical sources. Some drugs may belong to more than one group as they act in more than one way. These groups are:

- *Alkylating agents.* The alkylating agents, such as nitrogen mustard, cyclophosphamide and platinum compounds, were the first compounds identified to show anticancer activity. This class of drugs forms covalent bonds with nucleophilic sites on deoxyribonucleic acid (DNA), the most vulnerable site being guanine (Warwick, 1963). Alkylation of DNA bases results in various types of adducts such as inter-strand crosslinks that can subsequently lead to the induction of apoptosis.
- *Antimetabolites.* Antimetabolites are chemicals with structures similar to metabolites in cells but different enough to interfere with normal cellular functions. In cancer chemotherapy, most antimetabolites are analogs of either folate or nucleotides. Folates are cofactors necessary for the synthesis of nucleotides, which are the building blocks of DNA and ribonucleic acid (RNA). Folate and nucleotide analogs block the production of nucleotides. Nucleotide analogs may also be incorporated into DNA or RNA. These mechanisms cause impaired DNA and RNA synthesis and induce cell apoptosis.
- *Antimicrotubule agents.* Many antimicrotubule agents are natural products derived from plants such as taxanes and vinca alkaloids. Antimicrotubule agents exert their effects by the interaction with the tubulin/microtubule system required for mitosis (Rowinsky and Donehower, 1991). Binding of this class of agents suppresses microtubule dynamics and leads to mitotic block and eventually cell death by decreasing or increasing microtubule polymerization.
- *Antitumor antibiotics.* Antitumor antibiotics refer to anticancer substances produced by microorganisms, which interact with DNA primarily. The cytotoxic effects of the antibiotics are mediated by different mechanisms, yet many remaining unclear, including covalent complex formation with DNA, DNA cleavage and topoisomerases inhibition (Takeuchi, 1995). Examples of clinically used antitumor antibiotics are dactinomycin, doxorubicin, and bleomycin.
- *Hormonal agents.* Hormonal therapy is used in the treatment of certain types of cancers derived from hormonally responsive tissues such as breast cancer and prostate cancer. Since hormones drive gene expression in certain cancer

cells, manipulation of the activity of certain hormones can lead to cell growth arrest or even cell death. Generally, hormonal therapy involves exogenous administration of specific hormones or hormone antagonists.

- *Other drugs.* There are chemotherapy drugs that act in different ways and do not fall into any of these classes above. Elucidation of the molecular circuitry responsible for abnormal proliferation of cancer cells has provided novel targets for drug development. This field is also where the most effort has been put recently. To list a few developed using more rational approaches, this category includes L-asparaginase (an anticancer enzyme) (McCredie et al., 1973), imatinib (a small-molecule tyrosine kinase inhibitor) (Toga et al., 2003), bortezomib (a proteasome inhibitor) (Chen et al., 2011) and alemtuzumab (monoclonal antibody targeting CD52 antigen) (Havla et al., 2015). Combination chemotherapy also opens new avenues for cancer treatment.

### **1.2.2.3 Drug toxicity**

Drug toxicity is a critical factor in drug development, which is evaluated in preclinical and clinical studies. Drug toxicity may be graded from 1 (mild adverse event) to 5 (death related to adverse event) according to the criteria set by the National Cancer Institute (Institute, 2009). Toxicity, also termed side effect or adverse effect, is a consequence of the mechanism of drug action, drug dose and the patients' conditions. In general, the toxicity profile is dependent on the margin of safety between the dose necessary for efficacy and the dose causing adverse effects, the so called therapeutic window. If the safety margin is small, adverse effects could manifest at the therapeutic dose.

Cancer chemotherapeutic agents are often associated with some adverse effects because these drugs affect normal cells as well, especially cells with rapid turnover for example epithelium cells and blood cells. Common chemotherapy-related adverse effects are fatigue, gastrointestinal distress, immunosuppression, hair loss, anemia, bone marrow suppression and organ damage. Targeted therapies interfering with proteins or processes utilized predominantly by cancer cells generally show less severe side effects than cytotoxic agents, but life-threatening effects can result from their binding to unintended targets (off-targets adverse effects). Although

some advances have been made in optimized treatment scheme and combination chemotherapy (Mayer and Janoff, 2007; Ramsay et al., 2005), drug toxicity remains one of the major limiting factors for the efficacy of cancer chemotherapeutic agents.

#### **1.2.2.4 Drug resistance**

Drug resistance is another key limiting factor for drug development in cancer chemotherapy. Clinically, there are generally two major forms of resistance based on tumor response to the initial treatment, intrinsic and acquired resistance. Intrinsic resistance is an innate property of the tumor cells and is not triggered by drug exposure, while acquired resistance arises as a result of exposure to the drug(s). Owing to the genetic instability and multi-clonal nature of tumors, frequently resistance is intrinsic to cancer cells (Zahreddine and Borden, 2013). Cancer cells undergo mutation at a much higher rate, which grants them the ability to adapt to environmental changes. When cancers are treated with first-line chemotherapy, the clones that are vulnerable to the drugs are eliminated. However, the clones possessing the mutations that make them resistant to the drugs survive and continue to proliferate, leading to recurrent tumors. As chemotherapy has become more effective and the mainstay of cancer therapy, acquired resistance has become prevalent. Many mechanisms of acquired resistance have been identified: decreased drug uptake, decreased drug activation, increased drug metabolism, altered activity and expression of drug targets, enhanced DNA repair and blockage of cell apoptosis (Gottesman, 2002). Apart from these, the physiological conditions to which cancer cells are exposed, termed tumor microenvironment, have been reported to contribute to drug resistance: the hypoxic conditions in solid tumors significantly reduce the effectiveness of many cytotoxic drugs (Mellor et al., 2005); and the extracellular matrix proteins promote tumor cell growth and suppress cell apoptosis (Sethi et al., 1999).

Elucidation of the molecular mechanisms of drug resistance has paved the way to overcome drug resistance. A number of strategies have been investigated to minimize the likelihood of drug resistance. Development of new drugs against drug-refractory tumors has been a major focus in drug development programs (Fojo and Bates, 2003). Meanwhile, there are quite a few promising cases where old drugs are relaunched in terms of clinical dosing, treatment schedule, drug formulation and

delivery to overcome drug resistance (De Souza et al., 2011; Hennenfent and Govindan, 2006; Mamot et al., 2003; Zeng et al., 2014). Application of combination chemotherapy has been found to be effective in overcoming drug resistance in certain cases as well (Hu and Zhang, 2012; Yardley, 2013). In terms of apoptosis-related drug resistance, gene therapy has been explored to restore the abnormal growth caused by mutations in apoptosis-regulating genes (Piche and Rancourt, 2001). With various methods being explore to overcome drug resistance, it has proved that clinical application is more difficult, partly because clinical drug resistance is probably multi-factorial. In spite of great advances, overcoming drug resistance arguably remains the biggest challenge in cancer chemotherapy.

## 1.3 Chemotherapy targeting nucleotide synthesis

### 1.3.1 The roles of nucleotides

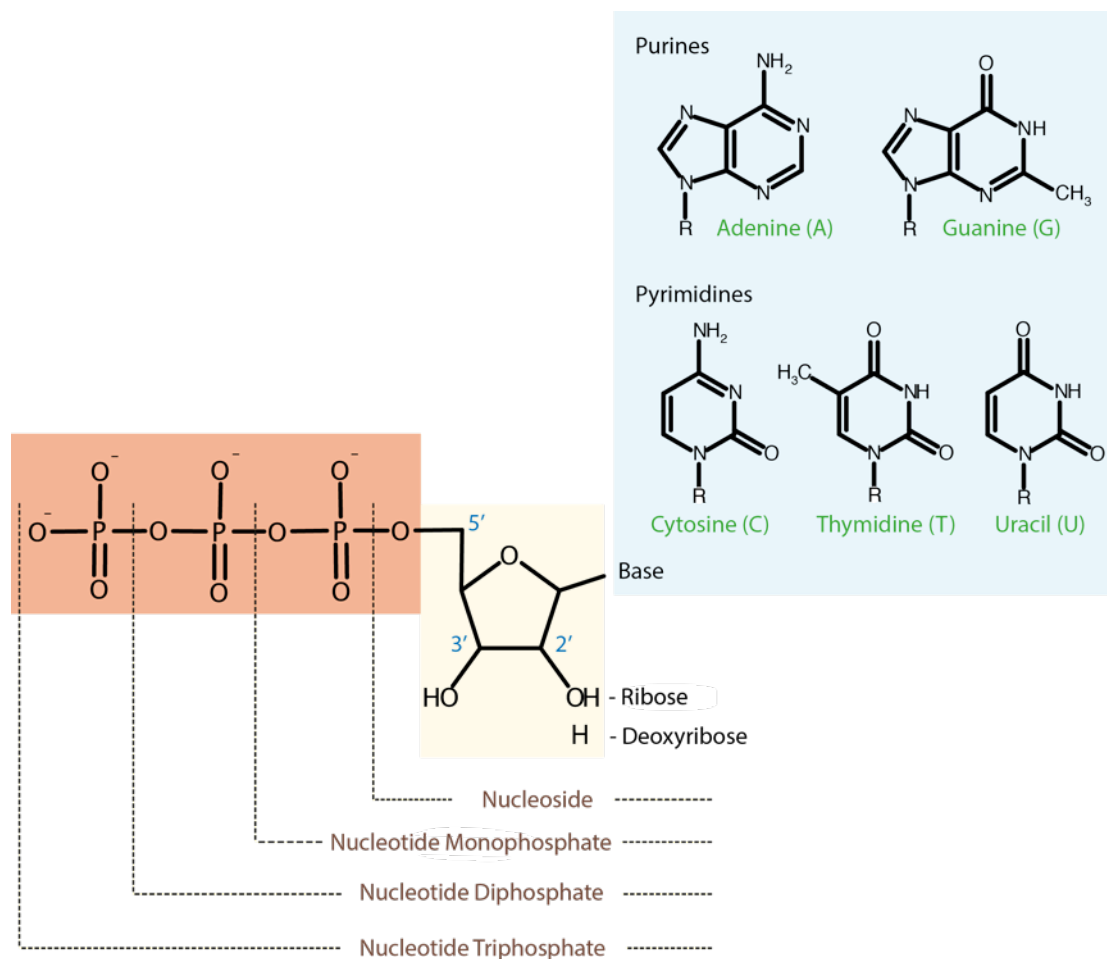
Nucleotides are the basic structural units of nucleic acids. They have a distinctive structure composed of three components: a nitrogenous base of either a pyrimidine or purine; a pentose (ribose or 2'-deoxyribose); and one or more phosphate groups (Figure 1.3). Nucleotides are involved in a wide range of biochemical processes that are vital to cellular metabolism:

1. *As building blocks of nucleic acids.* Nucleotides are the building units of DNA and RNA. DNA and RNA are polymers of deoxyribonucleotides and ribonucleotides, respectively, linked with an alternating sugar-phosphate backbone. They are essential to long-term storage, transmission and expression of genetic information.
2. *As biological regulators.* 3'-5'-cyclic adenosine monophosphate (cAMP) and 3'-5'-cyclic guanosine monophosphate (cGMP) are the oldest and probably the best-known second messengers in intracellular signal transduction. Both cAMP and cGMP can function through their associated protein kinases to regulate and coordinate metabolism within cells, including the regulation of glycogenolysis, lipid metabolism, cellular apoptosis and ion channel conductance (Kelley Bentley and Beavo, 1992; Stanley McKnight, 1991; Thompson and Garbers, 1990).
3. *As co-enzyme components.* Several important co-enzymes are composed of adenine nucleotides (Baddiley and Hughes, 1960). Key players in oxidation-reduction reactions, nicotinamide adenine dinucleotide ( $\text{NAD}^+$ ) and nicotinamide adenine dinucleotide phosphate ( $\text{NADP}^+$ ), are metabolites of adenosine triphosphate (ATP). Additionally, the adenosine monophosphate (AMP) moiety of flavin adenine dinucleotide (FAD) and co-enzyme A (CoA) originates from ATP.
4. *In biosynthetic pathways.* Nucleotides and their derivatives are activated intermediates involved in various biosynthetic pathways. For instance, nucleotide sugars are the energy-rich metabolites for various anabolic processes (Connolly and Duley, 1999): uridine diphosphate (UDP) is a key carrier of sugar residues for polysaccharides biosynthesis; UDP-glucose is the

specific supplier of glucose for glycogen; cytidine diphosphate-choline (CDP-choline) is an intermediate in phospholipids biosynthesis.

5. *In energy metabolism.* Nucleoside triphosphates (NTPs) are a source of chemical energy for a number of biochemical reactions, through losing their terminal phosphate groups (Knowles, 1980). ATP is the universal currency of energy in living systems. When coupled to a reaction, ATP hydrolysis can drive less favorable biochemical processes.

Nucleotides play various essential roles in cellular function. Maintenance of nucleotide levels is therefore fundamental to biological processes. Unsurprisingly, the supply of nucleotide pools should be, and is, tightly regulated at multiple levels.



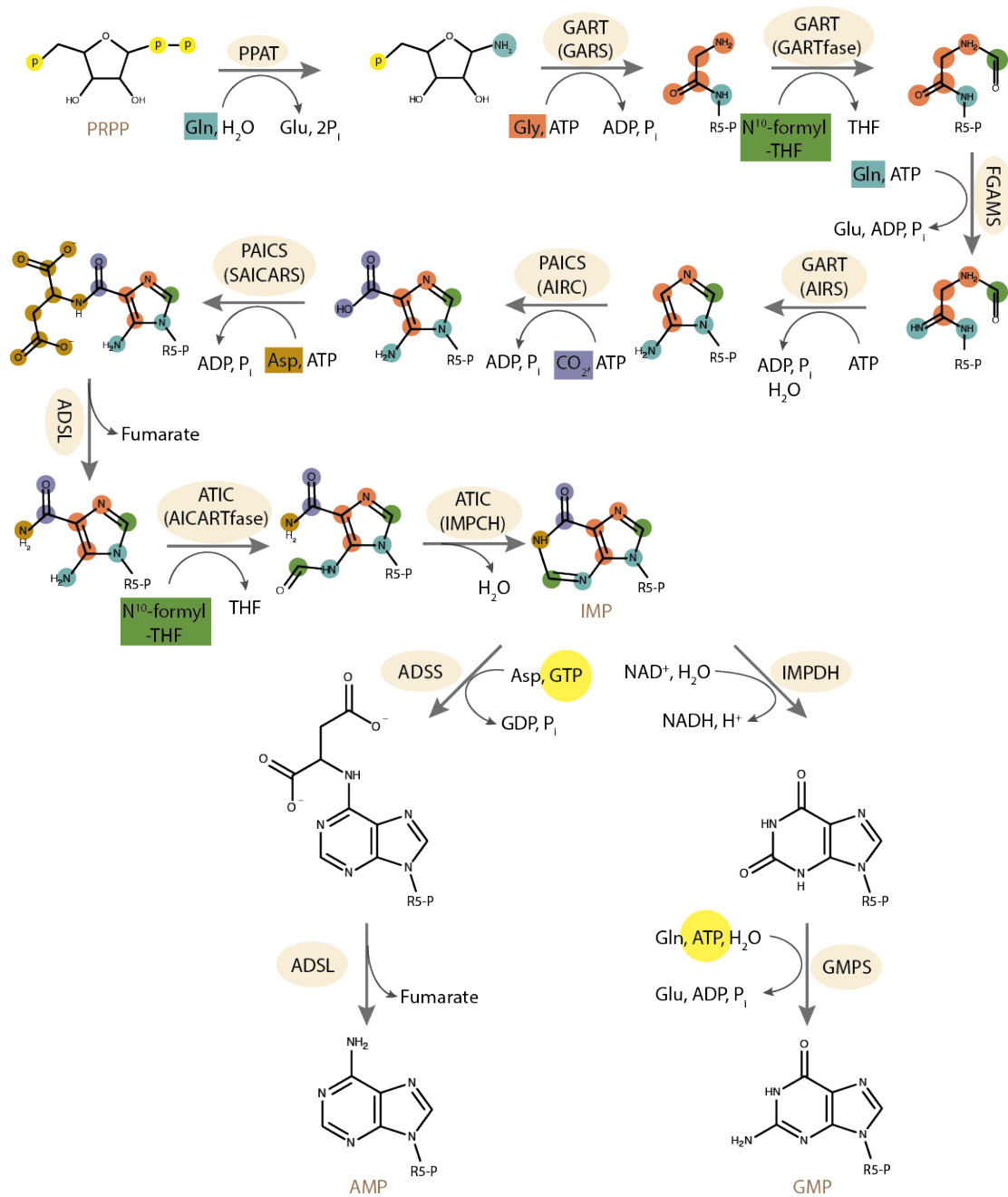
**Figure 1.3. Nucleotide chemical structures.** Nucleotide is composed of a pentose, a purine or pyrimidine base and phosphate(s). Depending on the group at the 2'-position of pentose, it can be ribonucleotide (building blocks for RNA) or deoxyribonucleotide (building blocks for DNA).

## 1.3.2 Nucleotide *de novo* biosynthesis

### 1.3.2.1 *De novo* synthesis of purine nucleotide

Purines are nine-membered heterocyclic compounds, consisting of a pyrimidine ring fused with an imidazole. Adenine and guanine are purine bases in nucleic acids, which form hydrogen bonds with their complementary pyrimidines. Most organisms are capable of synthesizing purine nucleotides *de novo*. In humans, purine nucleotides are mainly synthesized in the liver. The synthesis is entirely cytoplasmic, with different parts derived from various carbon and nitrogen sources in the cell (Moffatt and Ashihara, 2002; Zalkin and Dixon, 1992). The purine ring is built up atom by atom on the activated ribose-5-phosphoribosepyrophosphate (PRPP) (Figure 1.4). The first nucleotide from *de novo* purine synthesis pathway is inosine monophosphate (IMP), which is composed of carbon from glycine, CO<sub>2</sub> and N<sup>10</sup>-formyl-TetraHydroFolate (N<sup>10</sup>-formyl-THF) and nitrogen from glycine, glutamine and aspartate. IMP serves as a central node for purine biosynthesis and can be interconverted into both AMP and GMP. The production of AMP or GMP from IMP requires energy from GTP or ATP, respectively. This cross energy utilization enables the cell to maintain an appropriate balance between AMP and GMP.

Typically, the cellular concentrations of deoxyribonucleoside triphosphates (dNTPs) are relatively low compared with NTPs in resting cells. Only at the time of DNA replication, dNTP pools increase dramatically to meet the demand of DNA synthesis. DNA synthesis requires four dNTPs, which are derived from deoxyribonucleoside diphosphates (dNDPs) via reduction of ribonucleoside diphosphates (rNDP) at the 2'-OH position of the ribose (Kolberg et al., 2004). The reduction takes place in the cytoplasm and is catalyzed by the enzyme ribonucleotide reductase (RNR). AMP and GMP can be phosphorylated to their diphosphate forms, ADP and GDP, which are reduced to deoxyadenosine diphosphate (dADP) and deoxyguanosine diphosphate (dGDP) by RNR, respectively. Further phosphorylation by ATP-dependent kinases produces purine deoxyribonucleotides, dATP and dGTP, for DNA synthesis.



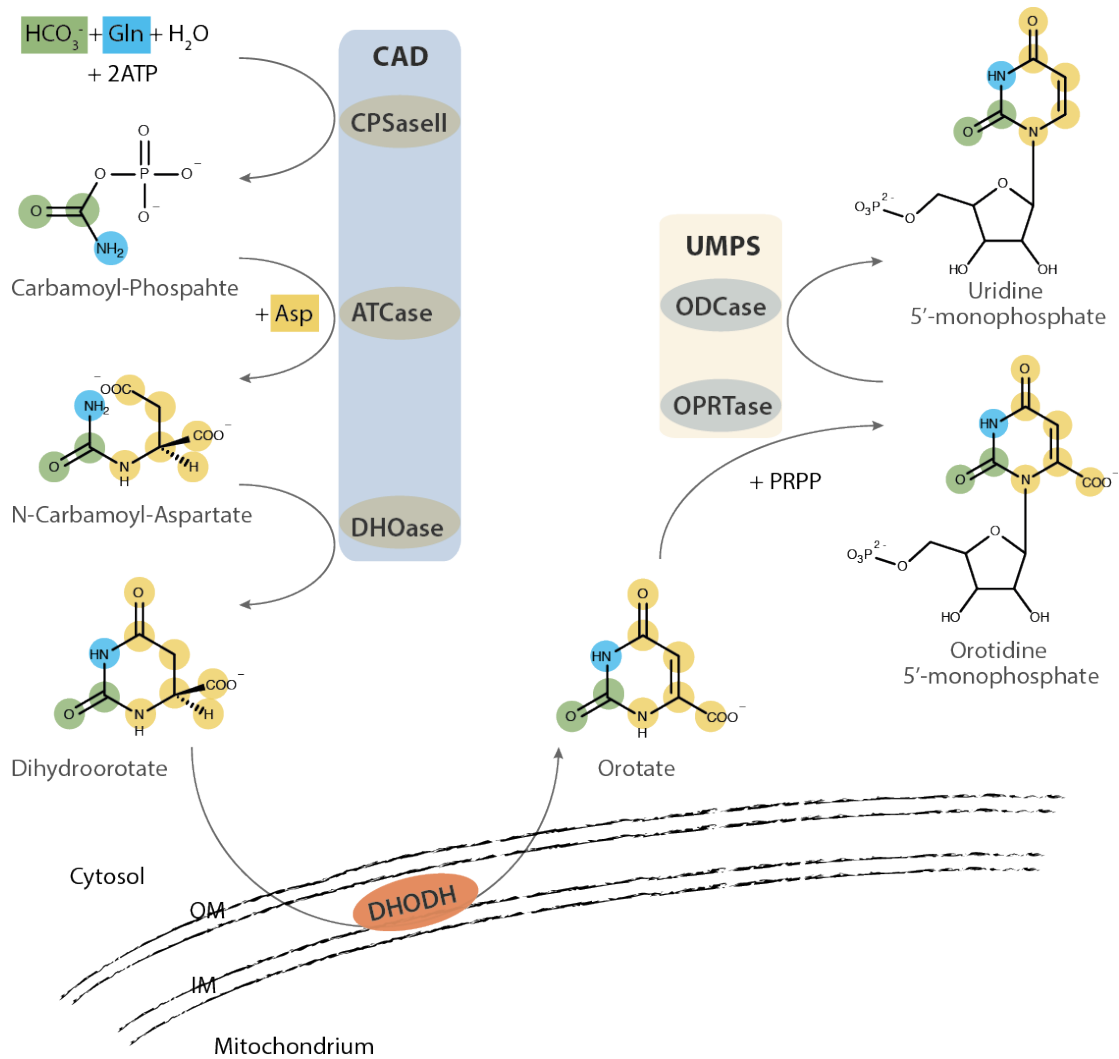
**Figure 1.4. *De novo* synthesis of purine nucleotides.** Purine nucleotide is built up on PRPP, with the purine ring atoms contributed by glutamine, glycine, N<sup>10</sup>-formyl-THF, CO<sub>2</sub> and aspartate. The synthesis process from PRPP to IMP involves three monofunctional enzymes (PPAT, FGAMS and ADSL) and three multifunctional enzymes (GART, PAICS and ATIC). IMP is converted to AMP or GMP by different sets of enzymes. Adapted from (Moffatt and Ashihara, 2002).

Abbreviations: PPAT, phosphoribosyl pyrophosphate amidotransferase; GARS, phosphoribosyl pyrophosphate amidotransferase; GARTfase, glycinamide ribonucleotide transformylase; AIRS, aminoimidazole ribonucleotide synthetase; FGAMS, phosphoribosyl-formylglycinamide synthetase; AIRC, aminoimidazole ribonucleotide carboxylase; SAICARS, succinylaminoimidazolecarboxamide ribonucleotide synthetase; ADSL, adenylosuccinate lyase; AICARTfase, 5-aminoimidazole-4-carboxamide ribonucleotide formyltransferase; IMPCH, IMP cyclohydrolase; ADSS, adenylosuccinate synthase; IMPDH, IMP dehydrogenase; GMPS, GMP synthetase.

### 1.3.2.2 *De novo* synthesis of pyrimidine nucleotide

Pyrimidines are six-membered heterocyclics containing two nitrogen atoms in the ring. In nucleic acids, three bases are pyrimidine derivatives: cytosine, thymine and uracil. Unlike purines, the *de novo* synthesis of pyrimidine nucleotides occurs in a variety of tissues. The synthesis is compartmentalized, with one step occurring in the mitochondria and other steps residing in the cytoplasm (Figure 1.5) (Löffler and Zameitat, 2013). Pyrimidine nucleotides are synthesized *de novo* from simple precursors, while the synthesis differs from purine synthesis in that the pyrimidine ring is synthesized independent of ribose-5-phosphate. The pyrimidine ring is assembled first and then linked to PRPP to form the first pyrimidine nucleotide in cells, uridine monophosphate (UMP). The pyrimidine ring of UMP is synthesized from CO<sub>2</sub>, glutamine, ATP and aspartate. Uridine di- and triphosphate (UDP and UTP) are produced from UMP by ATP-dependent kinases. Cytidine triphosphate (CTP) synthase utilizes UTP and the amide group of glutamine to form CTP (Lieberman, 1956).

Similar to purine deoxyribonucleotides, pyrimidine deoxyribonucleotides are derived from NDPs reduced by RNR, though in a slightly more complex way (Löffler et al., 2005). CTP is dephosphorylated to CDP, which is converted to deoxycytidine diphosphate (dCDP) by RNR and then to deoxycytidine triphosphate (dCTP) via phosphorylation. The DNA-unique nucleotide deoxythymidine is synthesized through methylation of deoxyuridine monophosphate (dUMP) catalyzed by the enzyme thymidylate synthase (TS) (Harrap et al., 1989). dUMP comes either from a specific phosphatase activity on dUDP and dUTP or from deamination of dCMP (Maley and Maley, 1963). dUTP is effectively converted back to dUMP by dUTP nucleotidohydrolase, in order to prevent its incorporation into DNA (Persson et al., 2001). The resulting deoxythymidine monophosphate (dTMP) derived from dUMP is subsequently converted to deoxythymidine triphosphate (dTTP) through two sequential phosphorylation activities (Figure 1.6).



**Figure 1.5. *De novo* synthesis of pyrimidine nucleotides.** The synthesis of pyrimidine nucleotide is compartmented into cytosol and mitochondria. The multifunctional enzyme CAD (CPSaseII + ATCase + DHOase) and UMPS (OPRTase + ODCase) are cytosolic. The enzyme DHODH resides on the outer face of the inner mitochondrial membrane. Adapted from (Löffler and Zameitat, 2013).

Abbreviations: CPSaseII, carbamoylphosphate synthetase II; ATCase, aspartate transcarbamoylase; DHOase, dihydroorotate dehydrogenase; UMPS, uridine monophosphate synthase; OPRTase, orotate phosphoribosyltransferase; ODCase, orotidine decarboxylase.

### **1.3.3 Regulation of nucleotide biosynthesis**

*De novo* synthesis of nucleotides is an energy-costly process that involves multiple metabolic pathways across different cell compartments. The processes are strongly regulated at multiple levels to meet the metabolic demands. Biosynthesis of RNA and DNA is a critical component of cell division. The regulation of nucleotide biosynthesis is therefore tightly linked to the cell cycle. A number of studies have addressed nucleotide biosynthesis and its regulation (Blakley and Vitols, 1968; Cosgrove, 1998; Kelley and Andersson, 2014; Lane and Fan, 2015; Moffatt and Ashihara, 2002). It has been shown that the biosynthesis processes are regulated at both the transcription level by several transcription factors and the enzyme level via allosteric interactions and feedback inhibition mechanism.

#### **1.3.3.1 Regulation of purine nucleotide biosynthesis**

*De novo* synthesis of purine nucleotides (Figure 1.4) involves about 20 enzymes, of which the genes are distributed over nine chromosomes. Coordinated expression of these functionally linked enzymes on different chromosomes is a prerequisite for appropriate nucleotide supply. This is achieved by making use of a common set of transcription factors. Among various common transcription factors identified so far, c-Myc is a master regulator of gene expression in nucleotide biosynthesis. c-Myc directly regulates the expression of several key enzymes in purine synthesis (Lane and Fan, 2015), including phosphoribosyl pyrophosphate amidotransferase (PPAT), the trifunctional enzyme GART [glycinamide ribonucleotide synthetase (GARS) + aminoimidazole ribonucleotide synthetase (AIRS) + glycinamide ribonucleotide transformylase (GARTfase)] and IMP dehydrogenase (IMPDH1 and IMPDH2). As an example of cell cycle coordination, the R2 subunit of RNR is strongly induced in phase to enable the production of the dNTPs needed for cell division (Chabes and Thelander, 2000).

Besides transcription-level regulation, direct modulation of enzyme activities plays an important role in the pathway. In purine nucleotide biosynthesis, the key regulatory steps occur in two phases. The first rate-limiting and commitment step is the transfer of an amide group from glutamine to PRPP catalyzed by PPAT (Figure 1.4). The enzyme PPAT is subject to feedback regulation by its end products,

adenosine and guanosine ribonucleotides (Hartman, 1963). Notably, binding of specific purine nucleotide pairs inhibits the enzyme synergistically. On the other hand, PRPP is suggested to be able to overcome the feedback inhibition and activate PPAT. The second control step deals with the branched portions between AMP and GMP from IMP. AMP inhibits adenylosuccinate synthetase through feedback inhibition (Stayton et al., 1983), and similarly, GMP inhibits IMP dehydrogenase (Jackson et al., 1975). To make the regulation tighter and as already mentioned, GTP and ATP provide the energy for the synthesis of AMP and GMP, respectively. This reciprocal substrate relation contributes to the maintenance of an appropriate balance between adenosine and guanosine ribonucleotides.

### **1.3.3.2 Regulation of pyrimidine nucleotide biosynthesis**

Similar to purine synthesis, coordination of gene expression involved in pyrimidine nucleotide synthesis is under the control of transcription factors. c-Myc serves as the key player and regulates the expression of the trifunctional enzyme CAD [carbamoylphosphate synthetase II (CPSaseII) + aspartate transcarbamoylase (ATCase) + dihydroorotase (DHOase)], dihydroorotate dehydrogenase (DHODH) and the bifunctional enzyme UMPS [orotate phosphoribosyltransferase (OPRTase) + orotidine decarboxylase (ODCase)] (Lane and Fan, 2015).

The first step of CAD, synthesis of carbamoyl phosphate catalyzed by CPSaseII (Figure 1.5), has been described as the rate-limiting step in pyrimidine biosynthesis (Löffler and Zameitat, 2013). The activity of cytoplasmic CPSaseII controls the flux through *de novo* pyrimidine synthesis pathway. The enzyme is sensitive to feedback inhibition by UTP but is also allosterically activated by PRPP. PRPP is a feed-forward activator in pyrimidine production and also the substrate in the rate-limiting step of purine synthesis, thereby coordinating regulation of pyrimidine and purine biosynthesis. Additionally, the CAD activity is tuned by phosphorylation in response to growth stimulation (Graves et al., 2000). The process of *de novo* pyrimidine synthesis is compartmented into cytoplasm and mitochondria. Dihydroorotate (DHO), the final product of CAD in cytoplasm, is captured and oxidized by the enzyme dihydroorotate dehydrogenase (DHODH) located on the outer surface of the inner mitochondrial membrane. Because DHODH is functionally connected to the respiratory chain, pyrimidine biosynthesis becomes a pacemaker of

cell growth and proliferation under hypoxic conditions (Gattermann et al., 2004). The relative pool size of uridine and cytidine nucleotides is regulated by cytidine triphosphate synthase (CTPSase). The enzyme is inhibited by its product CTP, but is also activated by GTP, which represents another control locus to coordinate the regulation of purine and pyrimidine production (Noree et al., 2014).

A noteworthy feature of enzymes in nucleotide biosynthesis pathways is their multifunctional nature. Several enzymes perform multiple physiological functions including GART in purine synthesis and CAD in pyrimidine synthesis. Encoding multiple enzyme activities on a single polypeptide chain simplifies the coordination and regulation at both the transcription and enzyme activity level. In addition, multifunctional enzymes achieve a micro-compartmentation, which facilitates the intermediates channeling and thus enhances the efficiency considerably (Easterby, 1989; Roy, 1999).

### **1.3.3.3 Regulation of deoxyribonucleotide biosynthesis**

The reduction of ribonucleotides to deoxyribonucleotides by RNR is the essential step for all dNTPs production. The enzyme RNR has a sophisticated control mechanism, regulated at the substrate level by allosteric interactions of RNR with ATP and dNTPs (Reichard, 1985). RNR consists of two allosteric sites: one controls the overall catalytic activity and the other regulates substrate specificity of the enzyme complex. The enzyme is inhibited when dATP binds to the activity site as a feedback inhibitor to signal an abundance of dNTPs. The binding of ATP releases the inhibitory effect and activates RNR. Consequently, the ratio of cellular ATP to dATP regulates the net flux to dNDP: a high ATP/dATP ratio stimulates a reductive flux, while a low ratio indicates low demand for dNDPs in the cell. The substrate-specificity site binds ATP, dATP, dGTP and dTTP, which allows RNR to detect the relative abundance of the four dNTPs effectively and thus adjust its activity towards different substrates accordingly. The binding of ATP or dATP promotes the reduction of pyrimidine nucleotides, UDP and CDP. dTTP enhances the reduction on GDP and prevents further reduction of pyrimidine nucleotides. The binding of dGTP induces the reduction of ADP. Together, these regulatory interactions achieve an appropriate balance of the four dNTPs required for DNA synthesis and repair.

As mentioned above, dTTP is synthesized in a more complex way in order to prevent the accumulation of dUTP, which could be incorporated into DNA mistakenly. Thymidylate synthase (TS) catalyzes the synthesis of dTMP from dUMP by utilizing 5,10-methylenetetrahydrofolate (5,10-CH<sub>2</sub>-THF) as the methyl carbon donor. The activity of TS is subject to the availability of its folate co-substrate. Folate is also involved in the synthesis of IMP. Consequently, production of dTTP and purine nucleotides is tightly linked to the folate metabolism (Rowe and Lewis, 1973).

### 1.3.4 Nucleotide degradation and salvage

There is a significant turnover of RNA in cells, both during the processing of mature RNAs from larger precursors and for the regulation purpose, e.g. decreasing mRNA level to reduce certain protein production. DNA is very stable and does typically not turn over but portions of DNA are excised during DNA repair. Hydrolysis of nucleic acids by nucleases generates a mixture of polynucleotides, which are further cleaved by phosphodiesterases to a pool of mononucleotides. Those nucleotides released from the breakdown of polymeric RNA and DNA are degraded and excreted or salvaged to re-enter the nucleotide synthesis routes (Nyhan, 2001). The balance between degradation and salvage is linked to cellular demand of nucleotides and nutrients supply. In general, as *de novo* synthesis is energetically costly, its activity is low in resting cells and the demand for nucleotides is largely satisfied by the salvage pathways (Moffatt and Ashihara, 2002). When cells divide, however, *de novo* synthesis becomes highly active to meet the metabolic demands. The activity difference in *de novo* synthesis between resting and dividing cells offers an opportunity to selectively target cancer cells but spare normal cells. This is one of the key reasons for targeting *de novo* synthesis pathway in chemotherapeutic drug development.

#### 1.3.4.1 Purine nucleotide degradation and salvage

In humans, purine nucleotides are catabolized to uric acid and excreted by the kidney into the urine. GMP is hydrolyzed to guanosine, which undergoes phosphorolysis to guanine and ribose-1-phosphate. AMP can be catabolized in a similar way, but to a less extent, due to the low activity of 5'-nucleotidases to AMP in humans (Spychala et al., 1989). However, AMP can be deaminated to IMP by AMP deaminase. IMP is degraded via hydrolysis to inosine and further phosphorolysis to hypoxanthine. Both guanine and hypoxanthine are converted to xanthine and then oxidized to uric acid for excretion (Fox, 1978).

Purine bases (guanine, hypoxanthine and adenine) can be recycled from degradation for salvage synthesis. The key enzyme taking purines to uric acid, xanthine oxidase is significantly active only in the liver and intestine (Watts et al., 1965). In non-hepatic tissues, purines are not readily catabolized to uric acid and thus

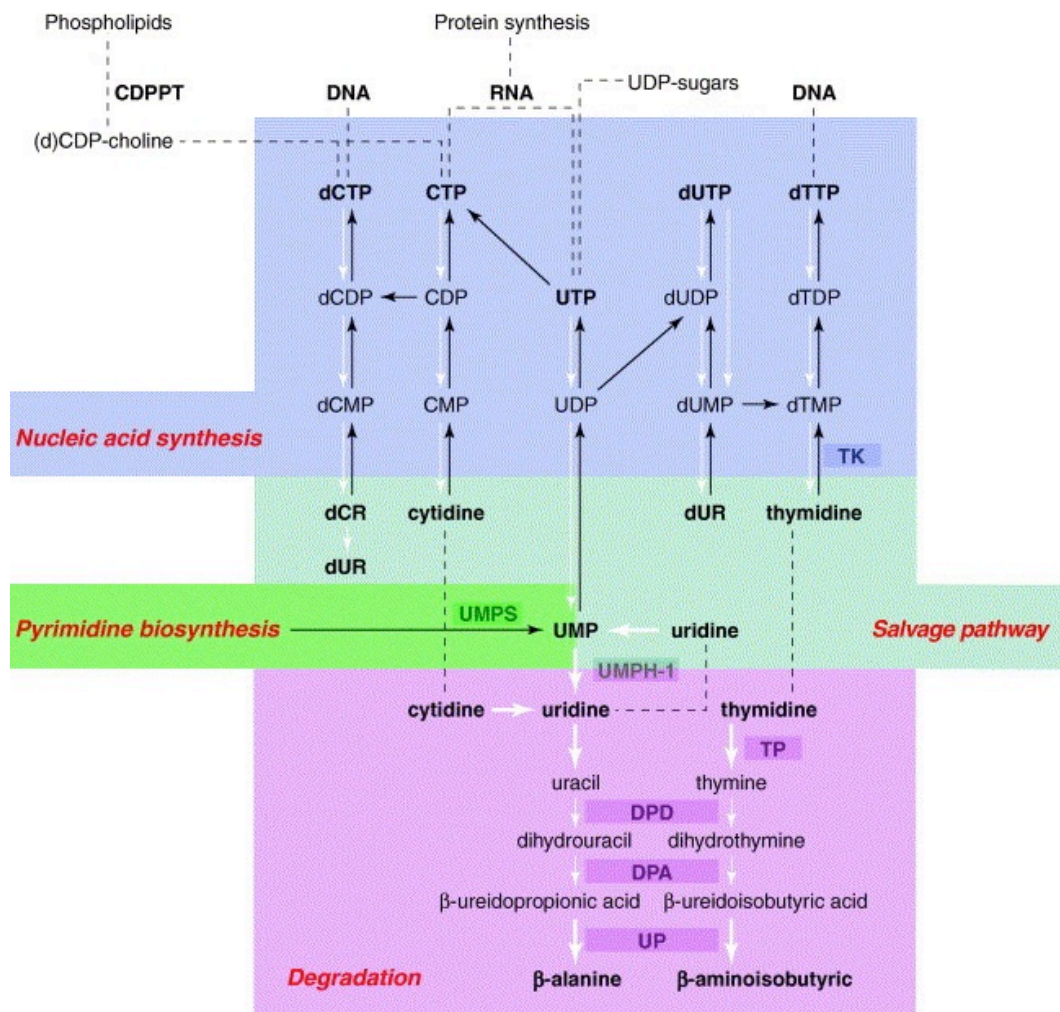
are available for salvage. Purine salvage in human is achieved mainly via the addition of ribose-5-phosphate to purine bases from PRPP by phosphoribosyl-transferases (Nyhan, 2001). Guanine and hypoxanthine are salvaged to their respective nucleotides, GMP and IMP, by hypoxanthine guanine phosphoribosyl-transferase (HGPRT). Adenine can also be converted to AMP by adenine phosphoribosyltransferase (APRT), but less significantly because very small amount of adenine is produced during catabolism. AMP is produced from IMP primarily or via phosphorylation of adenosine by adenosine kinase (AK). IMP, the branch joint of purine biosynthesis, can be converted to both AMP and GMP and thus maintains the balance of guanine and adenine nucleotide pools.

#### **1.3.4.2 Pyrimidine nucleotide degradation and salvage**

In contrast to purines, pyrimidines undergo ring cleavage in their degradation processes. The end products of pyrimidine catabolism are usually  $\beta$ -amino acids,  $\text{NH}_3$  and  $\text{CO}_2$ . In humans, unrecycled pyrimidine nucleotides are split to give free bases and ribose-1-phosphate by 5'-nucleotidases and pyrimidine nucleoside phosphorylases, which initiate the catabolic pathway. Free pyrimidine bases, uracil and thymine, are reduced by dihydropyrimidine dehydrogenase (DPD). Cleavage of pyrimidine rings follows the reduction reaction, from which  $\text{NH}_3$  and  $\text{CO}_2$  are released and the rest is left as a  $\beta$ -amino acid (Figure 1.6). The  $\beta$ -amino acids are excreted into urine or can be utilized and converted to malonyl-CoA and methylmalonyl-CoA for fatty acid synthesis and ATP production, respectively (Löffler and Zameitat, 2013).

In addition to the contribution from nucleic acids breakdown, pyrimidine nucleotide pools can stem from daily nutrients. Unlike purines, dietary pyrimidines are readily taken up by the gut. These pyrimidines can be rescued and diverted to pyrimidine biosynthesis. The relative significance of salvage versus *de novo* synthesis probably depends on the specific tissue of distinct enzyme expression profiles and on the dietary conditions. Prominent salvage capacity is observed in brain and lymphoid cells. Salvage of pyrimidine nucleotides commences with the dephosphorylation of pyrimidine (deoxy)ribonucleoside monophosphates catalyzed by 5'-nucleotidases, which is required for delivery of pyrimidines via the circulation (Nyhan, 2001). Pyrimidine salvage occurs on the nucleoside level instead of the bases in purine

salvage. Specific nucleoside kinases, utilizing ATP as the major phosphate donor, have been identified to rescue pyrimidine nucleosides from the circulation. Uridine kinase phosphorylates both uridine and cytidine (Cheng et al., 1986). Thymidine kinase participates in the salvage thymidine to dTMP (Munch-Petersen et al., 1995). Deoxycytidine kinase, though generally specific for lymphocytes, yields dCMP (Shewach et al., 1992). Orotate from daily nutrients can be absorbed and synthesized to UMP, but this occurs only in the liver and erythrocytes. UMP from salvage pathway is ready to be the common intermediate in pyrimidine *de novo* synthesis.



**Figure 1.6. Biosynthesis, interconversion and degradation of pyrimidine nucleotides.** UMP can be interconverted into (deoxy)cytidine and thymidine nucleotides. Salvage of pyrimidines occurs on the nucleoside level. The degradation pathway takes place in the liver. Retrieved from (Löffler et al., 2005). Abbreviations: DPD, dihydropyrimidine dehydrogenase; DPA, dihydropyrimidinase; TK, thymidine kinase; TP, thymidine phosphorylase; UMPH-1, UMP hydrolase; UMPS, UMP synthase; UP, ureidopropionase.

### 1.3.5 Pharmacological inhibition of nucleotide synthesis

Traditionally, cancer chemotherapy is based on the selective difference of cell proliferation rate between cancer cells and normal cells. Due to their rapid growth and proliferation, cancer cells are believed to be more susceptible to interference with DNA synthesis. These findings have led to the development of a class of anticancer drugs designed to interfere with the synthesis of nucleotides, which are the building blocks of DNA and RNA. Inhibition of nucleotide synthesis at a particular point of the pathways could result in accumulation of intermediates and consequently imbalance of the nucleotide pools, which causes arrest of DNA synthesis and usually cell death via apoptosis (Christopherson et al., 2002; Munoz-Pinedo et al., 2012).

Anticancer drugs interfering with nucleotide synthesis can be categorized into two major groups: nucleotide analogs and folate analogs. Some known chemotherapy agents targeting nucleotide synthesis are summarized in Table 1.2. Nucleoside analogs are typically anabolized to their active nucleotide forms by cellular enzymes. The broad substrate specificity of certain nucleoside kinases (e.g. deoxycytidine kinase) offers the channel to activate many nucleoside analogs. Nucleoside kinases, moreover, have been reported to be up-regulated in certain malignant cells, which provides a beneficial activation of nucleoside analogs in cancer cells (Arner and Eriksson, 1995). The active nucleotide analogs exert their cytotoxic effect mainly through one or more of the following three mechanisms: 1) inhibition of enzyme(s) involved in nucleotide synthesis pathways; 2) incorporation into DNA, resulting in chain termination and/or inhibition of DNA synthesis and function; 3) incorporation into RNA, interfering with RNA processing and translation.

Folate, as mentioned, is an important nucleotide precursor. It is involved in both purine and pyrimidine nucleotide synthesis. Folate analogs, also termed antifolates, antagonize the actions of folate in several key folate-requiring enzymes. Dihydrofolate reductase (DHFR) is one of the primary targets of antifolates. Inhibition of DHFR causes functional folate deficiency as a consequence of a decreased intracellular reduced folate pool, including N<sup>10</sup>-formyl-THF (the co-substrate in purine synthesis) and 5,10-CH<sub>2</sub>-THF (the co-substrate in dTMP synthesis) (Walling, 2006).

---

**Table 1.2. Examples of chemotherapy agents targeting nucleotide synthesis**

<b>Drug</b>	<b>Drug class</b>	<b>Mechanism of action</b>	<b>References</b>
5-fluorouracil	Pyrimidine analog	Monophosphate form inhibits TS for dTMP synthesis; DNA and RNA incorporation.	Nature Reviews Cancer. 3, 330-338 (2003)
Gemcitabine	Purine nucleoside analog	Triphosphate form inhibits RNR for deoxyribonucleotide synthesis; DNA incorporation.	Seminars in Oncology. 4, 3-10 (1995)
Methotrexate	Folate analog	Inhibits dihydrofolate reductase (DHFR) primarily resulting in functional folate deficiency and DNA synthesis inhibition	Nature Reviews Cancer. 5, 65-72 (2005)
Raltitrexed	Folate analog	Inhibits TS primarily leading to thymineless state and DNA synthesis inhibition	Advances in Experimental Medicine and Biology. 309a, 19-23 (1991)
Pemetrexed	Folate analog	Multi-targeted antifolate; Inhibits GARTfase and AICARTfase (in purine synthesis), TS, DHFR.	Lancet Oncology. 2, 298-306 (2001)

---

## 1.4 Human thymidylate synthase: a critical target in cancer chemotherapy

### 1.4.1 Overview of thymidylate synthase

Thymidylate synthase (TS, EC 2.1.1.45) catalyzes the reductive methylation of dUMP to dTMP using 5,10-CH<sub>2</sub>-THF as a carbon donor (Figure 1.7). The product dTMP (also known as thymidylate) is subsequently phosphorylated to dTTP, which is a critical precursor for DNA synthesis and repair. The reaction catalyzed by TS represents the sole intracellular source of *de novo* thymidine nucleotides. Therefore, inhibition of TS exploits one of the few metabolic bottlenecks in DNA synthesis and leads to a thymineless state, which is cytotoxic to actively dividing cells (Goulian et al., 1986).

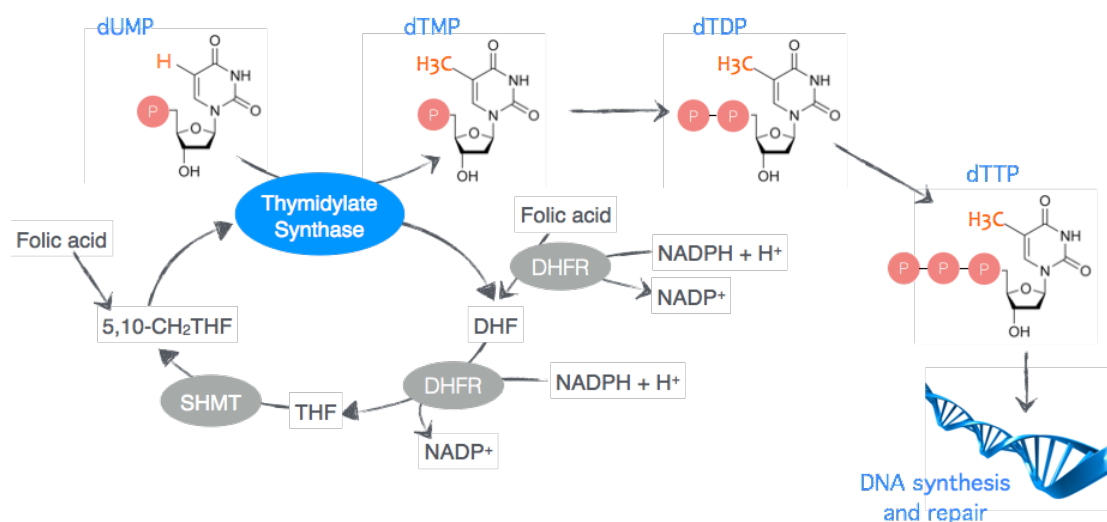
Among those folate-utilizing enzymes, TS is unique in that the cosubstrate 5,10-CH<sub>2</sub>-THF serves as both a one-carbon donor and a reductant. The 5,10-CH<sub>2</sub>-THF pool is replenished by the sequential action of DHFR and serine hydroxymethyl transferase (SHMT) (Rowe and Lewis, 1973). In humans, animals, fungi and some bacteria, TS and DHFR are translated separately and have their own existence. TS functions as a homodimer. In protozoa and many plants, TS and DHFR are encoded on a single polypeptide chain with the DHFR domain at N-terminal and TS domain at C-terminal. The bifunctional DHFR-TS enzyme forms dimer through the TS dimer interface (Ivanetich and Santi, 1990). Organization of these two protein domains in a single polypeptide chain is believed to be an evolutionary design for genetic economy of the organism and to enhance the substrate channeling and thus the efficiency of the two enzymes (Elcock et al., 1996; Landau, 2013).

The protein sequences of TS from different sources are quite similar except for some differences in three regions: the N-terminal region and two insert regions at position 117<sup>1</sup> and 146 based on human TS (hTS) residue number (Figure 1.8). Compared with *E. coli* and *L. casei* TS, the N-terminus of TSs from human, human herpes virus 3 (HHV3) and mouse is extended with various lengths (*P. falciparum* TS domain is not included as it has the junction region with DHFR at the N-terminus). In addition, *E. coli* and *L. casei* TS differ from other TSs in the two insert regions: *E. coli* lacks the two insertions; *L. casei* lacks one insert region but has a much longer

---

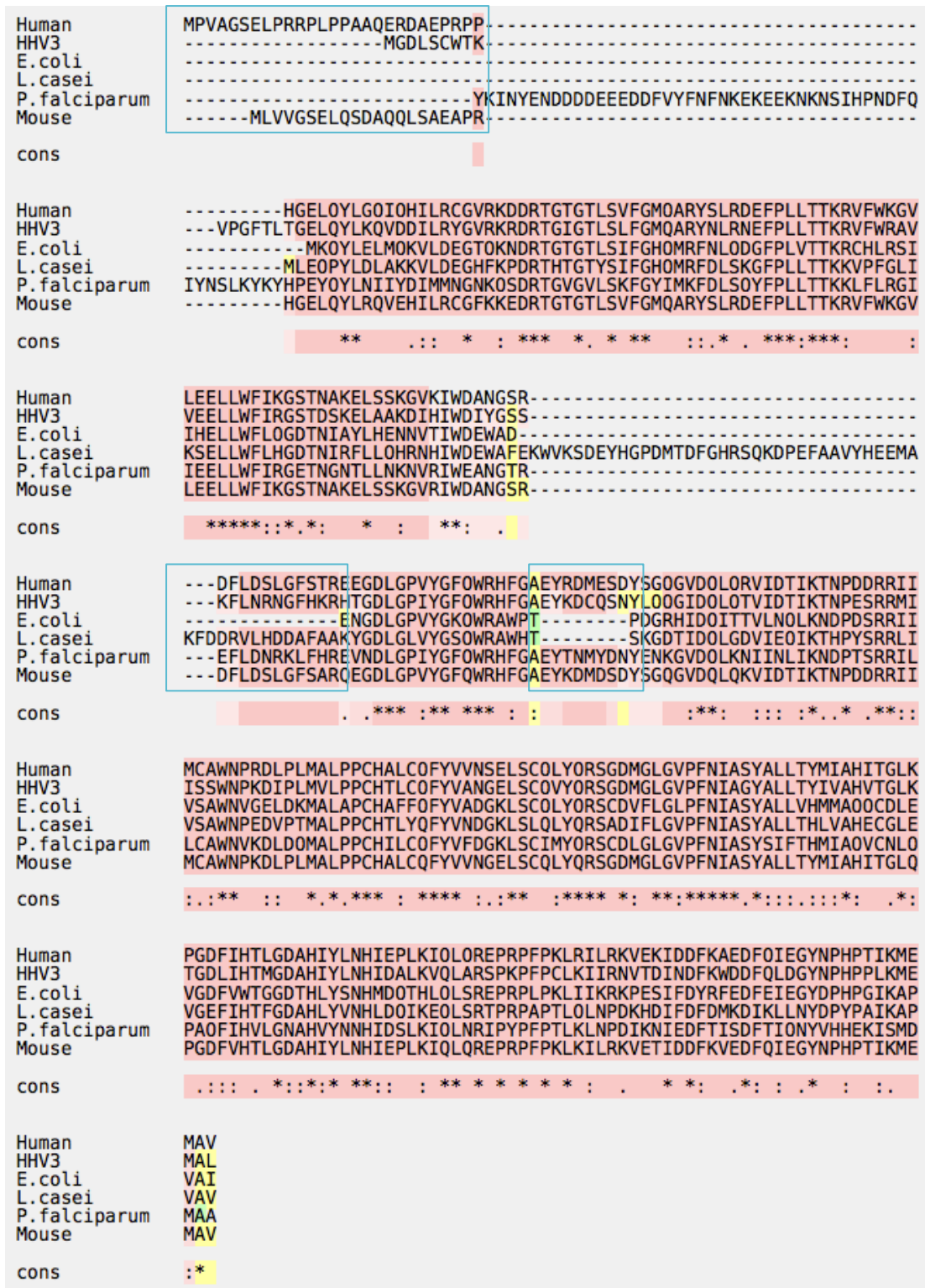
<sup>1</sup> The numbering scheme used here is based on human TS.

insertion in the other region. Despite extensive studies on TS as a drug target in the last few decades, the function and importance of these three regions have received limited attention. In recent years, the N-terminus of TS has been demonstrated to be a primary determinant of the enzyme's intracellular stability (Huang et al., 2010; Peña et al., 2006; Peña et al., 2009). The role of the two insert regions remains unclear.



**Figure 1.7. Function of TS and regeneration of folates.** TS catalyzes the production of dTMP, which is subsequently phosphorylated to dTTP for DNA synthesis and repair. TS utilizes 5,10-CH<sub>2</sub>-THF as the carbon donor. The folate pool is replenished by sequential actions of DHFR and SHMT.

Abbreviations: DHF, dihydrofolate; THF, tetrahydrofolate; SHMT, serine hydroxymethyl transferase.



BAD AVG GOOD

**Figure 1.8. Protein sequence alignment of TS from different organisms.** The multiple protein sequence alignment is performed using Expresso (Di Tommaso et al., 2011), which aligns sequence using structural information. Protein sequences are retrieved from UniProt. The three regions with significant difference among those organisms are highlighted within boxes. Note: *P. falciparum* has a bifunctional DHFR-TS. The sequence of TS domain here includes part of its junction region. Abbreviations: HHV3, human herpes virus 3; *E. coli*, *Escherichia coli*; *L. casei*, *Lactobacillus casei*; *P. falciparum*, *Plasmodium falciparum*.

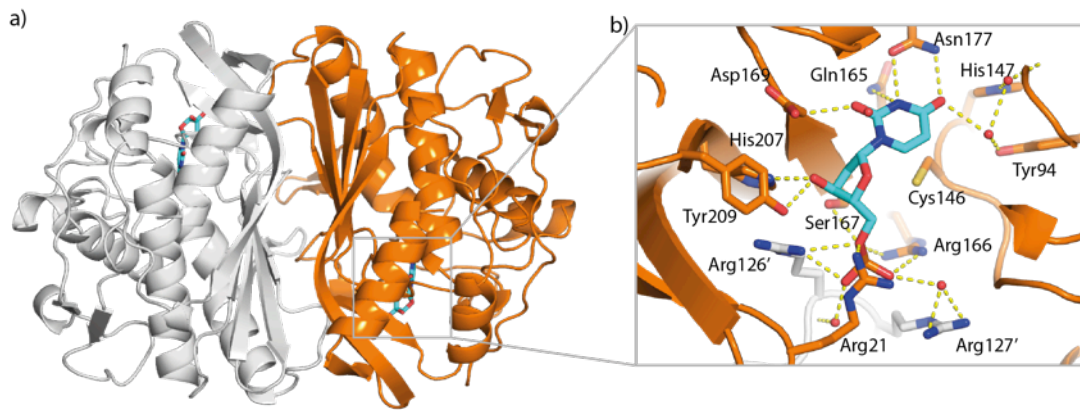
## 1.4.2 Structure and catalytic mechanism of TS

### 1.4.2.1 Binding of substrate dUMP

The understanding of the structure and molecular mechanism of hTS has improved immensely since this enzyme was recognized as an anticancer drug target. The first TS structure solved was *L. casei* TS in 1987 (Hardy et al., 1987). Subsequently, TS structures from different sources were solved and confirmed the striking structural homology of this enzyme. Among those TS enzymes from different species, *E. coli* TS is most extensively explored. Stroud's group has contributed immensely to the understanding of *E. coli* TS on the structure and catalytic mechanism over the years, with a number of X-ray crystal structures of TS and TS complexes (Kamb et al., 1992a; Kamb et al., 1992b; Perry et al., 1990). The native enzyme is a homodimer, of which the dimer interface is formed by a six-stranded twisted  $\beta$ -sheet from each subunit (Figure 1.9a). Each subunit contains a deep active site cavity with one substrate dUMP binding site and one folate binding site. Compared with the ligand-free enzyme (phosphate-bound TS-P<sub>i</sub>), the TS-dUMP binary complex is virtually identical (Finer-Moore et al., 1993; Hardy et al., 1987). The substrate dUMP binds to the active site cavity in an extended conformation. The binding affinity of dUMP has a strong contribution from the electrostatic interaction of the phosphate moiety with a conserved four-arginine binding site. Each subunit of TS homodimer contributes two arginines to this phosphate-coordinating site (Figure 1.9b). Additionally, one of the phosphate oxygens is hydrogen bonded to a conserved Ser167<sup>2</sup> (Ser216). The deoxyribose of dUMP forms hydrogen bonds with His207 (His256) and Tyr209 (Tyr258) via the 3'-OH group. The C-2 of deoxyribose lies close to the backbone amide of Asp169 (Asp218) and the side chain of Ser167 (Ser216) and His207 (His256). These residues would constitute steric hindrance to the binding of nucleotides with a 2'-OH, which explains the low affinity for ribose nucleotides to TS (Montfort et al., 1990). The pyrimidine ring of dUMP is involved in an extensive hydrogen bond network in the active site including His147 (His196), Asp169 (Asp218), Asn177 (Asn226) and several ordered water molecules (Finer-Moore et al., 1993).

---

<sup>2</sup> The numbering scheme used here is based on *E. coli* TS. The numbering of hTS is given in parentheses.



**Figure 1.9. Crystal structure of *E. coli* TS complexed with dUMP (PDB: 1BID).** a) TS is a homodimer with dimer interface formed by a six-stranded twisted  $\beta$ -sheet from each subunit. b) A close view of dUMP binding site. The phosphate moiety of dUMP is coordinated by four arginines, with two of them from the other subunit (Arg126' and Arg127').

#### 1.4.2.2 Binding of folate

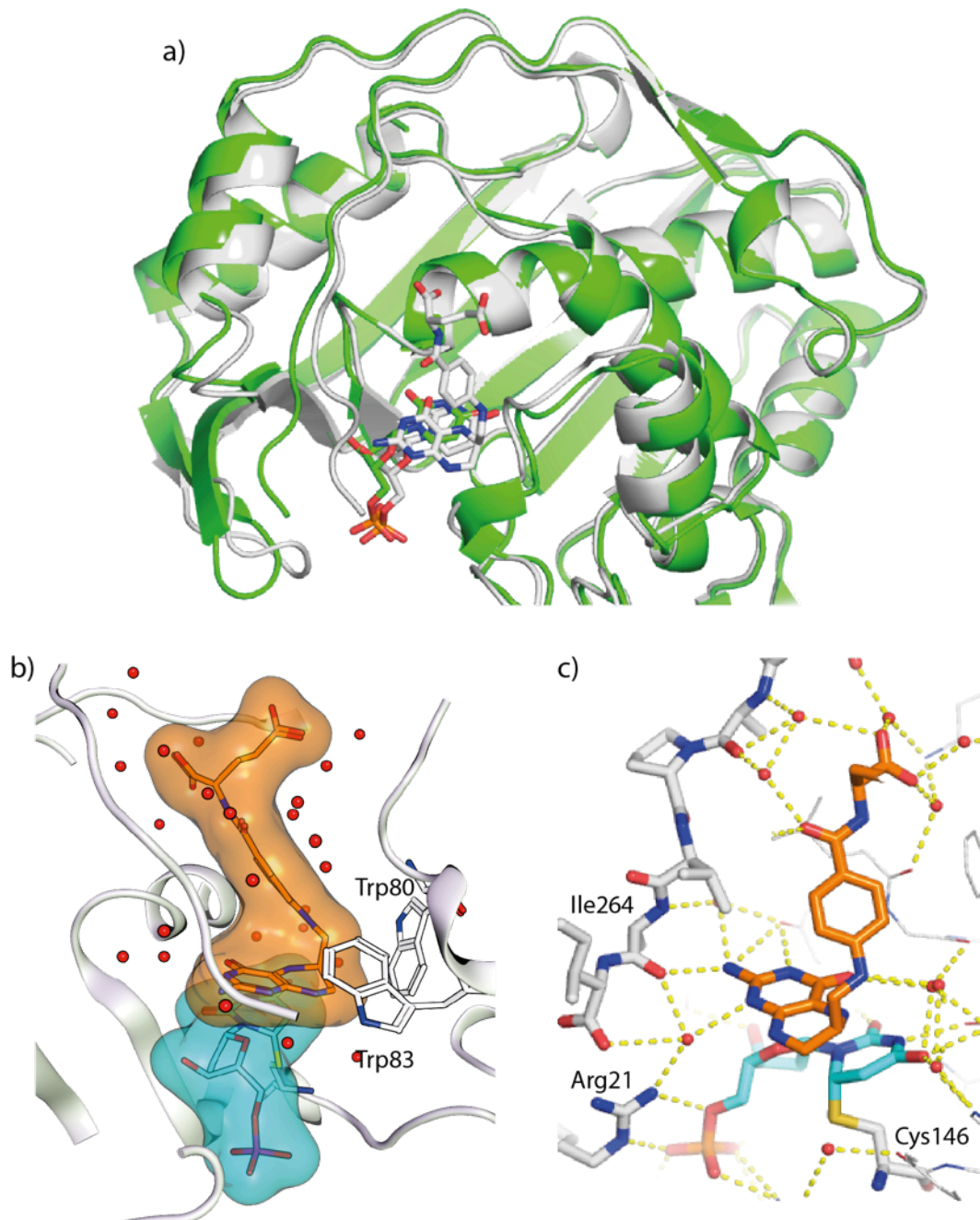
Many structures of the ternary complexes of TS together with nucleotides and folates or antifolates have been solved. Folate and antifolates in the ternary complexes adopt a similar folded conformation, in which the pterin ring or quinazoline ring (in some antifolates) is approximately perpendicular to the *para*-aminobenzoic acid (PABA) moiety (Fauman et al., 1994). The ternary complex is more ordered compared to the TS-dUMP binary complex. In response to the binding of folate, the enzyme undergoes a rearrangement of several protein segments by shifting toward the active site to maximize the interaction with folate. The largest change in position triggered by folate binding is the C-terminal loop, which moves toward the active site and serves as a lid to close the cavity (Figure 1.10a) (Kamb et al., 1992a; Montfort et al., 1990).

The folate binding site is composed of a large hydrophobic surface provided by dUMP, a solvent-exposed tunnel that contributes extensive water-mediated hydrogen bond interactions with the glutamate moiety, and the C-terminal residues in contact with both the pterin ring and the PABA moiety. The pterin ring maximizes its contact with the pyrimidine ring of dUMP via stacking (Figure 1.10b). The stacking interaction is further packed by massive interactions with hydrophobic residues,

including the conserved tryptophan residue, Trp80<sup>3</sup> (Trp109) (Montfort et al., 1990). Previous studies have shown that dUMP and folate bind TS in a sequential order (Finer-Moore et al., 1990). Binding of mono-glutamated folate requires the presence of dUMP in the catalytic cavity. This sequential binding indicates that the stacking interaction between dUMP and folate plays an essential role in folate binding. Following the conformational change, the C-terminal segment forms direct and water-mediated hydrogen bonds with the pterin ring. The C-terminal segment also interacts with two conserved residues in the active site, Arg21 (Arg50) and Asp169 (Asp218) via water-mediated hydrogen bonds, which stabilizes the closed conformation of the C-terminal loop (Figure 1.10c). The PABA moiety binds perpendicularly to the pterin ring, with the aromatic ring bound in a hydrophobic pocket and the glutamate ligated within nice water network hydrogen bonding extensively to the C-terminal segment. *In vivo*, cellular folates typically function in their polyglutamated forms, which are of higher affinity to folate-requiring enzymes as well as better retained inside cells (Shane, 1989). The polyglutamate chain adopts an extended conformation and binds TS in a positively changed binding site. It has been confirmed that polyglutamated folates can bind TS in the absence of nucleotides and induce similar conformational changes to the ternary complexes (Kamb et al., 1992a).

---

<sup>3</sup> The numbering scheme used is based on *E. coli* TS. The numbering of hTS is given in parentheses.



**Figure 1.10. Crystal structure of *E. coli* TS complexed with dUMP and THF.** a) Binding of folate THF induces conformational changes, mainly in the region of and near to the C-terminal loop. Green, *E. coli* TS-dUMP (PDB: 1BID); white, *E. coli* TS-dUMP-THF (PDB: 1KZI). b) A close view of dUMP and folate binding site. Both dUMP and THF are shown as surface. Stacking interaction between the pyrimidine and pterin ring is maximized. c) The glutamate moiety of THF interacts extensively with the C-terminal loop via water-mediated hydrogen bonds. The pterin ring forms direct or water-mediated hydrogen bonds with the residues in C-terminal loop.

### 1.4.2.3 Catalytic mechanism of TS

TS has been the subject of extensive efforts to dissect its catalytic mechanism using chemical models, kinetic studies, mutagenesis and structural information (Carreras and Santi, 1995; Finer-Moore et al., 1990; Montfort et al., 1990; Schiffer et al., 1995; Sotelo-Mundo et al., 2006). The overall mechanism of action was deduced before the three-dimensional structure of TS had been determined. The chemical steps in the reaction have been further validated with variant TS structures and site-directed mutagenesis. The mechanism of reaction catalyzed by TS can be divided into four events:

1. *Activation of C-5 of dUMP for substitution.* Upon the formation of a ternary complex of TS with dUMP and 5,10-CH<sub>2</sub>-THF, the thiol of Cys195<sup>4</sup> performs a nucleophilic attack at the C-6 of dUMP, which converts C-5 to a nucleophilic enol(ate) (intermediate I in Figure 1.11). This step can take place in the absence of 5,10-CH<sub>2</sub>-THF and is supported by the observation that 5-H of dUMP exchanges for solvent protons when folate is not bound (Pogolotti et al., 1979).
2. *Activation of 5,10-CH<sub>2</sub>-THF.* In the TS reaction, 5,10-CH<sub>2</sub>-THF is converted to its reactive form, the N-5 iminium ion, before undergoing reaction with nucleophilic intermediate I. The reactive iminium ion is formed by the opening of the imidazolidine ring of 5,10-CH<sub>2</sub>-THF. TS enzyme assists the ring opening by protonating N-10 via a general acid mechanism, which is proposed to be promoted by the conserved Glu87 based on the structural information (Hardy et al., 1995).
3. *Methylene transfer from 5,10-CH<sub>2</sub>-THF to dUMP.* With the two reactive species in proximity in the active site, covalent bond forms as a consequence of nucleophilic attack of C-5 of dUMP on the N-5 iminium ion of 5,10-CH<sub>2</sub>-THF. This leads to intermediate II in which C-5 of dUMP is attached to folate cofactor through a methylene bridge. Close analog of this bridged intermediate has been observed in the crystal structure of *E. coli* TS-FdUMP-5,10-CH<sub>2</sub>-THF complex (Hyatt et al., 1997).

---

<sup>4</sup> The numbering scheme used is based on human TS.



### **1.4.3 Targeting hTS in cancer chemotherapy**

#### **1.4.3.1 hTS-a chemotherapeutic target standing the test of time**

As the critical enzyme responsible for *de novo* biosynthesis of thymidylate, hTS is a key metabolic bottleneck in DNA synthesis required for proliferating cancer cells. This enzyme is vital to the proliferation and survival of cancer cells. Additionally, hTS is oncogenic and overexpressed in a variety of cancer cells (Rahman et al., 2004). hTS can also repress the expression of tumor-suppressor protein p53 (Ju et al., 1999). Therefore, hTS represents a very valuable intervention point in cancer chemotherapy. Since the 1940s, inhibition of hTS function has been explored and found to be effective in the treatment of various cancers (Figure 1.12).

Over the past few decades, the remarkable advance in our understanding of cancer biology has unfolded and led to a new era of anticancer drug development. A considerable fraction of drug development efforts has shifted away from the conventional strategy of cytotoxic chemotherapy to focus on the development of drugs targeting cancer-specific alterations. New therapeutic targets identified as behaving aberrantly in cancer cells include signal transduction proteins, cell cycle regulators, growth factors, apoptosis regulators and modulators involved in angiogenesis. One great success story is the discovery of imatinib, an inhibitor of BCR-ABL tyrosine kinase, in the treatment of chronic myeloid leukemia (CML) (Toga et al., 2003). However, despite the enormous progress in our knowledge of biology in many cancer types and the overwhelming promises of new drug targets, the development of anticancer drugs still suffers from a very high failure rate in clinical trials. On the other hand, as new cancer-specific therapeutic targets gain attention, the use of chemotherapeutics such as hTS-targeted anticancer agents is expected to decline gradually in the coming years. Still, there are indications on the contrary, that hTS-targeted drugs not only remain as the integral components of therapeutic regimes in some of the most-difficult-to-treat cancers (e.g. pancreatic cancer, colorectal cancer and non-small-cell lung cancer), but also expand in terms of clinical use with new approvals and indications (Cohen et al., 2010; Conroy et al., 2011; Hazarika et al., 2004). Moreover, advancement in our understanding of the nucleotide biosynthesis and DNA repair pathways at the molecular level has led to the development of

rational approaches for pathways-based combination therapies targeting thymidylate biosynthesis (Emura et al., 2004; Van Cutsem et al., 2011).

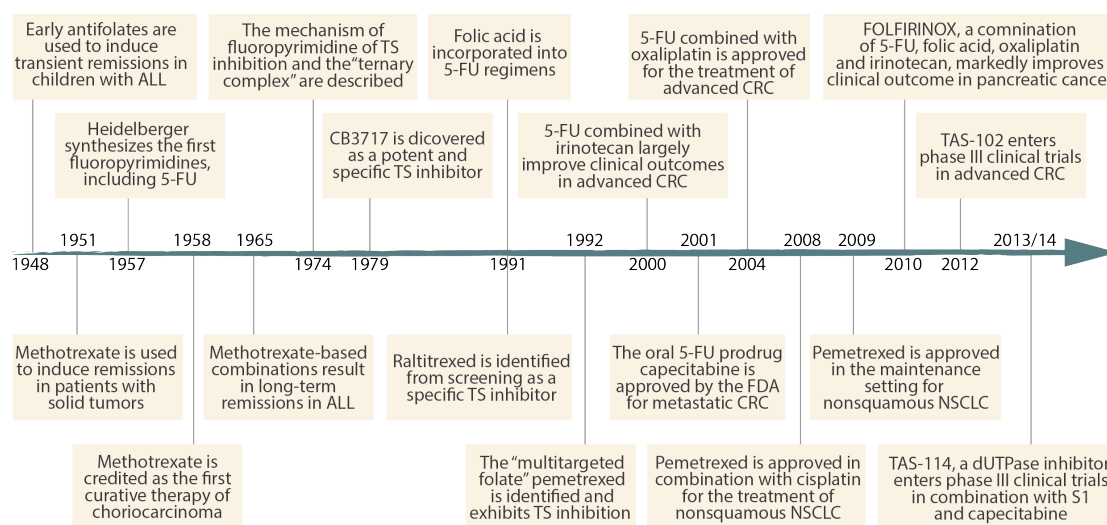
#### **1.4.3.2 hTS inhibitors and their modes of action**

The TS enzyme consists of two distinct binding sites for the nucleotide substrate and the folate co-substrate. These two binding sites have both been targets for the design of competitive inhibitors disrupting hTS activity, leading to the development of two structurally diverse classes of hTS inhibitors, nucleotide and folate analogs.

Fluoropyrimidine, the major hTS inhibitor in the class of nucleotide analogs, was developed in 1957 by Heidelberger (Heidelberger et al., 1957). 5-fluorouracil (5-FU) is the first and best-known fluoropyrimidine. It also represents one of the earliest examples of rationally designed drugs (Santi et al., 1974). Spurred by the notable success of 5-FU, the development of effective fluoropyrimidine-based hTS inhibitors continued with the introduction of fluoropyrimidines possessing improved pharmacokinetic properties, tumor specificity and toxicity profile, such as capecitabine (Ishikawa et al., 1998) and S-1 (Saif et al., 2009). Most fluoropyrimidines are prodrugs of 5-FU that rely on the metabolic activation via one or more enzymatic reactions. 5-FU is further converted to several active metabolites by cellular enzymes, including fluorodeoxyuridine monophosphate (FdUMP), fluorodeoxyuridine triphosphate (FdUTP) and fluorouridine triphosphate (FUTP) (Figure 1.13) (Miller, 1971). The antitumor effect of 5-FU, however, is mediated through inhibition of hTS by FdUMP (Santi et al., 1974) and misincorporation of its metabolites (FdUTP and FUTP) into DNA and RNA (Kufe and Major, 1981; Noordhuis et al., 2004).

The other class of hTS inhibitors is folate analogs or antifolates. It was in the late 1940s that inhibition of dTMP synthesis was found to be one of the major mechanisms of action of antifolates for their anticancer activity (Chabner and Roberts, 2005; Farber et al., 1948). A few years later, methotrexate proved to have curative effects on patients with choriocarcinoma and led to complete remissions in patients with acute lymphoblastic leukemia (ALL) (Freireich, 1967). This clinical success prompted the development of more potent hTS-targeted antifolates. For

example, raltitrexed (Tomudex®) was discovered to inhibit hTS specifically (Jackman et al., 1991). Similar to methotrexate, raltitrexed is transported into cells via the reduced folate carrier (RFC) and undergoes polyglutamation by folylpolyglutamate synthase (FPGS) (Jackman et al., 1993). Upon polyglutamation, the potency of raltitrexed increases by about 100-fold and its cellular retention is very much extended. Although not yet approved in the USA, due to a lack of survival benefit, raltitrexed is used clinically for different indications in many other countries. Following the discovery of raltitrexed, the multi-targeted folate pemetrexed was developed. Pemetrexed inhibits not only hTS but also several other folate-requiring enzymes including DHFR, and GARTase and AICARTase involved in purine nucleotide biosynthesis (Curtin and Hughes, 2001). Different from other antifolates, cellular uptake of pemetrexed involves two transporters, RFC and proton-coupled folate transporter (PCFT), which has important implications for the development of resistance.

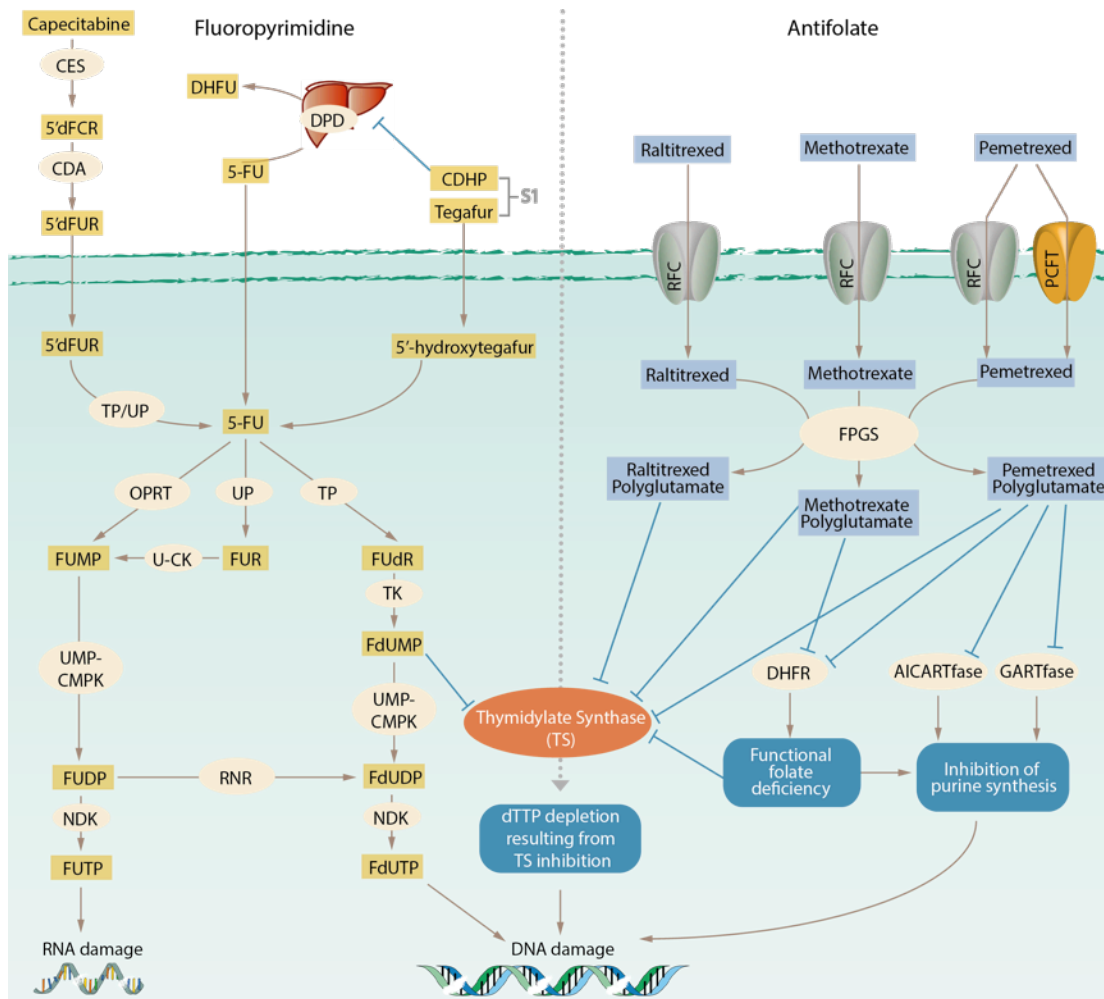


**Figure 1.12. Historical development in targeting hTS in cancer chemotherapy.** The regulatory approvals described here refer to the approvals made by the FDA in USA. Adapted from (Wilson et al., 2014).

Abbreviations: 5-FU, 5-fluorouracil; ALL, acute lymphoblastic leukemia; CRC, colorectal cancer; dUTPase, deoxyuridine triphosphate nucleotidohydrolyase; NSCLC, non-small-cell lung cancer.

Molecular downstream effects of hTS inhibition are mediated by numerous mechanisms depending on the action mode of inhibitor and the cancer type (Figure 1.13). The common immediate effect of hTS inhibition is depletion of the dTMP pool,

which results in subsequent depletion of dTTP and a concomitant increase in dUMP and dUTP levels (Goulian et al., 1986). Accumulation of cellular dUTP causes an increased level of dUTP misincorporation into DNA and thus severe disruption of DNA synthesis. Furthermore, depletion of dTTP can induce perturbations in the levels of the other dNTPs, which severely disrupt DNA synthesis and repair as well. As a consequence of DNA damage, cell cycle halts and often cell death via apoptosis is initiated.



**Figure 1.13. Mechanism of action of fluoropyrimidines and antifolates.** 5-FU is converted to three main active metabolites: FdUMP, FdUTP and FUTP. FdUMP inhibits TS and leads to rapid depletion of dTTP. FdUTP and FUTP can be misincorporated into DNA and RNA, respectively. Capecitabine and S1 are oral fluoropyrimidines, both of which are prodrugs of 5-FU. Capecitabine is converted to 5'dFUR in the liver, which is further converted to its active form 5-FU. Tegafur is the cytotoxic component of S1, which is also converted to 5-hydroxytegafur in the liver and subsequently to 5-FU. 80% of synthetic 5-FU is subject to degradation by DPD in the liver. S1 contains CDHP and prevents DPD-mediated hepatic degradation of 5-FU. Raltitrexed and methotrexate enter the cell primarily via RFC, while pemetrexed can also be transported via PCFT. Their polyglutamated forms result in dTTP depletion and inhibition of folate recycling and purine synthesis. Adapted from (Wilson et al., 2014).

Abbreviations: 5'dFCR, 5'-deoxy-5-fluorocytidine; 5'dFUR, 5'-deoxy-5-fluoro-uridine; 5-FU, 5-fluorouracil; AICARTfase, 5-aminoimidazole-4-carboxamide ribonucleotide transferase; CDHP, 5-chloro-2,4-dihydropyridine; DHFU, dihydrofluorouracil; DPD, dihydropyrimidine dehydrogenase; dUTPase, deoxyuridine triphosphate nucleotide-hydrolase; FdUDP, fluorodeoxyuridine diphosphate; FdUMP, fluorodeoxyuridine monophosphate; FdUTP, fluorodeoxyuridine triphosphate; FPGS, foyllypolyglutamate synthetase; FUDP, fluorouridine diphosphate; FdR, 5'-fluoro-deoxyuridine; FUMP, fluorouridine monophosphate; FUR, fluorouridine; FUTP, fluorouridine triphosphate; GARTfase, glycinamide ribonucleotide transferase; NDK, nucleoside diphosphate kinase; OPRT, orotate phosphoribosyltransferase; PCFT, proton-coupled folate transporter; RFC, reduced folate carrier; RNR, ribonucleotide reductase; TK, thymidine kinase; TP, thymidine phosphorylase; U-CK, uridine-cytidine kinase; UMP-CMPK, uridine monophosphate-cytidine monophosphate kinase; UP, uridine phosphorylase.

### 1.4.3.3 Drug resistance to hTS inhibitors

Drug resistance continues to be a major impediment in cancer chemotherapy. The resistance mechanisms to cytotoxic anticancer agents are mainly found at three levels: the drug availability; target availability for drug action; and the ability of drug-target interaction to induce lethal damage in cancer cells.

Given that numerous branch-points and enzymes are involved in the metabolism of hTS inhibitors (both fluoropyrimidine and antifolate inhibitors) (Figure 1.13), it is not surprising that the long-term effectiveness of these drugs is more tumor-dependent and hindered by the development of resistance. Several metabolic enzymes have been implicated in the altered drug metabolism and thus resistance to fluoropyrimidines or antifolates. For example, dUTPase, which hydrolyzes dUTP to dUMP, is overexpressed in a number of cancers to protect tumor DNA from uracil-misincorporation and circumvent the subsequent cytotoxicity induced by fluoropyrimidines (Canman et al., 1993). As most antifolates are dependent on transporter proteins for cellular uptake and polyglutamation for activation and cellular retention, alterations of antifolate metabolism at these two steps render resistance to antifolates. The well-known resistance mechanisms are decreased uptake due to loss of RFC function (Kobayashi et al., 1998; Moscow et al., 1997; Rothem et al., 2003), impaired polyglutamation due to defective FPGS activity (Liani et al., 2003; Pizzorno et al., 1989) and increased  $\gamma$ -glutamyl hydrolase activity (Li et al., 1993), and increased antifolate efflux due to overexpression of ATP-driven multidrug resistance (MDR) efflux transporters (Assaraf, 2006).

Modification of drug targets can also convey resistance to anticancer agents. There are a few mutations of hTS identified to constitute resistance to certain hTS inhibitors (Barbour et al., 1990; Davis and Berger, 1993; Tong et al., 1998). Apart from that, the relationship between hTS expression levels and therapeutic responses has been investigated. In general, resistance to hTS-inhibitor-based therapy is associated with increased levels of hTS expression (Liu et al., 2013; Peters et al., 2002; Qiu et al., 2008). The proposed mechanism of treatment-induced elevation of hTS protein levels is the disruption of the auto-regulatory translation of hTS mRNA, whereby binding of hTS inhibitors liberates hTS from the hTS-mRNA complex and relieves translational repression (Chu et al., 1991). Nevertheless, a substantial

proportion of the tumors with low hTS expression do not respond to 5-FU treatment (Showalter et al., 2008). The overall relationship between tumor hTS expression levels and resistance and/or therapeutic has not been materialized.

Independent of the factors above, downstream cellular response to hTS inhibitors constitutes an established determinant of therapy outcomes. As presented in Figure 1.13, the primary cytotoxic effect of hTS inhibition is mediated through DNA damage-induced cell death. Therefore, enhanced activity of DNA repair machinery contributes to chemotherapy resistance. In support of this, some enzymes involved in the DNA repair pathway are overexpressed in various types of cancers and demonstrated to be associated with resistance to hTS-inhibitors-based chemotherapy, such as apurinic-apyrimidinic endonuclease 1 (McNeill et al., 2009) and uracil DNA glycosylase (Weeks et al., 2013). Generally, apoptosis has been considered as the ultimate mechanism of cell death resulting from chemotherapy. A common mechanism of resistance to cytotoxic agents including hTS inhibitors is the suppression of pathways leading to apoptosis. Though the clinical data is quite limited, p53 and p53-regulated genes in the apoptotic pathway have been implicated in resistance to hTS-targeted chemotherapy (Houghton et al., 1997; Longley et al., 2002; Tillman et al., 1999; Zhang et al., 2000).

Owing to the inherent genetic instability of cancer cells, they are able to generate a wide variety of resistance mechanisms to cytotoxic agents. With our improving understanding in the mechanisms of drug resistance at the molecular level, we are potentially able to devise strategies to circumvent certain mechanisms of drug resistance. Most attempts to overcome resistance involve the development of rationally designed new drugs and the use of combination drug regimes (Fojo and Bates, 2003; Holohan et al., 2013). Despite significant expansion in our knowledge of resistance mechanisms to hTS-targeted therapies, many unanswered questions remain, such as the relevance of drug-induced elevation of hTS protein and alteration of downstream pathways as predictive biomarkers for therapeutic responses. It is reasonable to expect that as we progress further in understanding these mechanisms, we should be able to formulate more efficient strategies to tackle resistance problems in cancer therapy.

## **1.5 Aims of the project**

The enzyme hTS is a valuable drug target for cancer chemotherapy. Many hTS inhibitors have been developed and several have been approved and remain as the integral components in the treatment of many cancers. However, efficacy of these anticancer agents is largely limited by the development of drug resistance. In this project, fragment-based approach is applied to hTS, setting out to identify novel starting points for structure-based drug design and overcoming drug resistance. In order to have a better understanding of hTS system and its binding sites for the following drug design, a comprehensive study of the binary and ternary complexes of hTS is performed driven by a combination of biophysical techniques. Previous studies have shown that hTS exists in an inactive conformation, while the biological relevance of this inactive state remains unclear. Therefore, another aim of this work is to gain insights into the potential roles of this inactive conformation of hTS, including its correlation with the mRNA binding activity of hTS in particular.

## CHAPTER 2 MATERIALS AND METHODS

### 2.1 Fragment library

Fragment library was purchased from Maybridge (Thermo Fisher Scientific Inc.). The size of the first library is 500 compounds and the second is 1500. The molecular weights of the compounds are in the range of 90 Da to 300 Da, with the cLogP values between -2.5 and 3.0. The compounds were dissolved in 100% dimethyl sulfoxide (DMSO) at 200 mM and stored at room temperature.

### 2.2 Molecular cloning and expression

Molecular cloning was performed by protein production platform in School of Biological Sciences, Nanyang Technological University, using method as described (Lim et al., 2013). Briefly, the gene encoding human TS (GeneBank accession code: NM\_001071.2) was subcloned into pNIC28-Bsa4, pNIC-CH2 or pNIC-CTHF vectors using ligation-independent cloning (Graslund et al., 2008). All gene constructs were expressed in Rosetta BL21-DE3 *Escherichia coli* (Novagen) in Terrific Broth media supplemented with 50 µg/ml of kanamycin and 34 µg/ml Chloramphenicol. Cells were grown at 37 °C until OD<sub>600</sub> reached about 2.0 and induced with 0.5 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG) at 18 °C overnight. The cells were harvested by centrifugation at 4,500 x g for 15 min at 15 °C. Cell pellet was re-suspended in lysis buffer [100 mM HEPES, 500 mM NaCl, 10 mM imidazole, 10% (v/v) glycerol and 1 mM tris-(2-carboxyethyl)phosphine (TCEP), pH 8.0] supplemented with 1 µl/ml ethylenediaminetetraacetic acid (EDTA)-free protease inhibitor cocktail (Calbiochem) and 125 U/ml of Benzonase (Merck). Cells were lysed by sonication on ice at 70% amplitude, 3 s on/off for 3 min. The lysate was clarified by centrifugation at 47,000 x g for 25 min at 4 °C, and the supernatant was filtered through a 1.2 µm syringe filter to remove cell debris.

### 2.3 Mutagenesis

The mutant constructs were generated by PCR-based methods using KOD Xtreme™ Hot Start DNA polymerase (Millipore) according to manufacturer's instructions. The sequence of forward primer used for M190K construct was 5'-CCAAGAGATCTTCTCTGAAGGCGCTGCCT-3'. The sequence of forward

primer used in A191K site-directed mutagenesis was 5'-GAGATCTTCCTCTG-ATGAAGCTGCCTCCATGCCATG-3'. All mutant constructs were verified by DNA sequencing.

## 2.4 Protein purification

All human TS constructs were purified using immobilized metal affinity chromatography (IMAC) followed by gel filtration chromatography on ÄKTExpress system (GE Healthcare).

The cell-free extract was loaded on a pre-equilibrated HisTrap<sup>TM</sup> HP column (GE Healthcare) in IMAC wash buffer 1 [20 mM HEPES, 500 mM NaCl, 10 mM imidazole, 10% (v/v) glycerol and 1 mM TCEP, pH 7.5] and subsequently washed with 20 column volumes (CVs) of IMAC wash buffer 1 and 15 CVs of IMAC wash buffer 2 [20 mM HEPES, 500 mM NaCl, 25 mM imidazole, 10% (v/v) glycerol and 1 mM TCEP, pH 7.5]. Bound protein was eluted with 5 CVs of elution buffer [20 mM HEPES, 500 mM NaCl, 500 mM imidazole, 10% (v/v) glycerol and 1 mM TCEP, pH 7.5] and loaded onto a HiLoad 16/60 Superdex-200 column (GE Healthcare) pre-equilibrated with HEPES buffer [20 mM HEPES, 300 mM NaCl, 10% (v/v) glycerol, and 1 mM TCEP, pH 7.5] or phosphate buffer [100 mM PO<sub>4</sub>, 100 mM NaCl, and 2 mM TCEP, pH 7.4]. The protein fractions collected were analyzed on 4-12% Nu-polyacrylamide gel electrophoresis (PAGE) gels (Invitrogen). Fractions with high protein purity were pooled and incubated with mutant (S219V) TEV-protease with a molar ratio 1: 20 (protease: target protein) to remove hexa-histidine tag at 4 °C overnight. The samples were analyzed on Nu-PAGE gel to check its cleavage completeness. Another round of IMAC was applied to tag-cleaved protein sample, where the flow through from the column and the sample from washing steps were collected and loaded to a pre-equilibrated HiLoad 16/60 Superdex-75 column (with buffer A or buffer B) (GE Healthcare). Based on Nu-PAGE gel result, pure protein fractions were pooled and concentrated using a 50 kDa cut-off centrifugal driven filter concentrator (Satorius Stedium Biotech). The protein concentration was determined by the absorbance at 280 nm using Nanodrop spectrophotometer (Thermal Scientific). Protein was aliquoted, frozen in liquid nitrogen and stored at -80 °C.

## **2.5 Size-exclusion chromatography (SEC) coupled with multi-angle light scattering (MALS)**

SEC-MALS was performed using a superdex 200 15/150 GL column (GE Healthcare) combined with a miniDAWN TREOs light-scattering detector coupled with an Optilab rEX refractive index detector (Wyatt Technology). All experiments were conducted at room temperature at a flow rate of 0.3 ml/min. The injected protein sample was typically 20  $\mu$ l of 6 mg/ml of protein. Molecular mass calculations were performed using the Astra6.1 software (Wyatt Technology). Input of the refractive increment (dn/dc values) was set at 0.186 ml/g in the molecular mass calculations. The molecular mass was determined across the protein elution peak.

## **2.6 Enzymatic assay**

Enzymatic activity was measured spectrophotometrically at 340 nm by monitoring the absorbance change during the conversion of CH<sub>2</sub>H<sub>4</sub>folate to H<sub>2</sub>folate using Infinite M200 spectrometer (Tecan). Measurements were carried out at room temperature and in a buffer of 50 mM Tris pH 7.5. For steady-state kinetic measurement, initial velocities were measured by utilizing 100 nM of purified protein, 100  $\mu$ M 5,10-CH<sub>2</sub>-THF and 15.625-100  $\mu$ M dUMP. Kinetic constants and initial rates were analyzed with the software package Prism (GraphPad Software).

## **2.7 Differential scanning fluorimetry (DSF)**

DSF was performed on the iCycler iQ Real Time PCR Detection System (Bio-Rad), using the 96-well thin-wall PCR plate (Bio-Rad). The experiment was conducted in buffer 20 mM HEPES pH 7.5 and 150 mM NaCl. A total volume of 25  $\mu$ l solution containing 0.2 mg/ml protein, compounds and 5 x Sypro Orange (Invitrogen) diluted from 5000 x stock was dispensed into each well of the 96-well plate using multi-channel pipette. The same amount of DMSO was added instead of compounds in the control wells. The plates were sealed with Microseal B adhesive sealer (Bio-Rad) and heated in iCycler from 25 to 80 °C (56 heating cycles in 28 min). Fluorescent filter was used for Sypro Orange with  $\lambda_{\text{excitation}} = 492$  nm and  $\lambda_{\text{emission}} = 610$  nm. The calculation of the midpoint of the curves ( $T_m$ ) was performed using software package XLfit (ID Business Solutions).

## 2.8 Differential static light scattering (DSLS)

Temperature-dependent protein aggregation was monitored using DSLS, on Stargazer instrument (Harbinger Biotech). 45  $\mu\text{l}$  of protein (0.2 mg/ml) with compounds covered with 45  $\mu\text{l}$  of mineral oil (Sigma-Aldrich) was heated at 1  $^{\circ}\text{C}$  per min, from 25  $^{\circ}\text{C}$  to 80  $^{\circ}\text{C}$ , in each well of a 384-well black plate (Nunc). The images were taken every 0.5  $^{\circ}\text{C}$ , with 100 msec exposure time. The pixel intensities in a preselected region of each well were integrated and plotted against temperature. The plotting was fitted to the Boltzmann equation by nonlinear regression and the inflection point of each resulting curve was calculated as  $T_{\text{agg}}$ .

## 2.9 Isothermal titration calorimetry (ITC)

Protein was dialyzed with a buffer containing 20 mM HEPES pH 7.5 and 150 mM NaCl at 4  $^{\circ}\text{C}$  overnight. The protein was loaded into the iTC200 (Microcal) cell (cell volume about 200  $\mu\text{l}$ ). The titration syringe (volume about 40  $\mu\text{l}$ ) was filled up with compound solution, which was diluted with the dialysis buffer. The protein and compound sample were matched in terms of buffer condition and DMSO concentration. Titrations were initiated with one 0.5  $\mu\text{l}$  injection, followed by 16 to 20 larger volume injections, injected at 180 s intervals, at 25  $^{\circ}\text{C}$ . Stirring speed was 700 rpm. The heat peaks integration and non-linear regression analysis were performed with the Origin software (Microcal) and fitted to a single-site binding model to determine the parameters stoichiometry  $N$ ,  $K_D$  and  $\Delta H$ .

## 2.10 Crystallization

All protein constructs of hTS were set up for crystallization screening in sitting drops using Phoenix liquid handling system (Art Robbins Instruments). Promising crystallization conditions were subject to further optimization.

hTS vc1 crystallized in a buffer of 0.075 M HEPES pH 7.5, 15% PEG10000 and 25% glycerol as thin needles. During the optimization course, crystals were grown by sitting drop vapor diffusion method, in a drop ratio of 1: 3 (protein: reservoir solution) at room temperature. Streak seeding was applied to obtain three-dimensional crystals.

Crystal form 1 of hTS vc5 was grown using sitting drop vapor diffusion method at 20  $^{\circ}\text{C}$ . The protein solution (20 mg/ml) was mixed in a 1: 2 ratio with

crystallization buffer to set up drops of total volume of 0.3  $\mu$ l and the reservoir contained 50  $\mu$ l of crystallization buffer in Intelli-plate 96-3 (Art Robbins Instruments). The crystallization buffer was composed of 0.2 M NaCl, 0.1 M Bis-Tris pH 5.5 and 25% PEG 3350.

Crystal form 2 of hTS vc5 was obtained during the optimization process of crystallization condition of crystal form 1. Crystals of different morphology grew in a condition of different pH (0.2 M NaCl, 0.1 M Bis-Tris pH 7.5 and 23% PEG 3350).

The initial crystallization buffer from screening for crystal form 3 of hTS vc5 contained 0.1 M sodium cacodylate, pH 6.5 and 25% PEG 4000. Protein crystallized in sitting drops comprising equal volume of protein (about 24 mg/ml) and reservoir solution at 20 °C. After optimization and applying streak seeding, crystals of better quality grew in a buffer with lower precipitant (0.1 M sodium cacodylate, pH 6.5 and 10% PEG 4000) in sitting drops using Intelli-plate 24-4 (Art Robbins Instruments) at room temperature. Co-crystals of hTS-dUMP was obtained using protein sample containing 2 mM dUMP and 5% DMSO.

### **2.11 Crystal soaking and freezing**

Before data collection, crystals were transferred to the corresponding reservoir solution supplemented with 10% glycerol as cryo-protectant. In the case of soaking using crystal form 3 of hTS vc5, compounds were dissolved in 100% DMSO to the desired concentrations and diluted 10 times with a buffer of 0.1 M sodium cacodylate, pH 6.5 and 15% PEG 4000. The final soaking buffer included 10% DMSO and 15% PEG 4000, which worked well as a cryo-protectant. Crystals were transferred into the soaking buffer and incubated for 10 to 15 min at room temperature, followed by dipping into liquid nitrogen.

### **2.12 Data collection and structure determination**

X-ray diffraction data were collected on the beamline MX1 and MX2 at Australian Synchrotron, and on the 13B1 and 13C1 beamline at National Synchrotron Radiation Research Center (NSRRC), Taiwan. Collected data were indexed and processed with the HKL2000 software (Otwinowski and Minor, 1997). The structures of hTS were solved by molecular replacement using Phaser (McCoy et al., 2007) with the hTS-dUMP-raltitrexed structure (PDB code 1I00) as the search model. All the

solvent molecules and ligands were removed prior to molecular replacement. Model building and structure refinement were carried out using ARP/wARP (Langer et al., 2008) and REFMAC5 (Murshudov et al., 2011), respectively. The model was checked and manually improved in Coot (Emsley and Cowtan, 2004). Structure validation was carried out using Ramachandran plot (Ramachandran et al., 1963). Ligand structures and restraints files were generated using Phenix.eLBOW (Moriarty et al., 2009).

### **2.13 Ribonucleic acids (RNAs)**

All RNAs used were from Sigma-Aldrich. RNAs were dissolved in a buffer containing 10 mM sodium cacodylate, pH 6.5 and 50 mM NaCl. Before usage, RNAs were heated at 95 °C for 10 min and subsequently annealed on ice for 15 min.

### **2.14 Native PAGE of protein-RNA complexes**

Precast NuPAGE® Novex® 4-12% Bis-Tris gels (Life Technologies) were used for native PAGE analysis of hTS-RNA complexes. The electrophoresis buffer contained 45 mM Tris, 45 mM boric acid and 1 mM EDTA. RNAs were incubated with protein for 30 min on ice before loading. Samples were loaded in a buffer containing 20 mM HEPES, pH 7.5, 150 mM NaCl and 10% glycerol. RNAs were visualized with SYBR® Safe DNA gel stain (Life Technologies). Protein was visualized with coomassie blue stain (Sigma-Aldrich).

### **2.15 Panning of the phage display library**

The phage display library of single-heavy chain variable-fragment antibody (VH) used was generated by our PEST team. Our target protein M190K (10 µg/ml) was coated onto an immune tube (Nunc) in Tris-buffered saline (TBS) and incubated at 4 °C overnight. The immune tube was then blocked with TBS containing 4% skim milk (TBSM) for 1 hour on roller at room temperature. The VH phage library was precipitated from solution using 20% polyethylene glycol (PEG) 8000 and 2.5 M NaCl and blocked with TBSM for 1 hour on roller at room temperature. When blocking of immune tube and phage library completed, wash the immunotube with TBS twice and transfer the blocked phage to immune tube. After incubation on roller for 1.5 hours at room temperature, wash the tube three times with TBSM, three times with TBS with 0.05% Tween 20 (TBST) and then twice with TBS to remove unbound

phage. After washes, add trypsin solution (1 mg/ml in TBS supplemented with 2 mM  $\text{CaCl}_2$ ) for 1 hour at room temperature to elute phage. The trypsinized phage was transferred to 5 ml TG-1 culture grown in 2xYT medium with an  $\text{OD}_{600}$  about 0.5 and incubated at 37 °C for 30 min in a water bath. The TG-1 culture was then diluted with 45 ml of 2xYT medium supplemented with 2% glucose in a conical flask and incubated in a shaker at 37 °C until  $\text{OD}_{600}$  reached about 0.5. The TG-1 cells were infected with 20  $\mu\text{l}$  of M13K07 helper phage (New England Biolabs) by incubating at 37 °C for another 30 min in a water bath. Centrifuge the culture at 4,000 x g for 10 min and resuspend the pellet in 0.5ml 2xYT medium. The resuspended TG-1 cells were plated on YT-agar plate supplemented with 100  $\mu\text{g}/\text{ml}$  carbenicillin and 25  $\mu\text{g}/\text{ml}$  kanamycin and incubated at 30 °C overnight. On the second day, soak the agar plate with 25 ml TBS for 15 min and scrape cells from the plate with a plastic spreader. Centrifuge at 4,800 x g for 15 min and transfer the supernatant containing the phage to 20% PEG 8000 and 2.5 M NaCl to incubate on ice for 15 min. Centrifuge the sample at 4,800 x g for 15 min again and resuspend the phage pellet in 10 ml TBS. Repeat phage precipitation using PEG solution and resuspend the phage pellet in 1 ml TBS supplemented with 20% glycerol. Prepare 1 ml of phage ( $\text{OD}_{268}$  about 1.0) from round 1 panning in TBSM for the next panning. Repeat the panning process of binding, elution and infection steps three times to enrich binding phage.

## **2.16 Phage enzyme-linked immunosorbent assay (ELISA)**

After completing four rounds of panning, enriched clones from each round were tested for binding to M190K and hTS using polyclonal phage ELISA. An ELISA plate was coated with 100  $\mu\text{l}$  of protein in TBS (10  $\mu\text{g}/\text{ml}$ ) at 4 °C overnight, followed by blocking with TBSM and subsequent washing with TBST. Phage from each round diluted in TBSM ( $\text{OD}_{268}$  in the range of 0.1-1) was added to the plate and incubated for 1 hour at room temperature. After washing with four times TBST, the bound phage was detected with an anti-M13 antibody conjugated to HRP at a 1:10,000 dilution in TBSM. 3,3',5,5'-Tetramethylbenzidine (TMB) (100  $\mu\text{g}/\text{ml}$ ) was then used for color development, and the reaction was stopped after 10 min by addition of 2 M  $\text{H}_2\text{SO}_4$ . Absorbance at 405 nm was measured with Infinite M200 spectrometer (Tecan).

After deciding which round of phage to use based on polyclonal ELISA data,

TG-1 cell glycerol stock of the chosen round was plated onto a YT-agar plate supplemented with 100 µg/ml carbenicillin and incubated at 30 °C overnight. On the second day, 190 single colonies were picked from the plate and dispensed into the 2xYT medium supplemented with 2% glucose and 100 µg/ml carbenicillin (500 µl per well) in individual wells of two 96-deep well plates. These monoclonal starter cultures were incubated at 37 °C in a plate shaker overnight and 10 µl of each was then transferred into two new 96-deep well plates containing 200 µl of 2xYT medium supplemented with 2% glucose and 100 µg/ml carbenicillin per well. The cells were infected with 200 µl of 200-fold diluted M13K07 helper phage per well by incubating at 37 °C for 1 hour in plate shaker. The cell cultures were then incubated at 37 °C overnight in plate shaker after adding 400 µl of 2xYT medium supplemented with 100 µg/ml carbenicillin and 50 µg/ml kanamycin per well. After incubation, the culture was pelleted by centrifugation at 4,000 x g for 10 min. Monoclonal ELISA was performed with 50 µl of supernatant from each well blocked with 50 µl TBSM.

### **2.17 Co-Immunoprecipitation (CoIP) of VH-hTS complexes**

The Thermo Scientific Pierce Recombinant Protein A Agarose was used for CoIP of VH-hTS complexes. The assay was performed according to manufacturer's recommendations. In brief, VH binders were incubated with target proteins in a buffer of 25 mM Tris, pH 7.2 and 150 mM NaCl at 4 °C overnight. The samples were loaded to equilibrated Protein A resin slurry and incubated with gentle mixing for 2 hours at room temperature. The resin was washed three times with assay buffer and then eluted with 0.1 M glycine-HCl buffer, pH 2.5. The eluates were resolved on SDS-PAGE and visualized with coomassie blue stain.

### **2.18 Pull-down assay of VH-hTS complexes**

Pull-down assay of VH-hTS complexes was conducted using Thermo Scientific Pierce Recombinant Protein A Agarose in a 96-well filter plate. Briefly, VH binders in a buffer of 25 mM Tris, pH 7.2 and 150 mM NaCl were added into equilibrated Protein A resin slurry and incubated with gentle mixing at room temperature for 15 min. The flow through from each sample was collected by spinning the plate at 200 x g for 20 sec. The resin was then washed twice with a buffer of 50 mM Tris, pH 7.2, 150 mM NaCl and 0.02% Tween 20 before the bound

complexes were eluted with 0.1 M glycine-HCl buffer, pH 2.5. Both the flow through and eluate sample from each well were resolved on SDS-PAGE and visualized with coomassie blue stain.

### **2.19 Cellular thermal shift assay (CETSA)**

The CETSA experiments were conducted as previously described (Molina et al., 2013). Human cancer cell line HeLa S3 cells (ATCC no. CCL-2.2) were cultured in Dulbecco's Modified Eagle Medium (DMEM, Life Technologies). All culture media were supplemented with 10% fetal bovine serum (FBS, Life Technologies), 100 units/ml penicillin, 100 units/ml streptomycin and 100 units/ml neomycin (Life Technologies). Equal cells (about 1.0 million/data point) were used for the following CETSA experiments.

For the cell lysate CETSA experiments, cultured cells were harvested and washed with phosphate-buffered saline (PBS). The cells were diluted in PBS buffer supplemented with protease inhibitor cocktail. The cell suspensions were subsequently freeze-thawed three times using liquid nitrogen. The soluble fraction (lysate) was separated from the cell debris by centrifugation at 20,000 x g for 20 minutes at 4 °C. The cell lysates were diluted with PBS buffer supplemented with a final concentration of 100 µM dUMP and divided into two aliquots, with one being treated with compound (1 mM **A27**) and the other with equal amount of DMSO (1%) as control. After 20 minutes incubation at room temperature the lysates were divided into smaller (100 µl) aliquots and heated individually at desired temperatures for 3 minutes on thermal cycler machine (Life Technologies) followed by cooling for 3 minutes at room temperature. The heated lysates were centrifuged at 20,000 x g for 20 minutes at 4 °C. The supernatants were transferred to new microtubes and analyzed by SDS-PAGE followed by western blot analysis. Primary antibody used was anti-TS (sc-33679, Santa Cruz Biotechnology) and secondary antibody was goat anti-mouse HRP-IgG (sc-2055, Santa Cruz Biotechnology).

For whole cell CETSA experiments, the cells were exposed to the compound for 3 hours in an incubator chamber at 37 °C and 5% CO<sub>2</sub>. Control cells were treated with an equal volume of DMSO. Following treatment the cells were harvested and washed with PBS in order to remove excess drug. Equal amounts of cell suspension in

PBS buffer supplemented with protease inhibitor cocktail were aliquoted into 0.2 ml PCR microtubes and were subjected to heating as described above followed by lysing using 3 cycles of freeze-thawing with liquid nitrogen. The soluble fractions were isolated and analyzed by western blot.

## **2.20 Compound screening on hTS using CETSA**

The screening of a library containing 10,398 compounds using CETSA was performed by our collaborators in Koralinska Institutet, Sweden. The library consists of a chemically diverse collection of compounds containing both commercial and internal compounds (donation from Biovitrum). The tested compounds include also 1199 known bioactive molecules from the Prestwick chemical library and 192 nucleosides from Barry and Associates.

Briefly, compounds and controls were dispensed to round-bottomed 384-well plates (Greiner) using an Echo 550 (Labcyte) and diluted with supplemented cell medium using a Multidrop Combi (Thermo Scientific). The diluted compounds were next transferred to 384-well PCR plates (Bio-Rad) using a Bravo liquid handling station (Agilent). K562 suspension cells were then added using a multipipette followed by preincubation for 2 hours in a cell incubator at 37 °C. The PCR plates were next heated transiently to 50 °C for three minutes followed by a controlled cooling to 20 °C using a Roche Lightcycler480 (Roche). The plates were then centrifuged briefly and subsequently cell lysis was accomplished by addition of lysis buffer using a Flexdrop IV (PerkinElmer). To ensure sufficient lysis the cell lysates were thoroughly mixed by ten repetitions of aspirations and dispenses using the 384-well head of the Bravo. Part of the lysates were then transferred to white 384-well plates (PerkinElmer) followed by the addition of antibodies (Santa Cruz Biotechnology) and AlphaScreen® acceptor and donor beads (PerkinElmer) using a Multidrop Combi under subdued light. The plates were incubated over night at room temperature prior to detection in an Envision plate reader (PerkinElmer).

## **2.21 Surface plasmon resonance (SPR)**

Protein hTS (25 µg/ml in a buffer of 10 mM sodium acetate, pH 5.0, 1 mM dUMP and 200 µM methotrexate) were captured on Sensor Chip S-CM5 via amine coupling to a level of approximately 5000 resonance units (RU) using Biacore T-200

(GE Healthcare). Raltitrexed was used as a positive control to ensure that the protein remained active after immobilization and during the run. A concentration series (19.53 nM to 10  $\mu$ M) of CBK115334 was injected over the prepared surface for 60 s and allowed to dissociate for 60 s with a flow rate of 70  $\mu$ l/min, at 25 °C. The assay buffer was 20 mM HEPES, 150 mM NaCl, pH 7.5 and 0.005% Tween 20 supplemented with 1% DMSO and 200  $\mu$ M dUMP. Response data was processed using BIAevaluation software (GE Healthcare). Responses were double referenced and solvent-corrected. The data sets were fitted to 1:1 steady state model for determination of binding constants.

## **2.22 Computational tools**

Instant JChem was used for structure database management, search and prediction (ChemAxon). The protein-ligand crystal structures of both initial fragment hits and follow-up compounds were analyzed using Molecular Operating Environment (MOE) version 2013.08 (Chemical Computing Group) for potential chemical expansion. The Amber12:EHT force field was used for structure preparation and calculation for proteins and small molecules. Potential molecular optimization was investigated using the LigX Ligand Explorer application in MOE.

## **2.23 Compound synthesis**

Compounds designed for hit optimization were submitted to Jubilant Chemsys for synthesis. The compounds were purified using silica gel flash column chromatography or reverse phase high-performance liquid chromatography (HPLC). The purity of compounds was more than 95% in general. Compound characterization was conducted using mass spectroscopy and <sup>1</sup>H-NMR.

## CHAPTER 3 RESULTS

### 3.1 Characterization of hTS protein

#### 3.1.1 Protein purification of hTS

To obtain better characterization of hTS and increase the success rate of crystallization, multiple protein constructs were submitted to Protein Production Platform (PPP) for cloning, protein expression and the first batch of protein purification. In total, 14 constructs with different starting and stopping positions and tags were submitted and purified, with two unexpected point mutations in vc10 and vc13 (Figure 3.1a). All protein constructs were subject to biophysical characterization and crystallization screening. The following crystallization and biophysical work was mainly focused on construct vc1 (full-length construct) and vc5 (giving better crystallization data). In an attempt to get different crystal forms, purification of hTS protein construct vc1 and vc5 was performed in three different buffers (HEPES, Tris and phosphate buffer). The  $K_m$  value of dUMP for purified protein in HEPES buffer was determined to be 8.85  $\mu\text{M}$  and  $K_{cat}$  was 0.22  $\text{s}^{-1}$ . In order to assess protein homogeneity and dimerization, hTS vc1 proteins purified from different buffer systems were analyzed using size exclusion chromatography coupled with multi-angle light scattering (SEC-MALS). The results indicated that all the protein samples were monodisperse and presented as dimers in the three buffer systems (Figure 3.1b). Additionally, the SEC-MALS data of the mutant vc10 showed that vc10 formed dimers as well (Figure 3.1c).

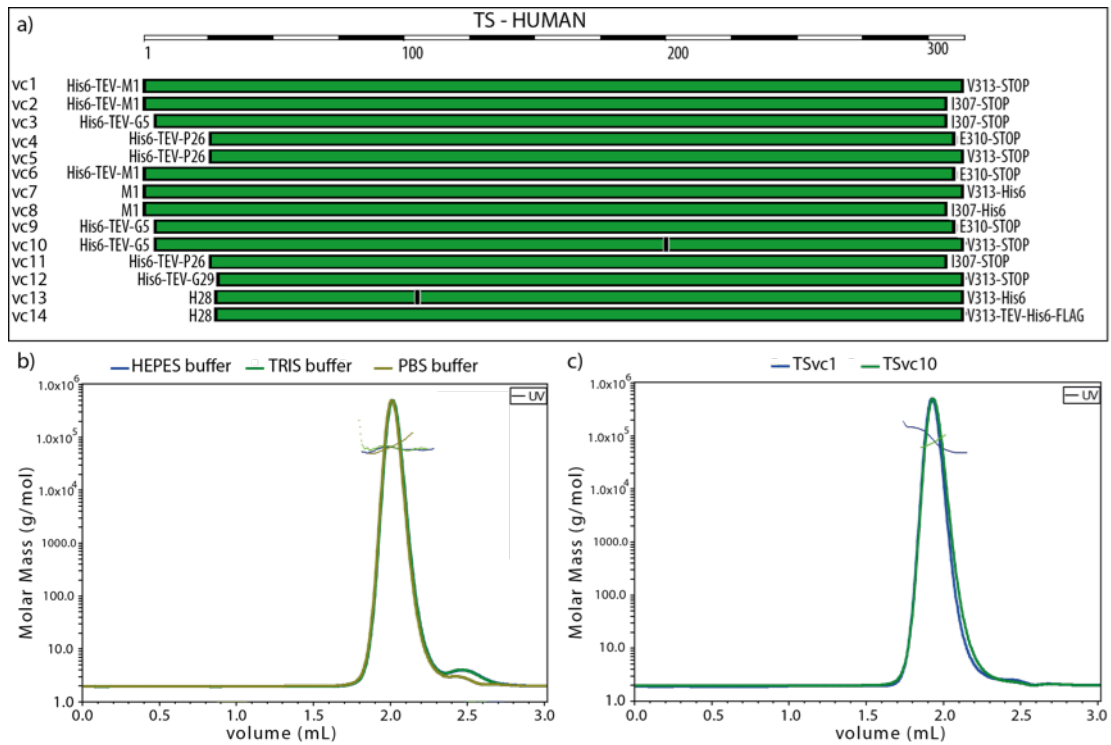
#### 3.1.2 Biophysical characterization of hTS

##### 3.1.2.1 Evaluation of different protein constructs using DSF

Apart from increasing the success rate of crystallization, multiple protein constructs were designed to facilitate characterization of hTS. We were interested in seeing how the N-terminal and particularly C-terminal residues affect the protein and the binding of dUMP and folate. All the protein constructs were examined using a thermal shift assay conducted with differential scanning fluorimetry (DSF). Interestingly, the melting profile of vc1 (hTS full-length construct) was biphasic (Figure 3.2a). The midpoint  $T_m$  for the first transition was about 37  $^{\circ}\text{C}$ , and for the

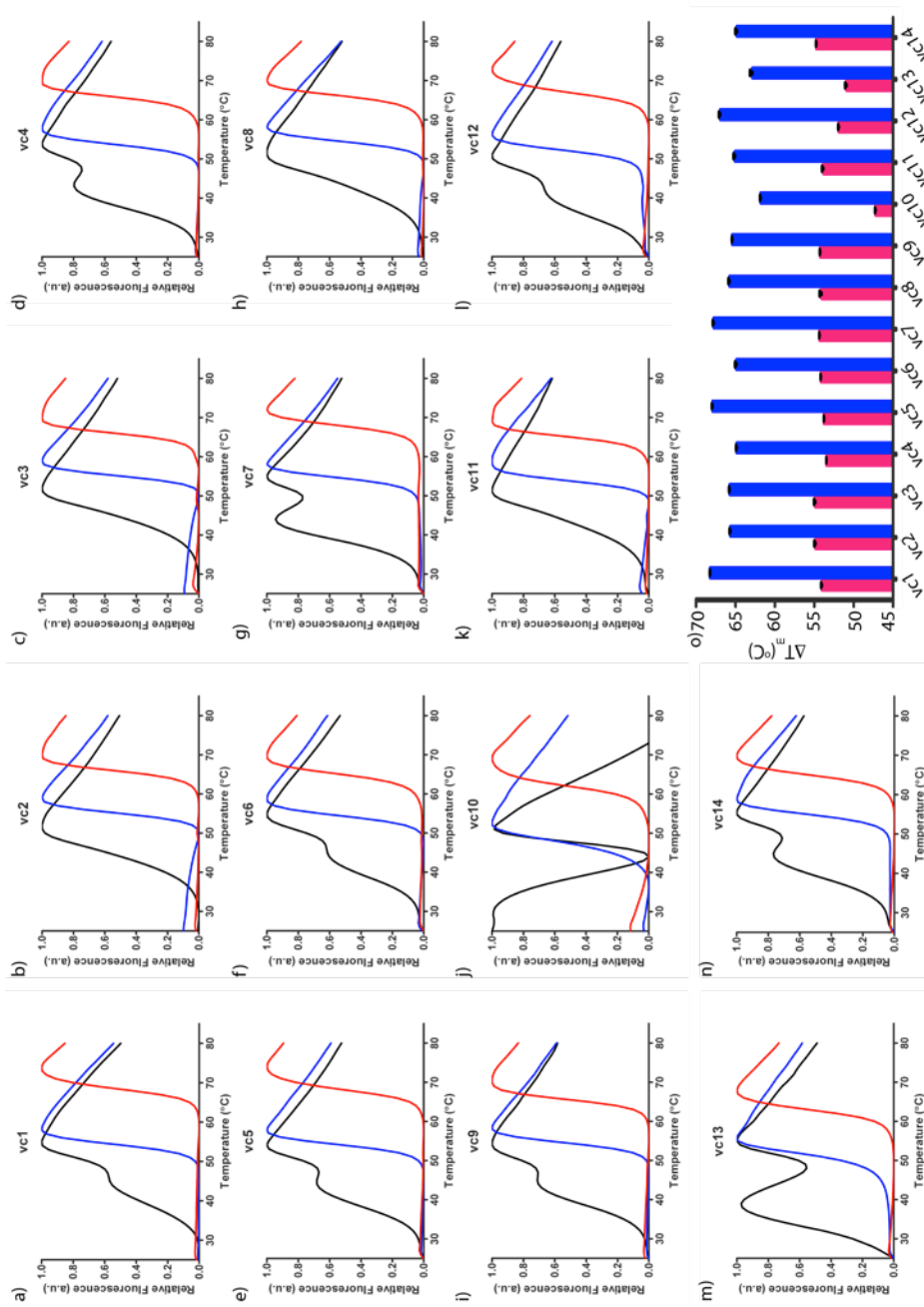
second transition about 52 °C. The addition of dUMP converted the melting curve to a typical monophasic profile. Excluding the two mutants vc10 and vc13, all constructs had apparent biphasic melting profiles, except for vc2, vc3, vc8 and vc11. The second phase of these four constructs became less pronounced. Relating the observation to the construct information, the data suggested that the second phase was closely associated with the residues at the C-terminus as all the four constructs lacked the last six amino acids. Even though there were six histidines extended in vc8, the second phase still diminished, indicative of the importance of sequence of the last six residues at the C-terminus. While comparing vc2 with vc6 or vc11 with vc4, it further pinpointed residues Lys308, Met309 and Glu310 the crucial role in the second phase of the melting curve. The two mutants had distinct melting profiles. Protein vc10 had a monophasic melting curve with very high initial fluorescence. Notably, the addition of high concentration of dUMP did not trigger large shift in vc10 as seen in other constructs but resulted in a significant decline of the initial fluorescence. Like wild-type proteins, vc13 melted in a biphasic profile, but the first phase was less stable.

To assess the binding of dUMP and folate, the melting temperature  $T_m$  was measured at high concentration of dUMP and antifolate raltitrexed. A significant increase of  $T_m$  was shown for all constructs upon the addition of raltitrexed in the presence of dUMP. However, proteins with complete C-terminus (vc1, vc5 and vc12) had higher  $T_m$ . Moreover, short extension at the C-terminus (vc7) did not have significant effect on the binding of raltitrexed, but long extension (vc14) resulted in lower  $T_m$ . It suggested that full native C-terminal residues were required for raltitrexed binding but too long extension at the C-terminal affected the binding negatively. Different from the correlation observed between C-terminal residues and raltitrexed binding, the binding of dUMP was not significantly affected by the variations at the C-terminus. Nevertheless, the data showed that vc12 had significantly lower  $T_m$  with dUMP, suggesting that N-terminal residues could play a role in the binding of dUMP. The two accidentally introduced mutants had much lower  $T_m$  values compared with wild-type proteins, implying that both mutations (C199Y in vc10 and K104N in vc13) affected ligand binding in the active sites.



**Figure 3.1. Different hTS protein constructs and SEC-MALS data.** a) 14 constructs have been designed. Unexpected point mutation is present in construct vc10 (C199Y) and vc13 (K104N). b) MALS data of hTS vc5 purified in HEPES buffer, Tris buffer and phosphate buffer plotted as chromatography of UV absorbance at 280 nm. The mean molar masses determined for protein in the three buffers are 69.8 kDa (in HEPES buffer), 70.1 kDa (in Tris buffer) and 71.3 kDa (in phosphate buffer), approximately the dimer size of hTS vc5 71.3 kDa. c) MALS data of hTS vc1 and hTS vc10 in HEPES buffer plotted as chromatography of UV absorbance at 280 nm. The mean molar mass determined is 75.4 kDa for vc1 and 75.2 kDa for vc5, about the dimer size of hTS vc1 76.8 kDa, and hTS vc10 76.1 kDa.

Abbreviation: His6, hexa histidine-tag; TEV, cleavage site for Tobacco Etch Virus protease (amino acid sequence: SSGVDLGATENLYFQ<sub>S</sub>); FLAG, FLAG octapeptide tag (amino acid sequence: DYKDDDDK).

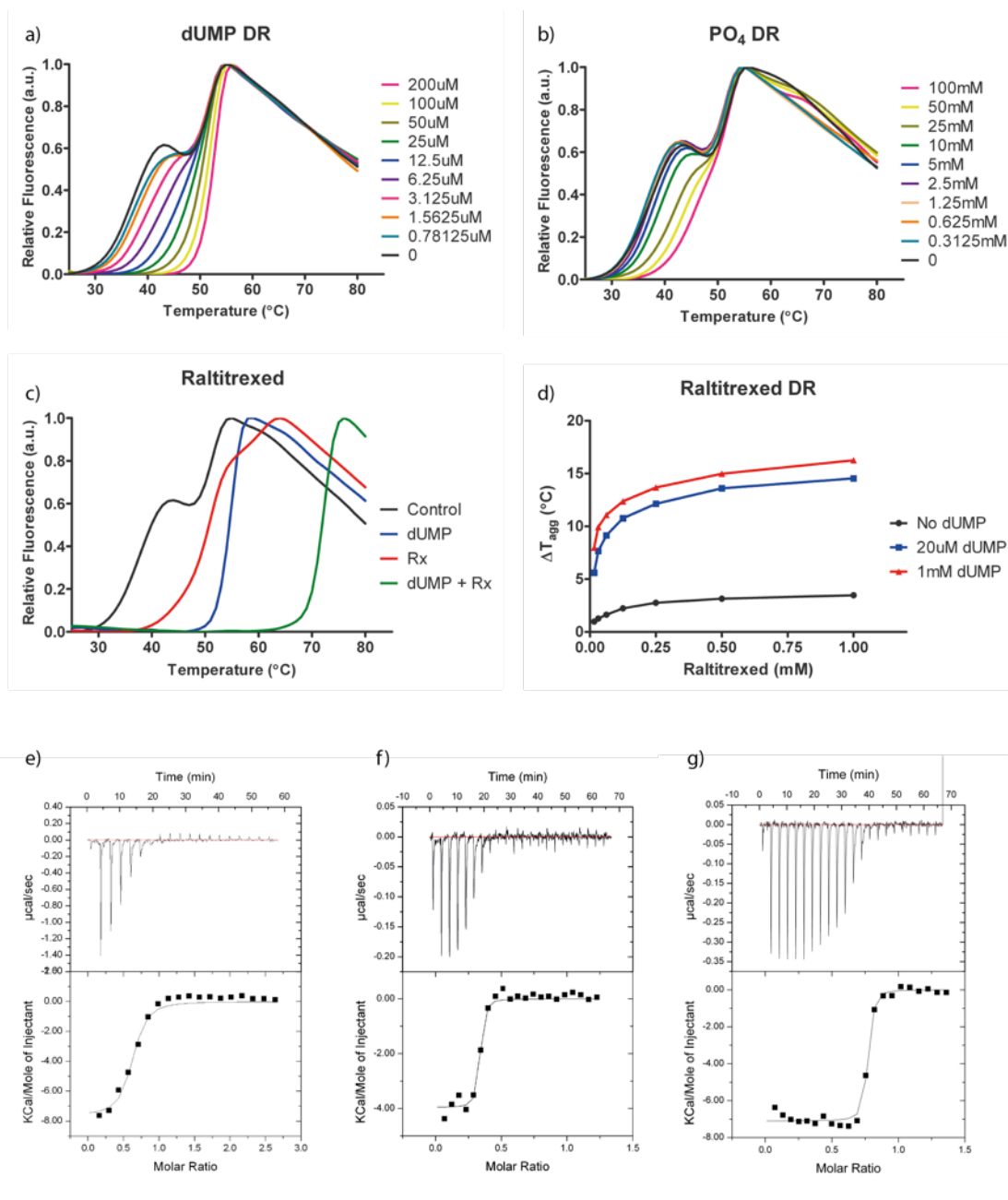


**Figure 3.2. Evaluation of hTS constructs using DSF.** a) to n) 14 protein constructs were analyzed alone (■), in the presence of 500 μM dUMP (■) and of 500 μM dUMP together with 50 μM raltitrexed (■). o) Melting temperature  $T_m$  of the binary and ternary complexes of different constructs: binary complex with 500 μM dUMP (■); ternary complex with 500 μM dUMP and 50 μM raltitrexed (■).

### 3.1.3.2 Binding of dUMP and antifolate

A series of biophysical assays were performed on vc1 to further explore the binding events in the catalytic cavity of hTS. According to the data obtained using DSF, dUMP stabilized only the first phase of hTS at low concentrations while keeping the second phase unaffected. The biphasic profile was converted to a single transition with about 25  $\mu\text{M}$  dUMP (Figure 3.3a). Similar to dUMP, phosphate was able to shift the first phase and yielded a monophasic melting curve when high concentration of phosphate was added (Figure 3.3b).

Antifolate raltitrexed alone induced a smaller shift of protein thermal stability, whereas a 17  $^{\circ}\text{C}$  stabilization was observed when raltitrexed was evaluated in the presence of dUMP (Figure 3.3c). This dUMP-enhanced stabilization effect was also observed using differential static light scattering (DSLS) (Figure 3.3d). With the addition of dUMP, shift of aggregation temperature ( $\Delta T_{\text{agg}}$ ) triggered by raltitrexed increased by about 12  $^{\circ}\text{C}$  at the saturation level. Following up the observations, the binding affinity of dUMP and raltitrexed to hTS was measured using ITC. The  $K_D$  of dUMP was determined to be 1.87  $\mu\text{M}$  and stoichiometry  $N$  about 0.578 (Figure 3.3e). In the absence of additional dUMP, the stoichiometry of binding for raltitrexed was determined to be 0.313. In the presence of 100  $\mu\text{M}$  dUMP, the stoichiometry increased to 0.808. It suggested that more raltitrexed binding site were available when hTS was bound with dUMP. The findings are in line with previous studies, where the folate pocket is fully formed only when dUMP is bound (Finer-Moore et al., 1990). On the other hand, the stoichiometry ( $N = 0.808$ ) at saturated level of dUMP and raltitrexed suggested that about 20% protein was inactive in the sample. Also, about 30% of the purified hTS was bound with dUMP, which should be from the expressing cells and co-purified with hTS.

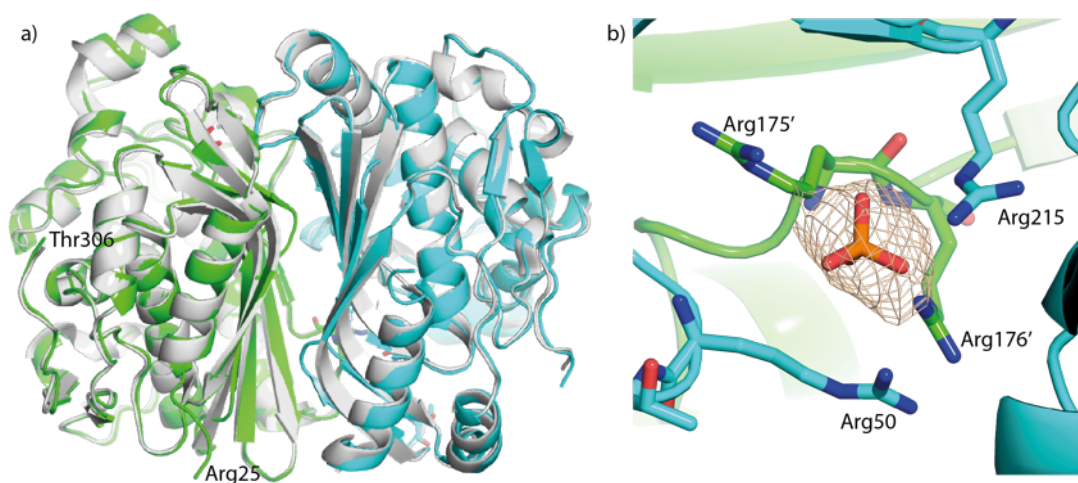


**Figure 3.3. Binding of dUMP and raltitrexed.** a) Dose-response data of dUMP using DSF. b) Dose-response data of phosphate using DSF. c) Stabilization effect of dUMP and raltitrexed on DSF. d) Dose-response data of raltitrexed using DSLS. e) ITC data of hTS titrated with dUMP:  $K_D = 1.87 \mu\text{M}$  and  $N = 0.578$ . f) ITC data of hTS titrated with raltitrexed alone:  $K_D = 69 \text{ nM}$  and  $N = 0.313$ . g) ITC data of hTS titrated with raltitrexed in the presence of 100  $\mu\text{M}$  dUMP:  $K_D = 15.7 \text{ nM}$  and  $N = 0.808$ .

## 3.2 Crystal structures of hTS

### 3.2.1 Structure of full-length hTS

Initial attempts to obtain diffracting crystals of full-length hTS (vc1) were not successful. Having confirmed that phosphate buffer enhanced the thermal stability of hTS, we pursued a second round of crystallization screening using protein purified in phosphate buffer. A 3.5 Å crystal structure of hTS in complex with phosphate ions was obtained after crystallization optimization and streak seeding. The structure is built up with four hTS as two dimers in the asymmetric unit (ASU). The  $\Delta$ SASA (change of solvent accessible surface area) upon dimer formation is about 5300 Å<sup>2</sup>, indicating an obligate dimer. In each subunit, one phosphate ion binds in a position that corresponds to the phosphate moiety of the substrate dUMP (Figure 3.4b). Overall, the structure is very similar to the hTS-P<sub>i</sub> structure of a 7-29 deletion mutant hTS (PDB: 1HZW) (Figure 3.4a), where the last 7 residues (307-313) at the C-terminus are disordered in both cases. Supporting the findings of intrinsically disordered nature (Peña et al., 2009), the N-terminal residues 1-24 are disordered and absent in the crystal structure.

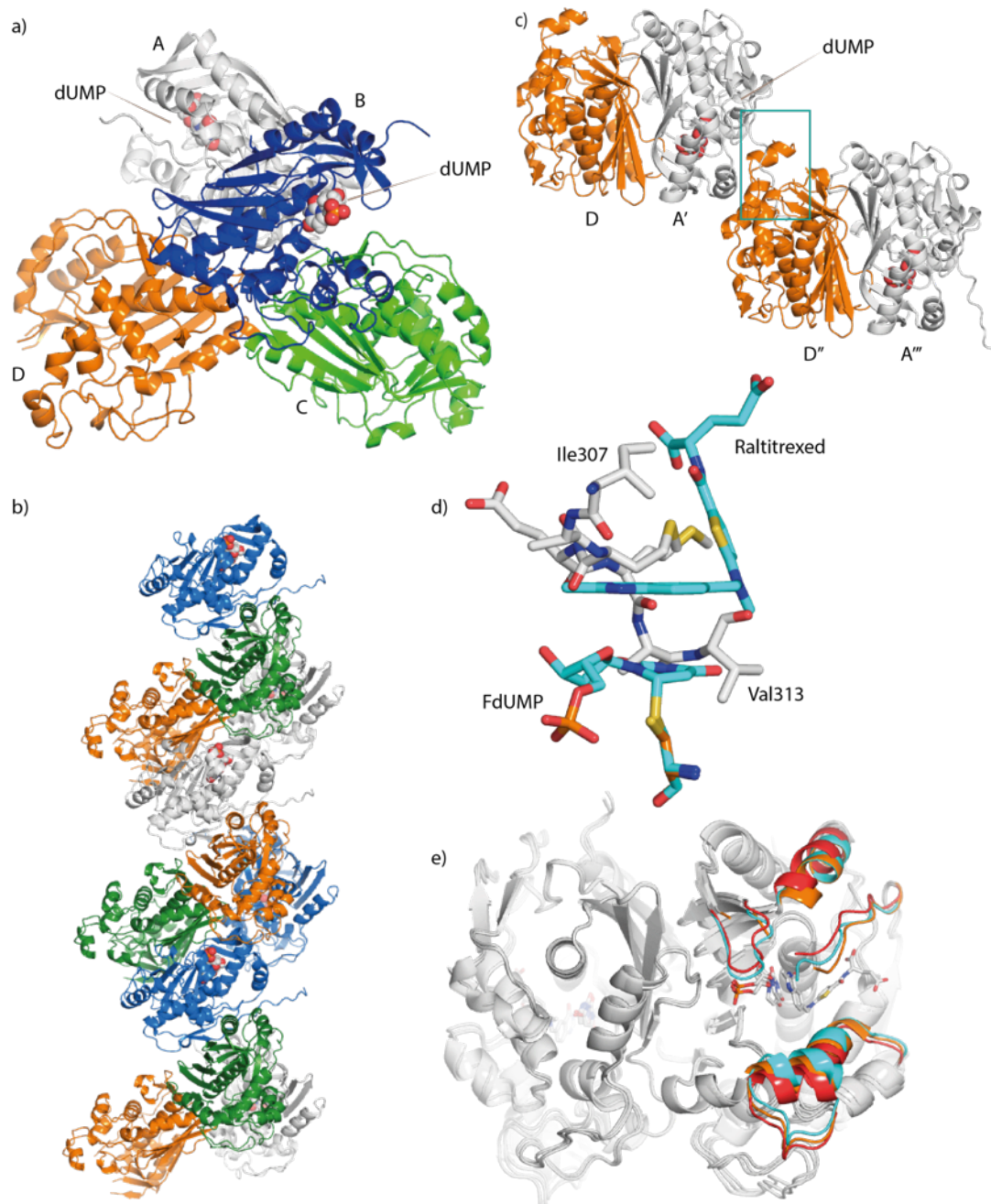


**Figure 3.4. Crystal structure of ligand-free full-length hTS (vc1).** a) Structural alignment of hTS vc1 (green and cyan) with the N-terminal truncated hTS structure 1HZW (white). b) 2Fo-Fc map around phosphate ion contoured at 1.0 sigma.

### 3.2.2 Structure of N-terminal truncated hTS

The N-terminal truncated construct vc5 crystallized in three different conditions, with the crystallization and structure determination details listed in Table 3.1. In the structures of crystal form 1 and 2, dUMP was found in the substrate binding site. The dUMP was co-purified during purification using HEPES buffer since no dUMP was added for crystallization, which was also consistent with the observation from ITC data. Both structures were built up with four monomers per ASU and each monomer formed a dimer with the monomer from the neighboring ASUs. In crystal form 1, subunit A and B had dUMP bound, while in subunit C and D, the active sites were occupied by the C-terminal loop of subunit B and A from the neighboring ASUs, respectively (Figure 3.5a and b). Consequently, crystal form 1 was composed of hTS dimers shown in Figure 3.5c: one subunit bound with dUMP and the other with a C-terminal loop embedded in the active site. The embedded seven C-terminal residues blocked both the dUMP and folate binding sites (Figure 3.5d). Interestingly, binding of the C-terminal loop triggered positional shifts of protein segments and resulted in a closed conformation, which was similar to the covalent ternary complex of hTS-FdUMP-raltitrexed (PDB: 1HVY) (Figure 3.5e). In crystal form 2, all the four monomers in ASU were bound with dUMP. The main chain of C-terminal residues was defined except for the last residue Val313. Comparison with the covalent ternary complex showed that crystal form 2 adopted an open conformation, where several protein segments were further away from the catalytic cavity (Figure 3.5e).

High-resolution crystals with accessible binding pockets are crucial in structure-based drug design programs. The first two crystal forms are not suitable for our following drug discovery studies, because the dUMP binding site is occupied in both crystal forms and the folate binding site is also blocked in crystal form 1. Setting out to obtain a ligand-free crystal form for compound soaking, subsequent crystallization efforts were focused on protein purified in phosphate buffer. The protein produced the third crystal form, crystal of unliganded hTS that diffracted X-ray within 2.5 Å. This crystal form contained 6 monomers per ASU. Similar to unliganded hTS vc1, phosphate ions bound in the crystal structure, residing where the phosphate moiety of dUMP was found.



**Figure 3.5. Crystal structures of N-terminal truncated hTS (vc5).** a) The asymmetric unit of crystal form 1 with 4 hTS monomers. Two monomers (A and B) have dUMP bound, while the other two (C and D) are dUMP-free. b) Symmetry mates of the ASU in a). Subunit A forms dimer with D from the neighboring ASU and C forms dimer with B from the other neighboring ASU. c) hTS dimers present in crystal form 1. The subunit (A in this dimer) that is bound with dUMP inserts its C-terminal loop into the other dimer. d) Close view of the C-terminal residues boxed in c). The residues (white) occupy both the dUMP and folate binding sites, compared with hTS-FdUMP-raltitrexed structure (cyan) (PDB: 1HVY). e) Structural alignment of crystal form 1 and 2 with covalent ternary complex (PDB: 1HVY). The ternary complex represents a closed conformation of hTS, with several protein segments shifted closed to the active site (cyan). Binding of C-terminal loop in crystal form 1 induces similar conformational changes (orange). Crystal form 2 adopts an open conformation, with the segments further away from the active site (red).

**Table 3.1 Data collection and refinement statistics of different crystal forms**

Dataset	hTS vc1	hTS vc5 crystal form 1	hTS vc5 crystal form 2	hTS vc5 crystal form 3
<b>Crystallization</b>				
Protein purification buffer	Phosphate buffer <sup>††</sup>	HEPES buffer <sup>‡‡</sup>	HEPES buffer	Phosphate buffer
Crystallization condition	0.075 M HEPES pH 7.5, 15% PEG10000 and 25% glycerol	0.1 M Bis-Tris pH 5.5, 0.2 M NaCl and 25% PEG3350	0.1 M Bis-Tris pH 7.0, 0.2 M NaCl and 25% PEG3350	0.1 M sodium cacodylate pH 6.5 and 15% PEG4000
<b>Data collection</b>				
X-ray source	NSRRC-13C1	NSRRC-13C1	NSRRC-13B1	AS-MX1
Wavelength, Å	0.9762	0.9762	0.9789	0.7749
Space group	P2 <sub>1</sub>	P2 <sub>1</sub>	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P4 <sub>3</sub> 2 <sub>1</sub> 2
<b>Cell parameters</b>				
a, b, c in Å	90.99, 105.30, 98.20	60.36, 93.74, 100.06	101.88, 108.51, 113.92	109.53, 109.53, 317.47
α, β, γ in degree	90, 108.55, 90	90, 101.64, 90	90, 90, 90	90, 90, 90
Resolution range, Å	30 – 3.50 (3.62 – 3.50)	30 – 2.40 (2.53 – 2.40)	30 – 2.20 (2.28 – 2.20)	30 – 2.40 (2.49 – 2.40)
No. observed reflections	56 559	285 728	462 737	672 784
No. unique reflections	19 878	42 784	65 322	66 303
Completeness, %	84.9 (84.4)	99.9 (100.0)	99.9 (100.0)	86.1 (81.6)
<i>R</i> <sub>merge</sub> , %	9.0 (55.8)	9.2 (30.7)	7.3 (49.3)	10.5 (49.8)
Average <i>I</i> /σ	10.9 (1.55)	11.0 (4.4)	25.7 (3.5)	19.7 (4.03)
<b>Refinement statistics</b>				
Resolution range, Å	30 – 3.50	30 – 2.40	30 – 2.20	30 – 2.40
<i>R</i> <sub>work</sub> , %	17.36	16.01	16.18	19.03
<i>R</i> <sub>free</sub> , %	24.34	22.74	21.94	25.30
Protein/asymmetric unit	4	4	4	6
Ligand	4 P <sub>i</sub>	2 dUMP	4 dUMP	6 P <sub>i</sub>
Water molecules	0	375	705	425
<b>Root mean square deviations</b>				
Bond lengths, Å	0.024	0.008	0.007	0.009
Bond angles, degree	1.488	1.224	1.048	1.143

<sup>††</sup> Phosphate buffer: 100 mM phosphate, 100 mM NaCl and 2 mM TCEP, pH 7.4

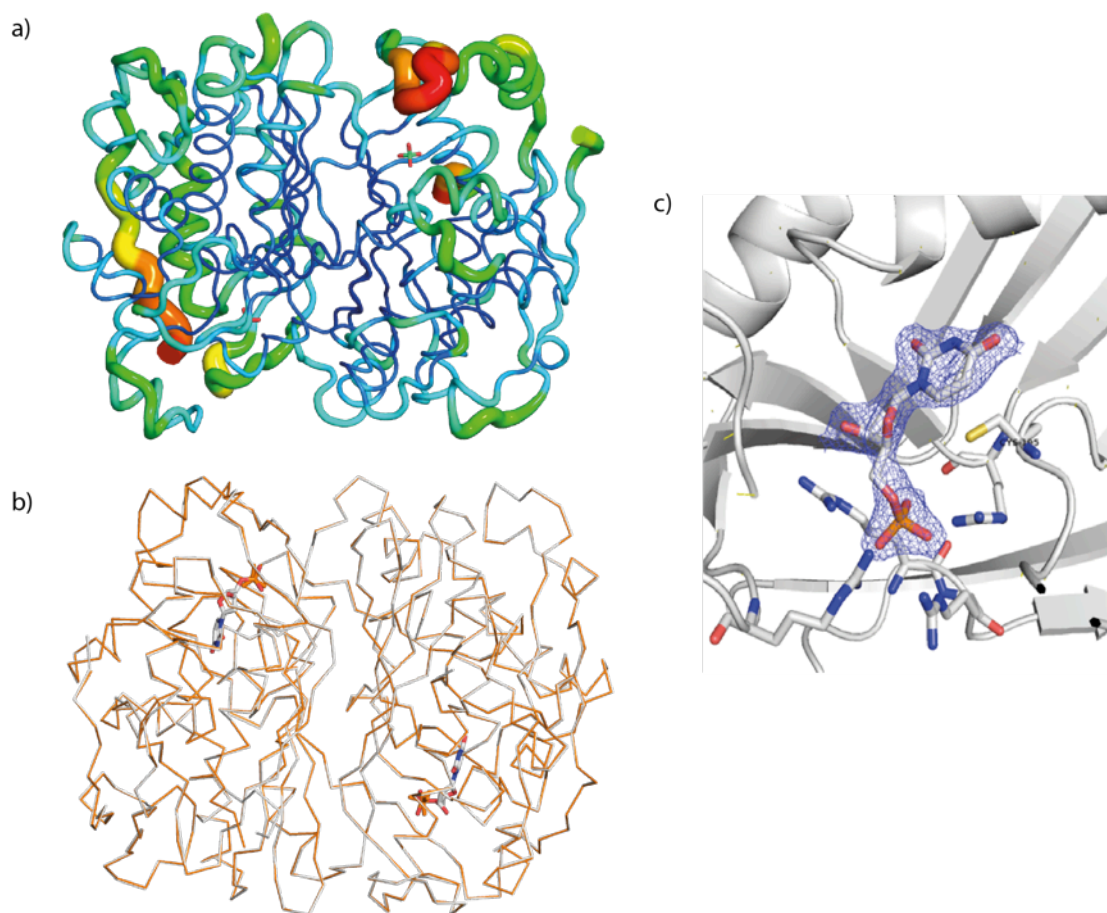
<sup>‡‡</sup> HEPES buffer: 20 mM HEPES, 300 mM NaCl, 10% (v/v) glycerol, and 1 mM TCEP, pH 7.5

### 3.2.3 Conformational changes upon dUMP and antifolate binding

To accurately describe conformational changes triggered by ligand binding, it is useful to compare differences using the same crystal form to avoid errors resulting from different crystal packing forces. The third crystal form of hTS vc5 survived soaking with compounds dissolved in DMSO and diffracted to good resolution (2.0 – 2.5 Å). The crystals enabled the acquisition of fast and reliable structures via soaking and thus were explored to provide additional insights into the structural biology of hTS.

The parent structure (unsoaked) was unliganded hTS bound with phosphate ion, representing the apo enzyme. As shown in Figure 3.6a, regions in the dimer interface had low temperature factor ( $B$ ), indicative of ordered structure and well-defined electron density, whereas the atoms surrounding the active site had higher  $B$  values of which the electron density was weaker. The asymmetric unit of the crystal form was six monomers. Differences in atomic position among the six subunits were small, except for the C-terminal residues. Only in two of the six monomers, the last six residues at the C-terminus were clearly defined with electron density. However, it is difficult and must be careful to draw any implications from the difference observed in thermal mobility of C-terminal residues in the structure. The difference could be the result of crystal packing interactions as the six subunits reside in non-identical crystallographic environment.

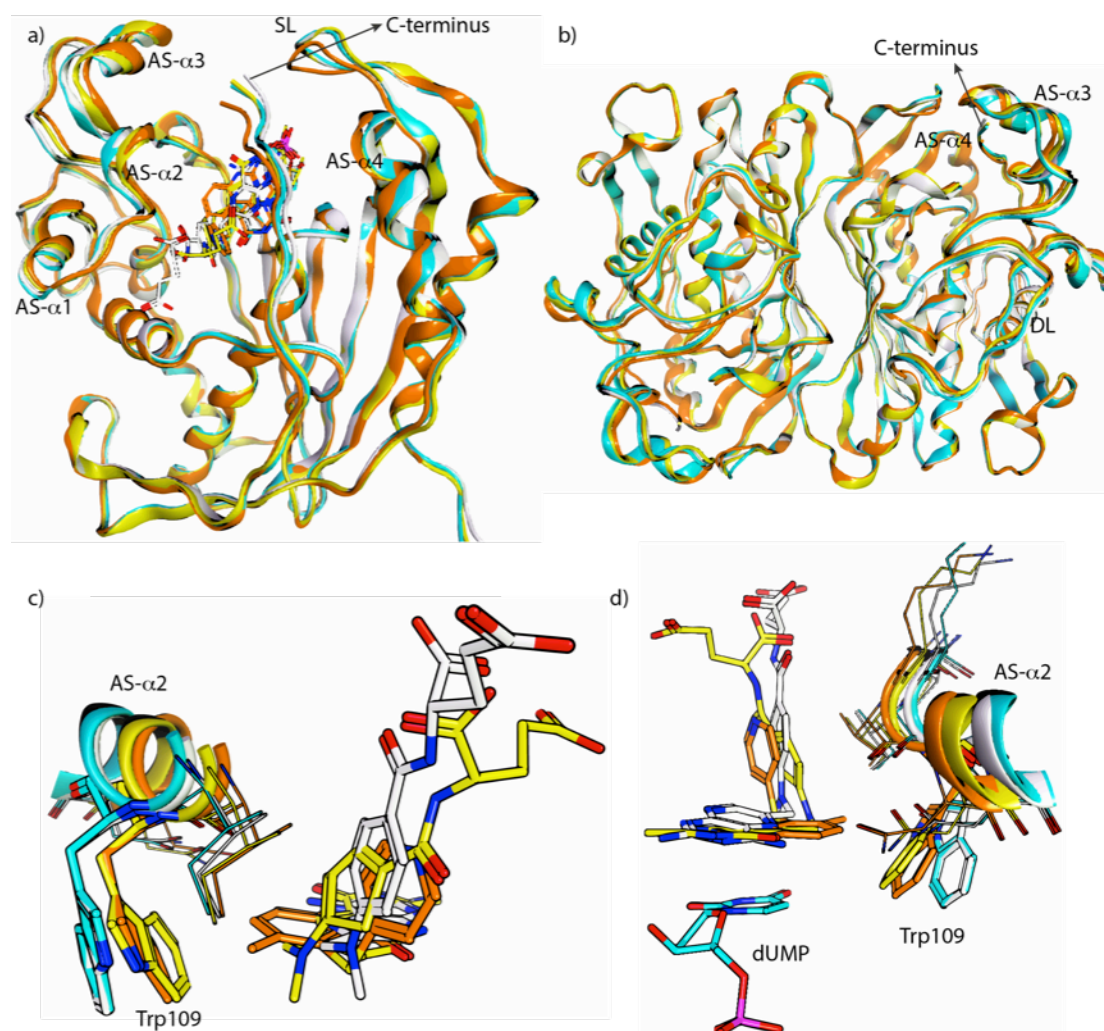
The binary complex structure hTS-dUMP was obtained by soaking with dUMP. The structure was virtually identical to the parent hTS-P<sub>i</sub> structure (Figure 3.6b). The root-mean-square deviation (RMSD) between the apo and binary complex dimer for C $\alpha$  atoms was 0.148 Å (0.241 Å for all atoms). In spite of overall lower thermal mobility, binding of dUMP did not elicit significant improvement in the electron density of residues at the C-terminus.



**Figure 3.6. Ligand-free and dUMP-bound crystal structures of vc5 in crystal form 3.**  
 a) The  $B$  factor putty representation of ligand-free structure created using PyMOL. Thicker secondary structure represents higher  $B$  value. b) Overlay of ligand-free structure (orange) and dUMP-bound structure (white). c)  $2F_o-F_c$  density map around dUMP contoured at 1.0 sigma.

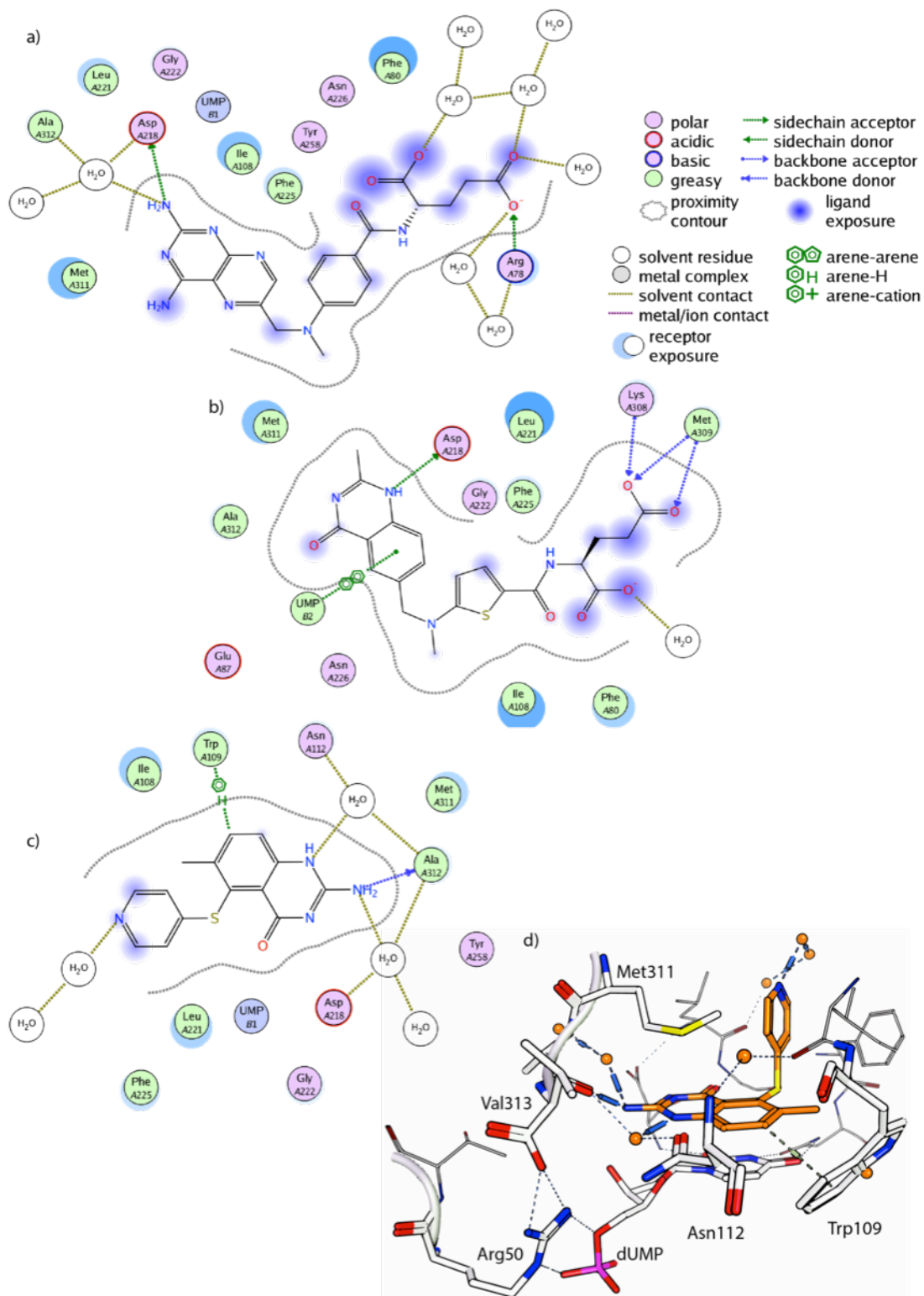
Three ternary complex structures with known antifolates were determined (Figure 3.7): hTS-dUMP-raltitrexed, hTS-dUMP-methotrexate and hTS-dUMP-nolatrexed. These are the first structures of hTS in complex with methotrexate and nolatrexed. An interesting finding is that soaking crystals with nolatrexed resulted in changes of crystal unit cell and space group (Table 3.2). The unit cell and space group conversion reflects the gain of symmetry and corresponds to a switch from six molecules to two molecules per ASU. The three antifolates bound in a similar manner to folate as previously described (Fauman et al., 1994; Fritz et al., 2002), in which the quinazoline ring packed face to face against the pyrimidine and ribose ring of dUMP. The L-glutamate moiety of raltitrexed and methotrexate bound in the entrance channel to the active site. Different from the well-determined electron map for quinazoline

ring, the electron density of L-glutamate moiety in both structures were not clearly defined, implying that the L-glutamate group retained certain flexibility in the binding site.



**Figure 3.7. Alignment of the binary and ternary complex structures.** a) Front view of superposition of complex structures of hTS-dUMP (cyan), hTS-dUMP-methotrexate (white), hTS-dUMP-raltitrexed (yellow) and hTS-dUMP-nolatrexed (orange). b) Back view of the complex structures. c) Back view of the folate binding pocket and positional shift of Trp109. d) Front view of the folate binding pocket and positional shift of Trp109.

In comparison with the apo and binary complex structures, several protein segments lied in slightly different positions in the ternary complexes, including mainly the  $\alpha$ -helix and connecting loops lining the active site (AS- $\alpha$ 1, AS- $\alpha$ 2, AS- $\alpha$ 3, AS- $\alpha$ 4, substrate binding loop SL, dimer interface loop DL and C-terminal loop) as shown in Figure 3.7a and b. It was apparent that these segments were shifted in

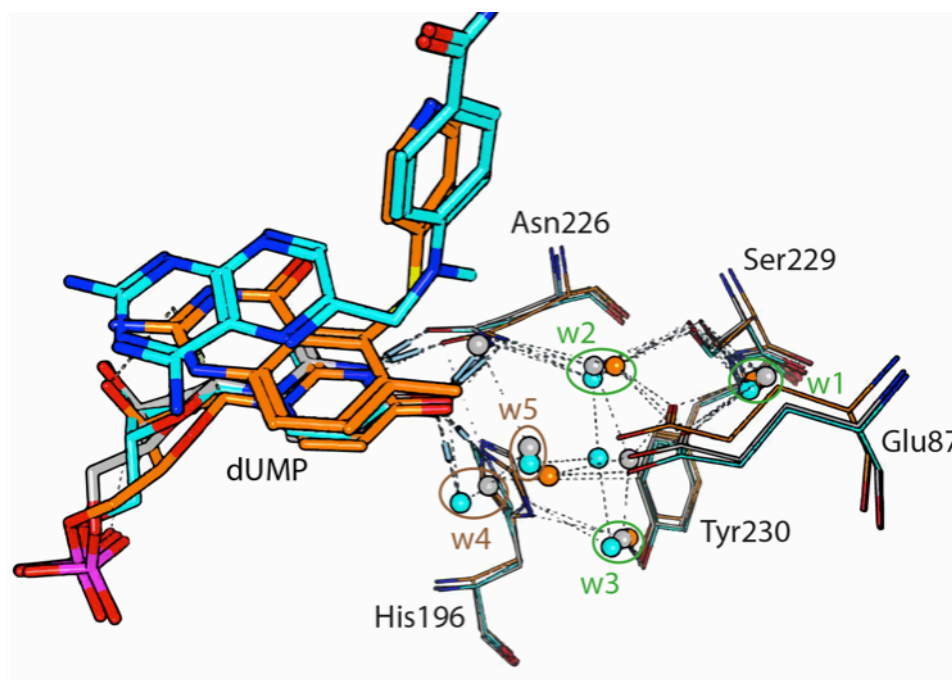


**Figure 3.8. Binding of antifolates in the active site.** a) 2D contact map of methotrexate. b) 2D contact map of raltitrexed. c) 2D contact map of nolatrexed. d) Close view of C-terminal residues in the crystal structure hTS-dUMP-nolatrexed. Met311 is located above the quinazoline ring. The terminus residue Val313 is hydrogen bonded to Arg50 upon nolatrexed binding.

concert toward the active center of hTS. The conformational changes induced generally served to close the active site. Overlaying the structures revealed that nolatrexed triggered the largest conformational changes, with methotrexate the smallest. The radius of gyration calculated for the ternary complex dimer was 23.96 Å for nolatrexed, 23.99 Å for raltitrexed and 24.08 Å for methotrexate, compared with 24.07 Å for the binary hTS-dUMP dimer. Binding of methotrexate was unable to elicit positional shifts of AS- $\alpha$ 1, AS- $\alpha$ 2, AS- $\alpha$ 3 and the connecting loops, which was probably because its quinazoline ring sat closer to the C-terminal loop to form hydrogen bonds and thus too far to make contacts with those helices and loops on the other side. In contrast, these segments moved into the active site upon the binding of raltitrexed and nolatrexed. The residues moved with the translational shift vectors pointing toward the cavity as units to avoid disruption of the favorable hydrogen-bond patterns in these segments as well as make interactions with the ligands. The only residue that underwent extra rotation of side chain was the absolutely conserved Trp109, of which the indole ring rotated into the active site by 6° in raltitrexed structure and 3° in nolatrexed structure to interact with the quinazoline ring (Figure 3.7c and d).

Another pronounced structural difference between the binary (or apo structure) and the ternary complexes involves the residues at the C-terminus. In the absence of ligands in the folate site, 2/3 of molecules in the structure had flexible C-terminal loop, where the electron density was not well defined. However, the C-terminal residues became ordered with well defined electron density in all molecules in the three ternary complexes. The structures showed that the three antifolates interacted with the backbone of residues of the C-terminus directly and/or via water-mediated hydrogen bonding, including Ala312 with methotrexate and nolatrexed and Lys308 and Met309 with raltitrexed (Figure 3.8). In addition, the methionine sulfur of Met311 was found to be in close contact with the pterin or quinazoline ring of the antifolates (Figure 3.8d). These interactions explain why the C-terminal loop becomes stabilized and thus defined with electron density upon the binding of antifolates. The largest shift of C-terminal loop occurred in the structure of hTS-dUMP-nolatrexed, where the C-terminal region moved up to 2.2 Å further into the active site and served as a lid to cover the cavity. The main chain carboxyl of Val313 formed a salt bridge with Arg50 that in turn formed hydrogen bonds to the phosphate moiety of dUMP.

Consequently accompanied with the shift of C-terminal loop, the substrate-binding loop (SL) in which Arg50 resided moved toward the active site by 2.18 Å in the structure (Figure 3.7a).



**Figure 3.9. Water molecules in the back pocket near Glu87.** Overlay of active site residues and water molecules in crystal structures of hTS-dUMP (grey), hTS-dUMP-methotrexate (cyan) and hTS-dUMP-nolatrexed (orange).

Due to the polar properties of the two substrates dUMP and folate, the active cavity of hTS is lined with many polar amino acids. In the hTS-P<sub>i</sub> structure, the cavity was filled with a large number of ordered water molecules, most of which were expelled by the binding of dUMP and antifolates. An unusual feature of the solvent observed in high-resolution structures of TSs is that an ordered water network is buried near to Glu87 at the back of the active site. Those waters have been proposed to assist proton transfers to and from dUMP during the catalytic reaction (Fauman et al., 1994). Several water molecules were structurally conserved in the binary and ternary complex structures (Figure 3.9). Water molecules w1, w2 and w3 were highly conserved. Water w1 and w2 were hydrogen bonded to Glu87 and bridged Glu87 to Ser229 and Asn226, respectively. Water w3 bridged Tyr230 to His196, which in turn was hydrogen bonded to O-4 of dUMP. In the structure of hTS-dUMP, two water

molecules (w4 and w5) were buried at the back pocket, constituting an ordered solvent chain between O-4 of dUMP, Glu87 and His196. This solvent structure remained unperturbed during the binding of methotrexate. However, binding of nolatrexed placed its methyl group in the back pocket and thus disrupted the water network, resulting in displacement of water w4 and positional shift of w5.

**Table 3.2 Data collection and refinement statistics of binary and ternary complex structures**

Dataset	hTS-dUMP	hTS-dUMP-methotrexate	hTS-dUMP-raltitrexed	hTS-dUMP-nolatrexed
Crystal form	hTS vc5 crystal form 3	hTS vc5 crystal form 3	hTS vc5 crystal form 3	hTS vc5 crystal form 3
Soaking details	2 mM dUMP, 10% DMSO, 20 mins soaking	2 mM dUMP and 10 mM methotrexate, 10% DMSO, 20 mins soaking	2 mM dUMP and 10 mM raltitrexed, 10% DMSO, 20 mins soaking	2 mM dUMP and 10 mM nolatrexed, 10% DMSO, 20 mins soaking
Data collection				
X-ray source	AS-MX1	AS-MX1	AS-MX2	AS-MX2
Wavelength, Å	0.9537	0.9537	1.0130	1.0130
Space group	P4 <sub>3</sub> 2 <sub>1</sub> 2	P4 <sub>3</sub> 2 <sub>1</sub> 2	P4 <sub>3</sub> 2 <sub>1</sub> 2	P4 <sub>1</sub> 2 <sub>1</sub> 2
Cell parameters				
a, b, c in Å	109.74, 109.74, 317.69	109.20, 109.20, 317.13	107.84, 107.84, 314.95	108.29, 108.29, 106.03
α, β, γ in degree	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Resolution range, Å	30 – 2.00 (2.07 – 2.00)	30 – 2.00 (2.07 – 2.00)	30 – 2.80 (2.90 – 2.80)	30 – 2.13 (2.20 – 2.13)
No. observed reflections	463 678	642 721	303 589	560 607
No. unique reflections	116 562	119 969	43 961	35 535
Completeness, %	88.2 (95.2)	91.6 (86.8)	93.0 (93.8)	99.5 (94.4)
<i>R</i> <sub>merge</sub> , %	8.3 (40.3)	9.3 (44.5)	12.6 (59.6)	17.3 (86.7)
Average <i>I</i> /σ	14.3 (3.2)	15.4 (3.2)	14.9 (3.2)	15.5 (4.3)
Refinement statistics				
Resolution range, Å	30 – 2.00	30 – 2.00	30 – 2.80	30 – 2.13
<i>R</i> <sub>work</sub> , %	16.99	18.89	18.50	15.02
<i>R</i> <sub>free</sub> , %	22.41	24.31	27.76	19.96
Protein/asymmetric unit	6	6	6	2
Ligand	6 dUMP	6 dUMP and 6 methotrexate	6 dUMP and 6 raltitrexed	2 dUMP and 2 nolatrexed
Water molecules	1770	1700	182	404
Root mean square deviations				
Bond lengths, Å	0.008	0.007	0.009	0.008
Bond angles, degree	1.129	1.123	1.238	1.207

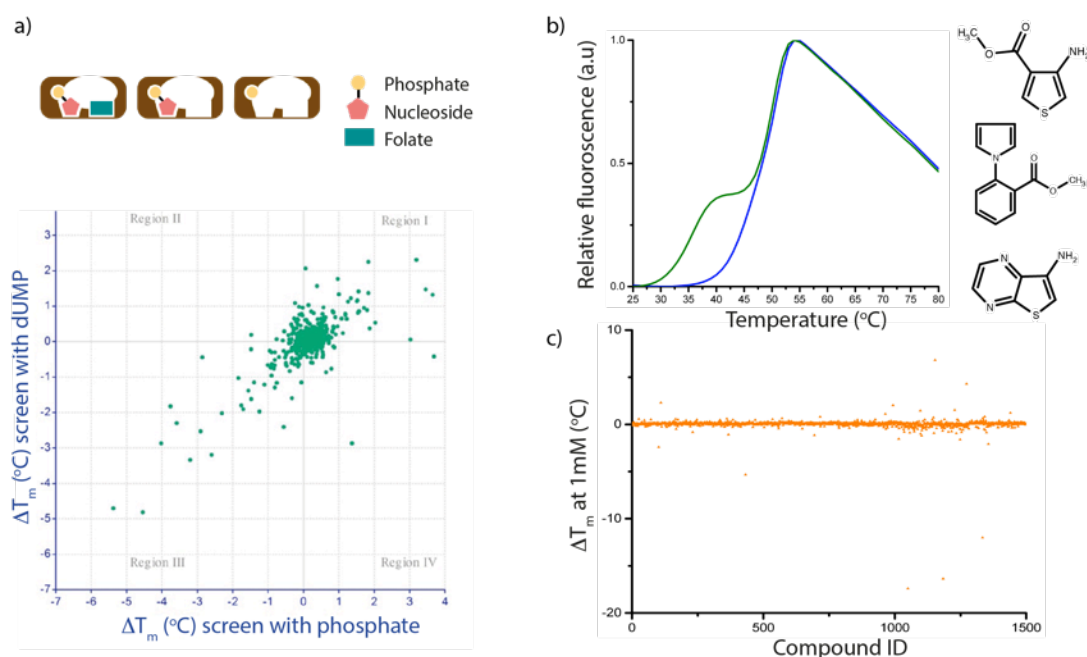
### 3.3 Fragment-based drug discovery targeting hTS

#### 3.3.1 Fragment screening using DSF

As discussed in the introduction sections, multiple strategies are possible for fragment screening. In this work, thermal shift assay-based screening was explored, a rapid and low cost strategy for identifying initial fragment hits (Larsson et al., 2013). The first set of fragment screening was performed against a library of 500 compounds. In an attempt to sample different binding sites of hTS, differential screening strategy was applied. The rationale of differential screening is illustrated schematically in Figure 3.10a: screening in the presence of substrate dUMP provides a more intact folate binding site and enables identification of complementary binders in this site; screening in the presence of phosphate allows recognition of fragments in both nucleotide and folate binding sites. The screening was performed in the presence of 20  $\mu$ M dUMP or in 100 mM phosphate buffer. Differential analysis of screening data is shown in Figure 3.10a. The majority of fragments behave similarly in the two conditions. Notably, certain fragments are protein stabilizers in one condition but destabilizers in the other condition (region II and in particular region IV), indicating the different availability of certain binding sites between these two conditions. Hits with positive  $\Delta T_m$  fall predominantly into region I as stabilizers in both conditions. Some fragments in region III turn out to be generic destabilizers, which also give large negative  $\Delta T_m$  on other proteins. Interestingly, there are several fragments shifting the monophasic profile back to biphasic melting curves by destabilizing the first phase (Figure 3.10b).

Using a threshold of  $\Delta T_m \geq 0.5$  °C, 46 primary screening hits were validated for dose-dependent behavior using DSF. In total, 29 hits stabilized hTS in a dose-dependent manner, which translated to a hit rate of 5.8% (Appendix Table 1). These positive hits were classified into three categories: 4 dUMP hits that exhibited dose-response only in the presence of dUMP; 6 phosphate hits that showed dose-response only in phosphate buffer; and 19 common hits as dose-dependent stabilizers in both conditions. Other than the positive hits, the profile-changing hits were also subject to dose-response validation, where three fragments transformed monophasic melting profile to biphasic with dose dependency (Figure 3.10b).

As outlined later that the success rate of translating dUMP hits into crystal structures is much higher than that of phosphate hits, the second set of fragment screening against another 1500 compounds was performed in the presence of 200  $\mu\text{M}$  dUMP (Figure 3.10c). 19 hits demonstrated dose-dependent stabilization, resulting in a hit rate of 1.27% (Appendix Table 2).

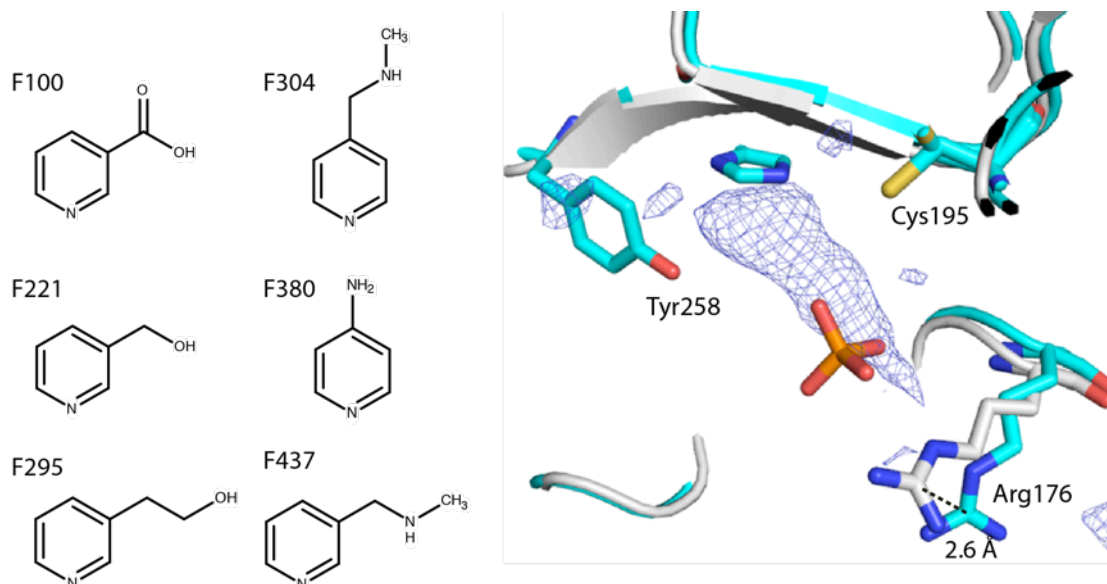


**Figure 3.10. DSF-based fragment screening data.** a) Differential screening and data analysis for the first fragment library of 500 compounds. b) Three fragments were identified to transform the melting profile back to biphasic curve. c) Fragment screening data of the second fragment library of 1500 compounds in the presence of 200  $\mu\text{M}$  dUMP.

### 3.3.2 Crystallographic screening of fragment cocktails

Considering that DSF might not be sensitive enough to detect binding of very small fragments in our library, crystallographic screening was applied to 96 small fragments using hTS- $\text{P}_i$  crystals (crystal form 3 of vc5). To improve the efficiency of fragment soaking, the 96 fragments were grouped into 16 cocktails of shape-wise similar fragments at high concentrations (Appendix Table 3). Of the 16 cocktail soaks, one resulted in a difference Fourier peak in dUMP binding site (Figure 3.11). Moreover, the phosphate-coordinating Arg176 was displaced from the active site by 2.6 Å in the structure, which gave further support for that the difference density could correspond to a bound fragment. Because identification of the ligand from the

electron density was difficult, we progressed to deconvolution by soaking with individual fragment. Unfortunately, neither of the fragments produced the same difference density and the displacement of Arg176. We suspected that failure to reproduce the density was probably because the cocktail sample was contaminated by degradation products after storage.



**Figure 3.11. Fragment cocktail 1 and difference Fourier peak in the crystal structure.** The crystal structure after soaking with cocktail1 (cyan) is superposed with unsoaked hTS-P<sub>i</sub> structure (white). The Fo-Fc difference map of soaked dataset in the active site is contoured at sigma level of 3.0.

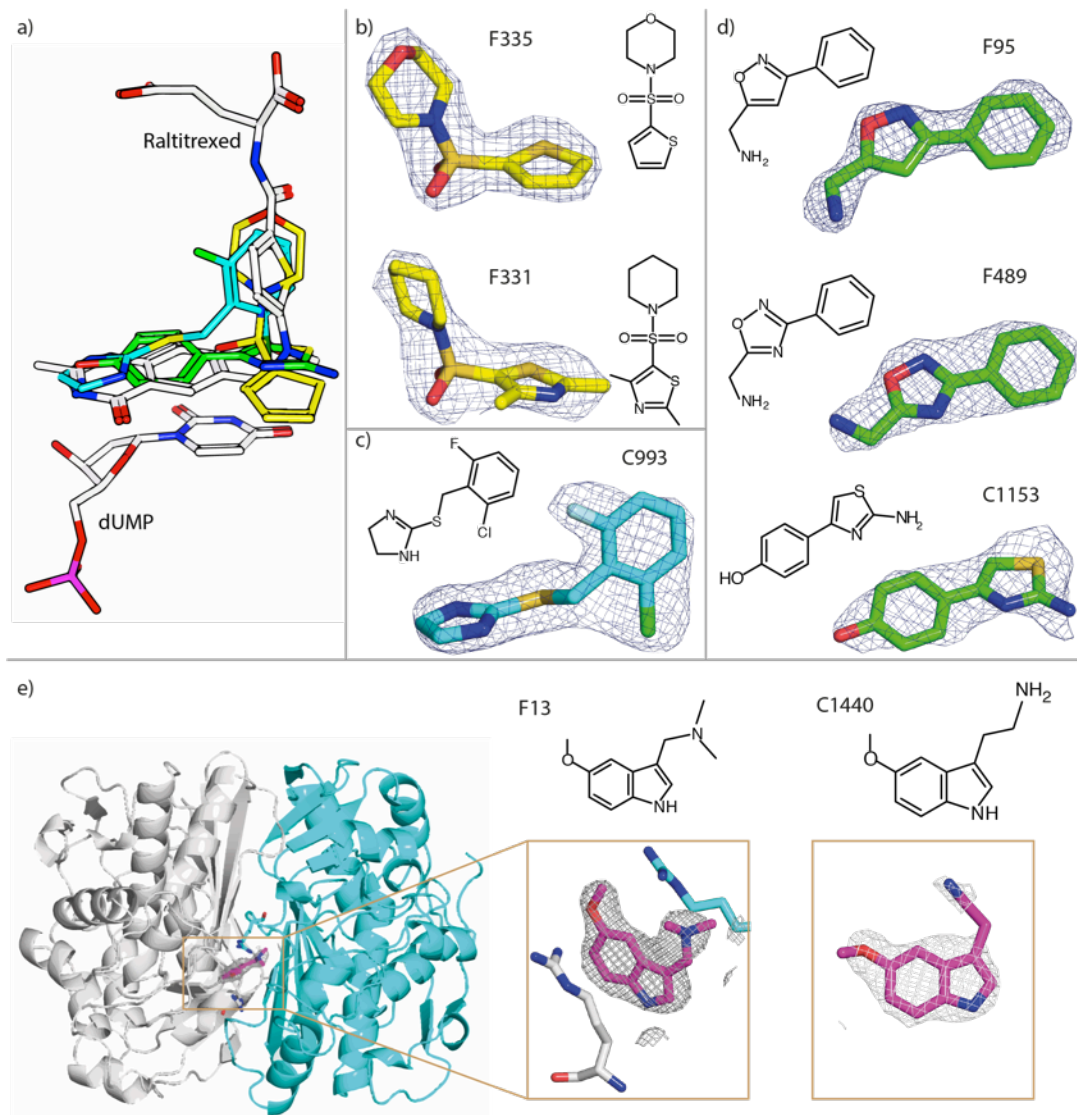
### 3.3.3 Translating fragment hits into crystal structure

All hits from DSF screening were progressed to crystal soaking to obtain structural information. For phosphate hits and profile-changing hits, soaking was performed using hTS-P<sub>i</sub> crystals (crystal form 3 of vc5). For dUMP hits, hTS-dUMP cocrystals (in the same condition of crystal form 3) were used because of better crystallization results and diffraction quality. For common hits, both hTS-P<sub>i</sub> and hTS-dUMP crystals were used for soaking. Ultimately, eight fragments were successfully translated into crystal structures in complex with dUMP. None of phosphate hits and profile-changing hits had been identified in crystal structure using soaking approach. Even though some datasets were collected from co-crystallization of these fragments, none of them was translated into the crystal structures. Reasons for this could be a combination of factors such as intrinsic weak affinities, the different buffer conditions

in the crystal as compared to screening, as well as the limited conformation flexibility of the structure in the crystal as compared to in solution.

Based on their binding modes, the eight hits in hTS-dUMP structures are categorized into four classes (Figure 3.12). Type I hits are two sulfonamides, which adopt an L-shape conformation with its horizontal part residing in the back pocket near Glu87 and the vertical part in the entrance tunnel of the active site. Type II hit binds partially in the entrance tunnel as well, but its head ring sits above the ribose ring of dUMP and overlaps the end site of quinazoline ring of raltitrexed. Three fragments are of type III binding mode, which bind in a planar conformation face to face with the pyrimidine ring of dUMP. Apart from hits identified in the active site or folate binding site, more specifically, two fragments bind in the dimer interface, which are about 25 Å away from the active site. The two type IV hits are sandwiched between Arg185 and Arg185' in the dimer interface.

We decided to progress type I and II hits into chemical evolution for drug development against hTS. The reason is that the binding modes of type I and II hits are of higher novelty value than type III hit in the folate binding pocket. Additionally, type I and II hits overlap in the entrance tunnel region, which offers an opportunity to merge the two hits and further adds value for exploration. Another reason for not pursuing type III hit at this point is that we are not confident in the conformation of the bound fragments when the fragments could only be fitted into the density with partial occupancies and the data indicated that there could be multiple binding modes. Regarding type IV hits binding distant from the active site, this is less interesting for drug development but the novel pocket identified might play a functional role and could therefore be of significant interest as is further studied in this work.

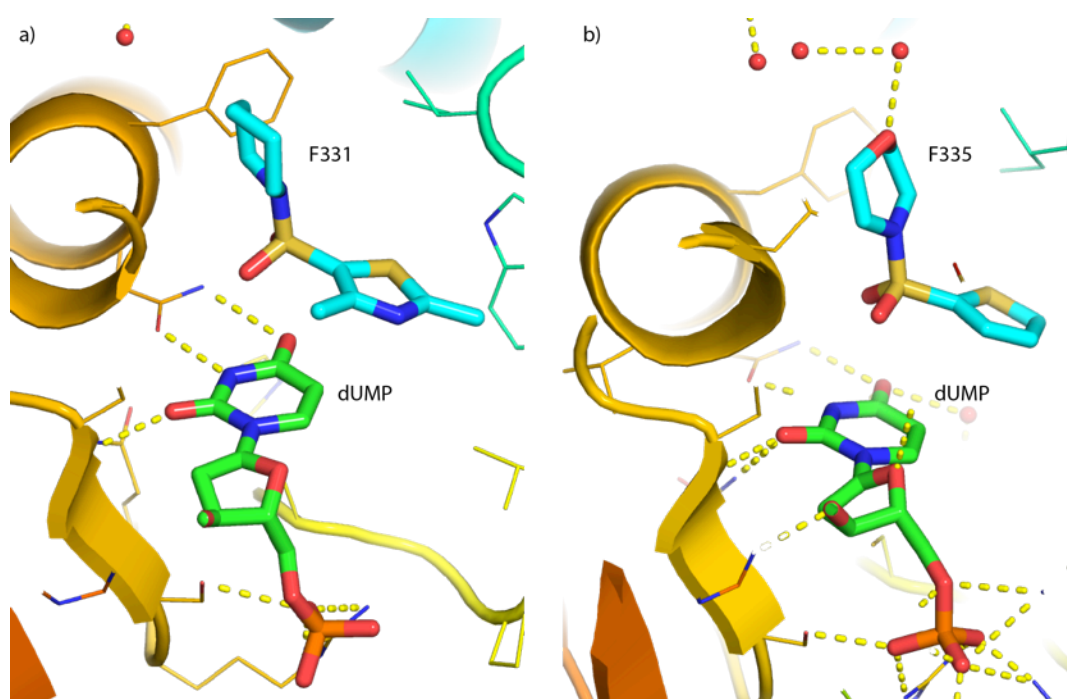


**Figure 3.12. Four types of fragment hits identified in the crystal structures.** a) Superposition of all four types of hits with raltitrexed. b) Type I hit: two fragments adopting L-shape. c) Type II hit: one fragment. d) Type III hit: three fragments with planar conformation. e) Type IV hit: 2 fragments binding in the hTS dimer interface. The 2Fo-Fc map around the fragments is contoured at 1.0 sigma.

### 3.3.4 Structure-based drug design

#### 3.3.4.1 Expansion and optimization of type I hit

Two sulfonamides from our fragment library bind in the folate binding site of hTS in the type I binding mode (Figure 3.13). The two fragments are structurally similar. Analysis of the crystal structures showed that the morpholine group of F335 participated in water-mediated interaction in the entrance channel while the piperidine in F331 did not. On the basis of this, our first round of expansion kept the morpholine group and optimized the other ring binding near Glu87.



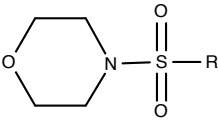
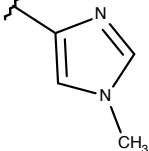
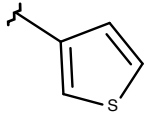
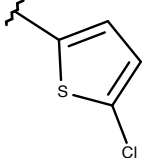
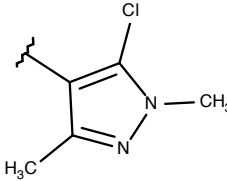
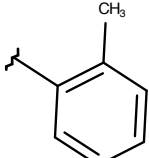
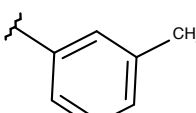
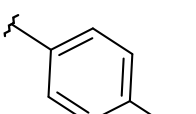
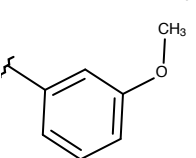
**Figure 3.13. Binding of type I fragments.** Close view of fragment F331 (a) and F335 (b) in the folate binding site of hTS: the oxygen from the morpholine group of F335 forms hydrogen bond with one water molecule in the entrance channel.

A set of 8 compounds was prepared and evaluated using DSF (Table 3.3). The compounds substituted with phenyl derivatives led to larger  $T_m$  shift than the ones with 5-membered rings. We subsequently solved the structures of protein-ligand complexes of **A5** and **A8**. The structures revealed that **A5** and **A8** bound in a similar manner to the parent compounds. The *o*-methyl of **A5** pointed toward the back of the active site (near Glu87) while the *m*-methoxyl group in **A8** pointed in the opposite direction (Figure 3.14a). It was apparent from the structures that **A5** rotated by about

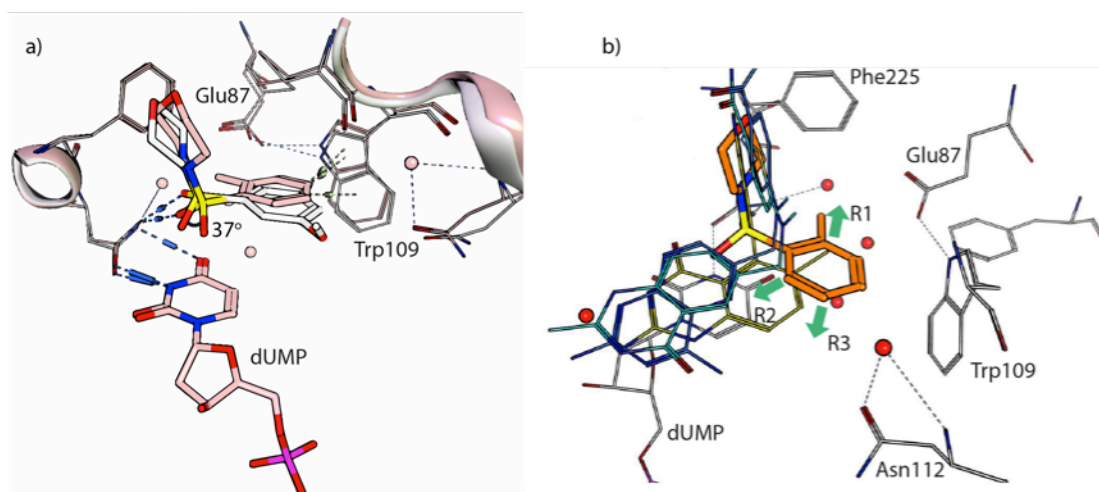
37° toward the back pocket, but anchor of oxygens of the sulfonyl group to Asn226 remained. Although no structure of **A7** was obtained, the two structures suggested that *para*-substituents could easily cause steric hindrance because the vector was very close to Tyr109 (about 3.4 Å to the nearest atom). This is supported by the fact that *para*-substituents larger than -CH<sub>3</sub> abolished the stabilization effect on hTS (data not shown).

Further analysis of the crystal structures suggested a few vectors that could be exploited in the next cycle of chemical expansion (Figure 3.14b). Initially, small modifications were made to R1, vectoring out toward the back pocket near Glu87. Introduction of hydroxymethyl group (**A11**) invoked larger  $\Delta T_m$  relative to other size-wise similar substituents (Table 3.4). An SAR study on vector R1 revealed that changing the hydroxyl group to an amino (**A12**) or amide (**A14**) resulted in significant decrease of  $\Delta T_m$ . Bulkier substituent with a 1-hydroxyethyl group (**A13**) also had a large decrease in  $\Delta T_m$ . We subsequently solved the crystal structure of hTS in complex with **A11**. The structure showed that the hydroxyl group had displaced the water molecule w5 of the solvent chain in the back pocket. Consequently, the hydroxyl group formed hydrogen bond with water w4 behaving as w5, but it was slightly far from Glu87 for hydrogen bonding (Figure 3.15a). Inspired by the displacement of water molecules in the water-filled back pocket, the length of carbon chain was investigated setting out to involve more polar interactions with water molecules and/or polar residues lining the pocket. Interestingly, extending the carbon chain by another atom (**A15**) gave strikingly better binding affinity of 19  $\mu$ M, leading to an approximate 6-fold improvement over **A11**, but longer linkers were suboptimal (**A16** and **A17**). As revealed by the crystal structure, the extended substituent in **A15** adopted a conformation maximizing its polar interactions with proximal solvent and residues. Aside from being hydrogen bonded to water w4, **A15** participated in two more hydrogen bonds formation with O-5 of dUMP as well as the side chain of Glu87 (Figure 3.15a). Furthermore, the structure suggested that the suboptimal binding of longer substituents **A16** and **A17** was likely due to introduction of steric clash and disruption of polar interactions.

**Table 3.3. Biophysical data of compounds A1 – A8**

Cpd	R1	DSF $\Delta T_m$ (°C) <sup>8</sup>
		
A1		0.05
A2		0.37
A3		0.29
A4		0.77
A5		1.31
A6		0.97
A7		0.52
A8		1.12

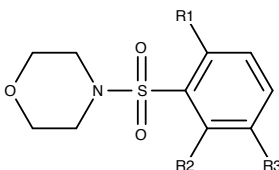



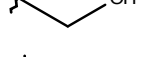
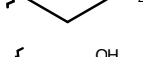
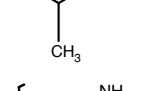
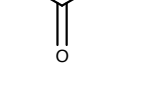
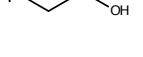

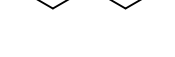

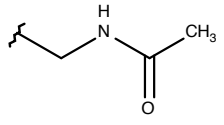

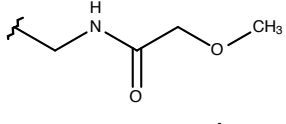

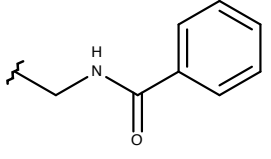

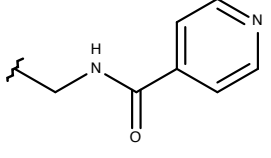


<sup>8</sup> DSF data was measured at 1 mM of compound in the presence of 500  $\mu$ M dUMP in HEPES buffer (20 mM HEPES and 150 mM NaCl, pH 7.5)

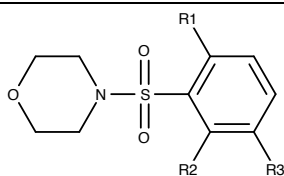


**Figure 3.14. Crystal structures of hTS complexed with compound A5 and A8.** a) Superposition of crystal structure of **A5** (pink) with the one of **A8** (white). b) Overlay of structure of **A5** (orange) with structures of the three antifolates (methotrexate, raltitrexed and nolatrexed). There is a “hot spot” region above dUMP where all the three antifolates occupy. Three vectors are suggested for expansion based on structural analysis. Vector R1 points toward Glu87. R2 and R3 vector out to the “hot spot” region.

Further modification also focused attention on R2 and R3 vectors. We reasoned that extension on the vectors would allow reaching the relatively spacious “hot spot” region above dUMP where other known TS inhibitors in the folate pocket occupied (Figure 3.14b). A series of compounds were prepared with this in mind. Initial trials of expansion on R2 failed to yield promising results. However, a range of small substituents were tolerated on R3 vector and led to improvement in  $\Delta T_m$  as listed in Table 3.4. The largest increase of  $\Delta T_m$  was observed with an amide substitution (**A27**), which had a  $K_D$  value of 18.6  $\mu\text{M}$ . The crystal structure implied that the gain of affinity was attributed to formation of polar interactions with the indole ring of Trp109 and side chain oxygen of Asn112 (Figure 3.15b). Several larger substituents with 5- or 6-membered rings were developed, aiming to anchor in the “hot spot” region via pi stacking with the pyrimidine of dUMP. But no improvement was observed for these substituents. In the crystal structure, the morpholine and the phenyl connected to sulfur were well defined with electron density, whereas the substituents on R3 vector could not be observed crystallographically, indicating failure of anchoring to desired pocket.

**Table 3.4. Biophysical data of compounds A9 – A33**

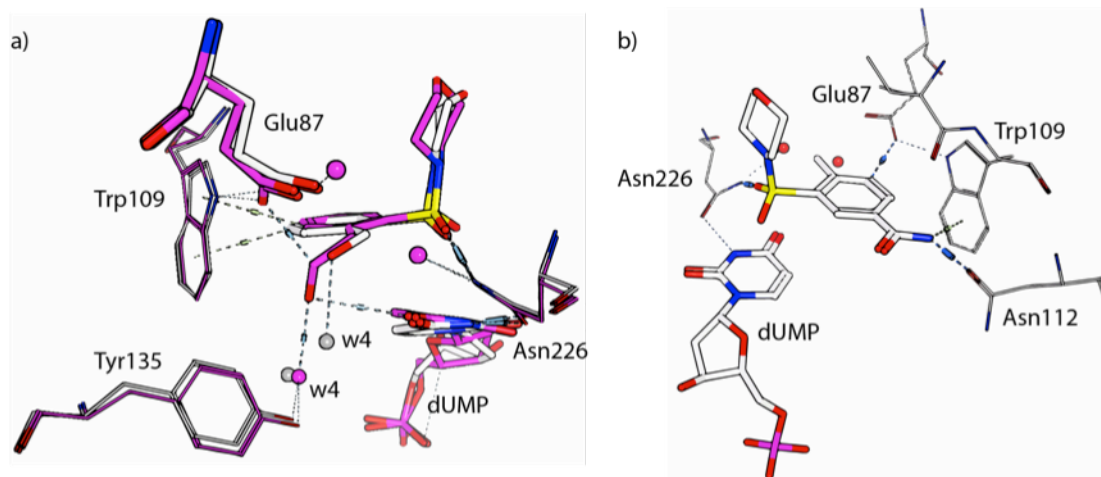
Cpd				DSF $\Delta T_m$ (°C)	$K_D$ ( $\mu$ M)
	R1	R2	R3		
A5		H	H	1.31	-
A9		H	H	2.14	-
A10		H	H	1.30	-
A11		H	H	2.97	121.5
A12		H	H	0.55	-
A13		H	H	0.60	-
A14		H	H	0.11	-
A15		H	H	4.00	19.1
A16		H	H	1.33	-
A17		H	H	0.81	-
A18			H	-0.72	-
A19			H	0.00	-
A20			H	-0.02	-
A21			H	0.66	-
A22		H		2.70	-



Cpd	R1	R2	R3	DSF $\Delta T_m$ (°C)	$K_D$ ( $\mu\text{M}$ )
A23		H		2.37	72
A24		H		2.53	50
A25		H		3.05	-
A26		H		2.14	-
A27		H		3.43	18.6
A28		H		1.70	-
A29		H		2.38	-
A30		H		1.77	-
A31		H		2.80	-
A32		H		1.99	-
A33		H		3.67	16.16

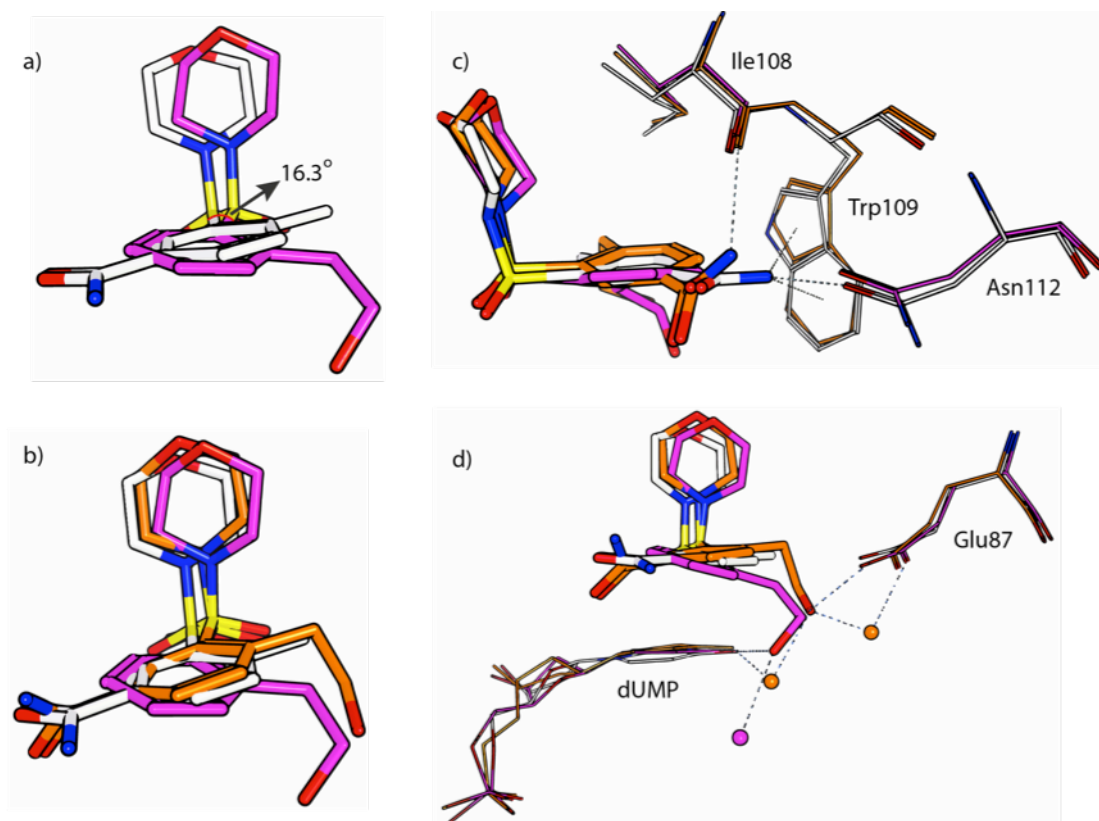
Note: DSF data was measured at 1 mM of compound in the presence of 500  $\mu\text{M}$  dUMP in HEPES buffer (20 mM HEPES and 150 mM NaCl, pH 7.5).

$K_D$  values were determined using ITC in the presence of 2 mM dUMP in HEPES buffer.



**Figure 3.15. Crystal structures of hTS in complex with A11, A15 and A27.** a) Back view of the binding site of **A11** (white) and **A15** (magenta) by expansion on R1 vector. **A11** displaces water w5 and forms hydrogen bond with water w4. **A15**, extension of **A11** by one more carbon atom, captures additional polar interactions with Glu87 and dUMP. b) Crystal structure of **A27** with an amide substituent on R3 vector. The amide NH serves as a hydrogen bond donor to Asn112 and also forms NH••• $\pi$  interaction with Trp109.

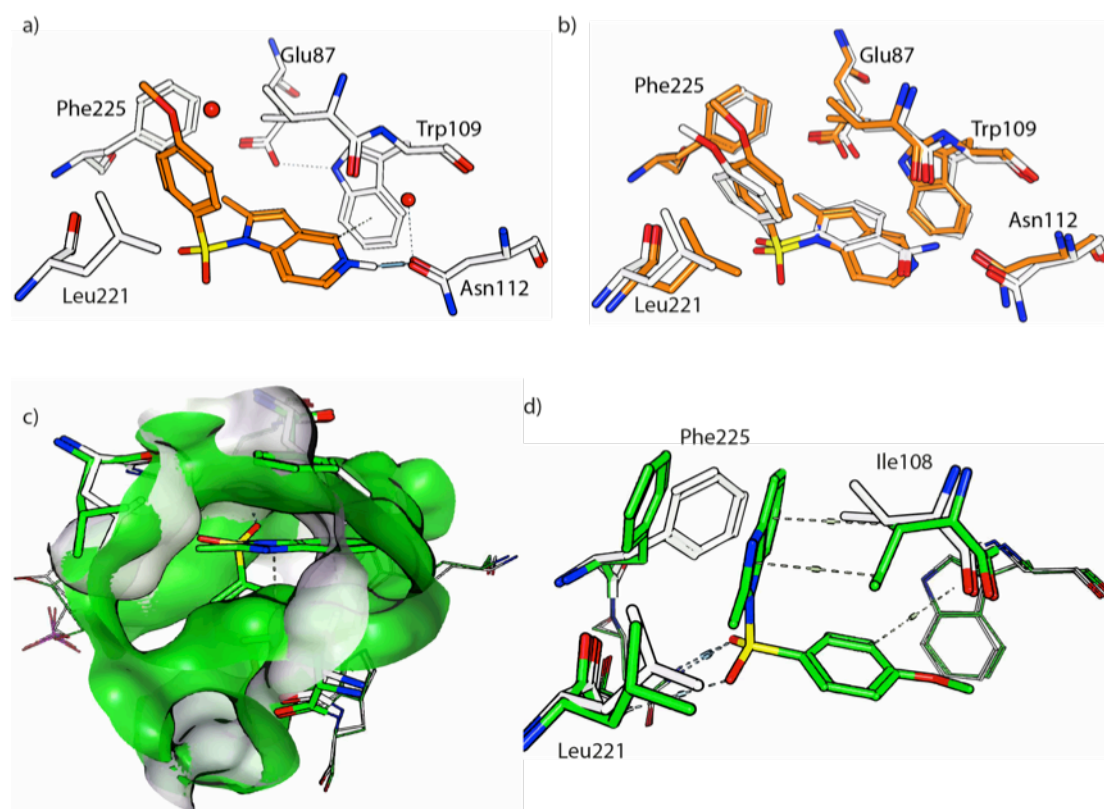
Having identified favorable substituents on vector R1 and R3, we superposed the crystal structure of compound **A15** with the one of compound **A27** (Figure 3.16a). The phenyl rings of the two compounds were not in the same plane. **A27** tilted the amide substituent down to interact with Trp109. Meanwhile, **A27** underwent a 16° rotation to get the amide closer to Trp109. Albeit not a satisfying overlapping between **A15** and **A27**, we designed **A33** in an attempt to merge the two compounds. Unfortunately, **A33** did not exhibit significant improvement in affinity compared with the parental compounds. The reason was subsequently revealed by the crystal structure. Possessing both R1 and R3 substituents, **A33** bound in a compromised conformation to accommodate them, which in turn led to a suboptimal binding conformation for both substituents. As shown in Figure 3.16d, resulting from the tilted phenyl ring, the 2-hydroxyethyl group on R1 was lifted up and sacrificed its hydrogen bond with dUMP. On the other hand, rotation of **A33** was hindered because of R1 substituent, resulting in the loss of amide NH••• $\pi$  interaction with Trp109 (Figure 3.16c).



**Figure 3.16. Crystal structure of hTS in complex with A33.** a) Superposition of **A15** (magenta) and **A27** (white) in the folate binding site. b) Superposition of **A15**, **A27** and **A33** (orange) in the binding site. c) Front view of compound binding site: due to the lack of rotation in **A33** (orange), the amide substituent on R3 vector is further from Trp109 and Asn112 as compared to **A27** (white). d) Side view of compound binding site: the phenyl ring of **A33** (orange) is tilted upward as compared to **A15** (magenta).

Since the results of SAR on R2 and R3 separately were not encouraging, we decided to explore a selection of fused heterocycles (Table 3.5). In order to reserve enough space for fused ring substituents as well as to merge type I and II binding modes, the morpholine in the entrance channel was replaced with the less bulky phenyl ring as in type II hit. It turned out that the phenyl scaffold behaved slightly better in the entrance channel (**A34**), reflected by the improvement in affinity ( $K_D$  of 9.09  $\mu\text{M}$  for **A34** compared with 18.6  $\mu\text{M}$  for **A27**).

Surprisingly, this series of compounds with a fused heterocycles induced comparable or even larger  $T_m$  shift than previous sets at a concentration of one magnitude lower. The DSF data suggested that the presence and position of nitrogen atom in the heterocycles affected the activity of compounds (**A35**, **A36** and **A37**) on hTS. We followed up with crystal structure determination and found out that the nitrogen atom affected the binding mode as well. The compound **A35** with an indole



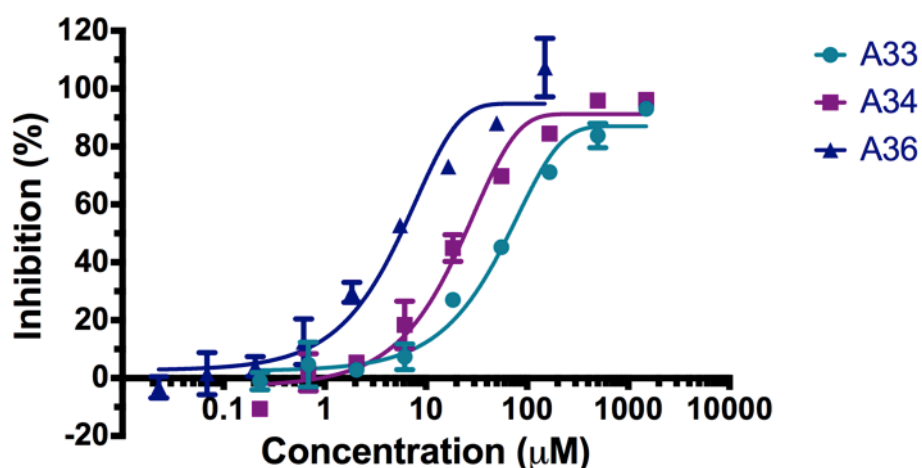
**Figure 3.17. Crystal structures of hTS in complex with compound **A36** and **A37**.** a) Crystal structure of **A36**: the N-7 of the heterocycle forms hydrogen bond with Asn112. b) Comparison of structure of **A36** (orange) with **A34** (white). c) Surface representation of the folate binding pockets in the structure of hTS-dUMP (white) and hTS-dUMP-**A37** (green). Shifting of Ile108 and Phe225 creates extra space for the binding of **A37**. d) Close view of contacts of **A37** with protein residues. The heterocycle is sandwiched between the shifted Ile108 and Phe225.

scaffold could not be clearly observed crystallographically and the conformation could not be identified, indicative of a lower affinity, which was in agreement with the lower  $\Delta T_m$  value. The introduction of a nitrogen atom at the 7-position of the heterocycle (**A36**) led to large increase of  $\Delta T_m$  and clear improvement of ligand electron density in the crystal structure, where the nitrogen introduced formed hydrogen bond with Asn112 (Figure 3.17a). Structural alignment of **A36** with **A34** revealed that the overall scaffold of the two compounds overlapped, with the methyl-substituent pointing toward Glu87 and NH acting as hydrogen bond donor to Asn112 (Figure 3.17b). Compared with **A34**, **A36** exhibited a 13-fold improvement in affinity ( $K_D = 0.69 \mu\text{M}$ ).

However, when the nitrogen was at the 4-position, the compound **A37** adopted a flipped conformation, of which the heterocycle resided in the entrance channel instead (Figure 3.17c). This was unexpected as the entrance was barricaded by three

hydrophobic residues (Ile108, Leu221 and Phe225) and thus too narrow to accommodate the heterocycle. In the crystal structure shown in Figure 3.17c and d, it was apparent that both Phe225 and Ile108 were displaced from their barrier positions, consequently assembling an elongated tunnel for the heterocycle. Furthermore, the new conformation of Ile108 and Phe225 constituted nice stacking interactions with the heterocycle. Taken together, compound **A37** represented a novel class of hTS inhibitor rendering conformational changes that have not been identified before.

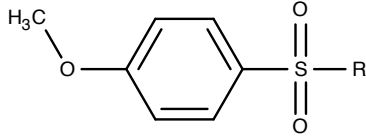
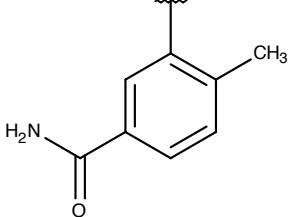
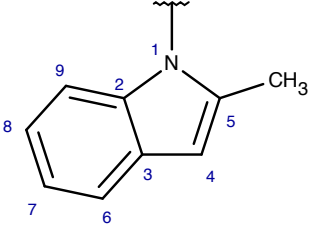
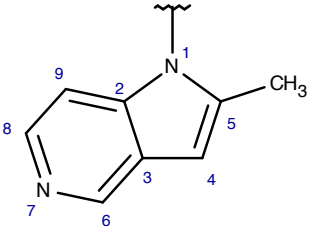
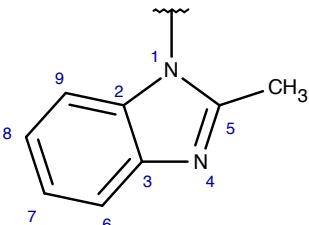
Apart from biophysical techniques, enzymatic assay was used as an orthogonal technique to evaluate follow-up compounds. The compounds were found to inhibit hTS enzymatic activity. The  $IC_{50}$  values for **A33**, **A34** and **A36** were determined to be 78.5  $\mu$ M, 26.1  $\mu$ M and 5.7  $\mu$ M, respectively (Figure 3.18). The data suggested that the rank potency of these compounds for inhibiting hTS activity correlated well with their potency in binding to the protein hTS.



	<b>A33</b>	<b>A34</b>	<b>A36</b>
$K_D$ ( $\mu$ M)	16.16	9.09	0.69
$IC_{50}$ ( $\mu$ M)	78.5	26.1	5.7

**Figure 3.18. Enzymatic activity inhibited by follow-up compounds A33, A34 and A36.** Enzymatic activities were measured in the presence of 100  $\mu$ M dUMP and 100  $\mu$ M 5,10- $CH_2$ -THF. The  $IC_{50}$  data correlated with the  $K_D$  values determined from ITC.

**Table 3.5. Biophysical data of compounds A34 – A37**

Cpd	R	DSF $\Delta T_m$ (°C) <sup>9</sup>	$K_D$ ( $\mu\text{M}$ ) <sup>10</sup>
			
A34		4.58	9.09
A35		1.92	-
A36		5.74	0.69
A37		2.88	-

<sup>9</sup> DSF data was measured at 125  $\mu\text{M}$  of compound in the presence of 500  $\mu\text{M}$  dUMP in HEPES buffer (20 mM HEPES and 150 mM NaCl, pH 7.5).

<sup>10</sup>  $K_D$  values were determined using ITC in the presence of 2 mM dUMP in HEPES buffer.

### 3.3.4.2 Expansion and optimization of type II hit

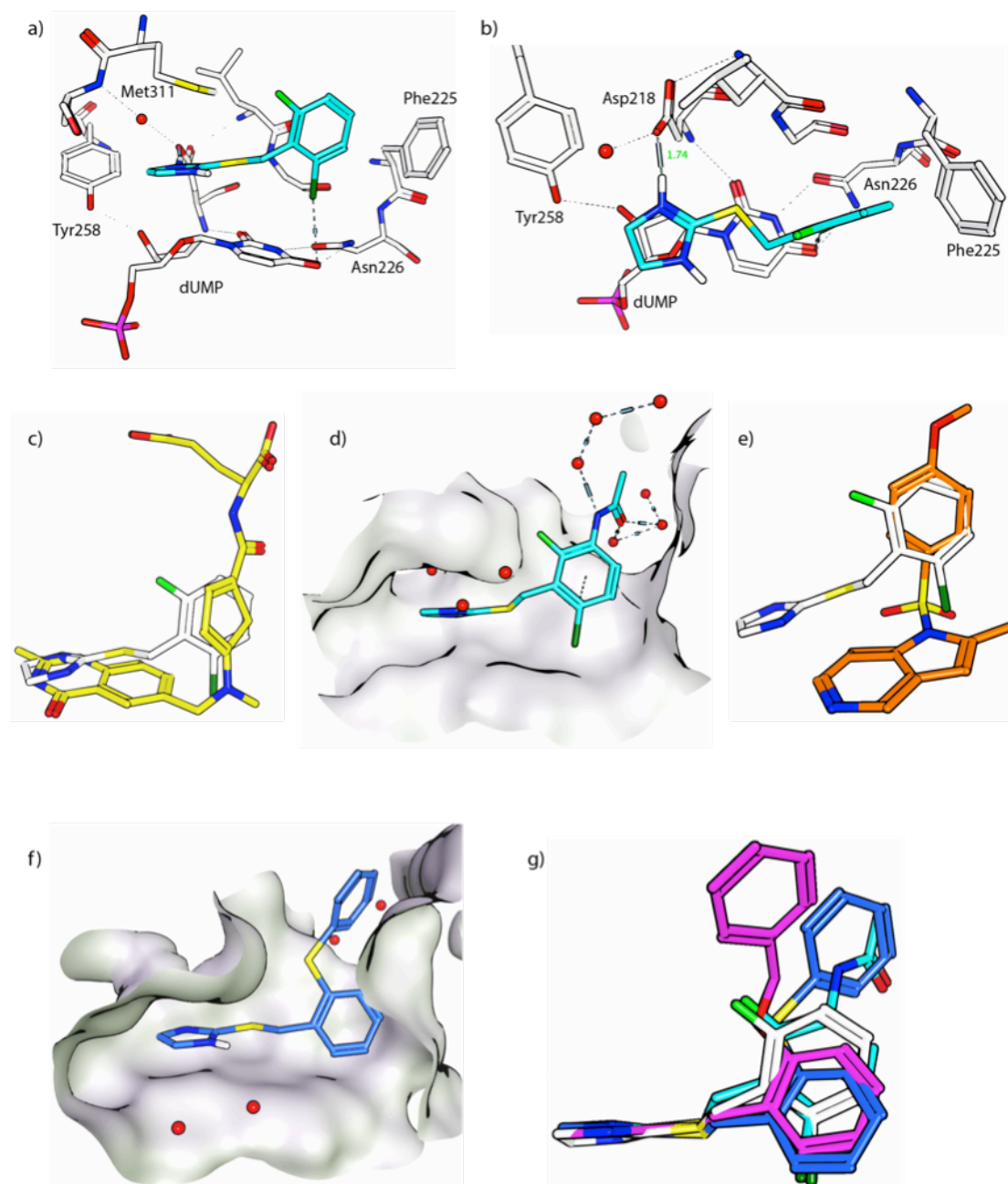
Similar to the three antifolates, binding of type II hit C993 involved interactions with the C-terminal residues and the C-terminal loop was clearly defined in the crystal structure. As illustrated in Figure 3.19a, the imidazoline ring of C993 was sandwiched between Met311 and the ribose of dUMP. The C-S-C fragment of Met311 was positioned approximately in a coplanar fashion above the imidazoline at van der Waals distance (3.68 Å to the nearest atom). This type of geometric arrangement has been found in the binding of known antifolates as well (Figure 3.8). Further analysis of the structure revealed that the NH–O (Asp218) distance was 1.74 Å, indicative of an electrostatic N–carbonyl oxygen interaction (Figure 3.19b). An investigation of effects of substituting the head ring was carried out with compound **B1-5** to assess the possibility of expansion on the ring (Table 3.6). Unfortunately, none of the five compounds produced better DSF data than the initial fragment. Replacing the head ring with planar imidazole derivatives (**B1-3**), the binding activity was totally abolished. If 1-methyl was introduced on the nitrogen of imidazoline ring, the compound **B4** could not induce positive  $T_m$  shift. But if 5-methyl was included (**B5**),  $T_m$  shift was partially recovered, giving a smaller shift size than the initial fragment, which was probably due to the steric clash of the methyl substituent with the protein. Taken together, it seems as the interaction between Met311 and the imidazoline ring is largely driven by the electronic nature of the ring system. Modifications on the nitrogen atoms or the electronic structure of the ring were not tolerated. We therefore focused our attention on the phenyl ring in the entrance channel for the following expansion efforts.

The phenyl scaffold sat perpendicular to the plane of the head ring, with its *ortho*-chloro pointing to dUMP and forming halogen bond with O-5 of dUMP (Figure 3.19a). Meanwhile, the *ortho*-fluoro sat close to Met311 with the acidic CH...F contact below van der Waals distance (about 2.4 Å), indicating a weak hydrogen bond character of this interaction. These two halogen bonds contributed largely to the ligand binding as replacing the halogens resulted in decrease of  $\Delta T_m$  values (**B6** and **B7**).

Superposing the complex structure of raltitrexed with the one of C993, it revealed that the *meta*-vector next to *ortho*-fluoro of the phenyl ring in C993

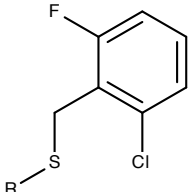
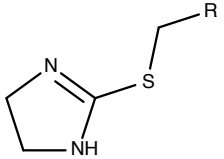
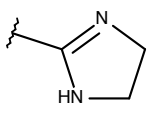
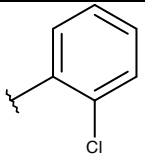
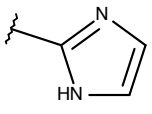
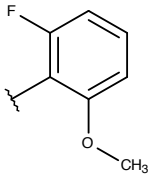
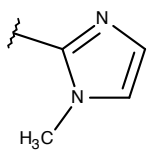
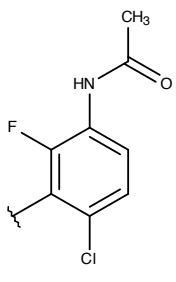
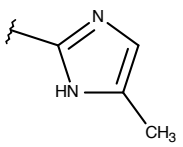
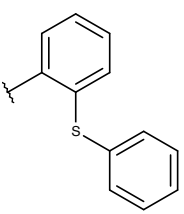
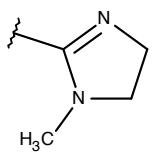
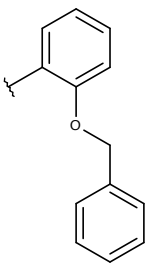
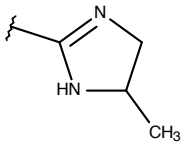
overlapped the anchor point of glutamate moiety in raltitrexed (Figure 3.19c). The crystal structure suggested that polar substituents should be well tolerated at this position as this vector pointed toward the polar entrance tunnel of the active site. Unsurprisingly, introducing an acetamide group at this vector (**B8**) led to an increase in affinity to 8.4  $\mu\text{M}$ . The structure showed that the acetamide group was hydrogen bonded to three water molecules in the entrance tunnel (Figure 3.19d).

Rather than continue with the polar space in the entrance tunnel, we started to explore the *ortho*-position for fragment merging. We reasoned that there was a great extent of atomic overlap between C993 and **A36** (Figure 3.19e) and it would make it possible to merge the key features of the two binding modes and generate novel compounds with high affinity. Owing to the difficulty in chemical synthesis, only two compounds (**B9** and **B10**) had been synthesized and evaluated at this point. Unfortunately, neither compounds showed improvement in binding according to DSF data. The crystal structures showed that both compounds failed to recapitulate the binding pose of **A34** as the *ortho*-substituents protruded into the water-filled entrance tunnel instead of the desired back pocket (Figure 3.19f). The *ortho*-substituents did not appear to make significant interactions directly with hTS other than van der Waals contacts. In the structures of **B9** and **B10**, there was also a notably different position for the phenyl core. The phenyl ring lied further down to dUMP as a consequence of steric hindrance from the *ortho*-substituents on binding (Figure 3.19g). The decrease of  $T_m$  shift could be the result of a combination of unfavorable perturbation of the water network with hydrophobic groups in the entrance tunnel and loss of interaction with dUMP.



**Figure 3.19. Crystal structures of hTS in complex with compounds of type II binding mode.** a) Side view of C993 binding mode. The head ring imidazoline is sandwiched between Met311 and dUMP. b) Top view of C993 binding mode. Salt bridge is formed between imidazoline and Asp218. c) Superposition of C993 (white) and raltitrexed (yellow) in the folate binding site. d) Crystal structure of hTS in complex with **B8**. The acetamide group binds into the water-filled entrance channel. e) Superposition of C993 and **A36**. f) The structure of hTS with **B9**. The phenyl ring anticipated for fragment merging binds in the entrance channel instead. g) Superposition of C993 (white) with elaborated compounds (**B8**, cyan; **B9**, blue; **B10**, magenta) in the folate binding site. Positional shifts are observed on the phenyl ring of **B9** and **B10**.

**Table 3.6. Biophysical data of compounds B1-B10**

							
Cpd	R	DSF $\Delta T_m$ (°C)	$K_D$ ( $\mu M$ )	Cpd	R	DSF $\Delta T_m$ (°C)	$K_D$ ( $\mu M$ )
<b>C993</b>		1.42	32.7	<b>B6</b>		0.82	-
<b>B1</b>		-0.01	-	<b>B7</b>		1.07	-
<b>B2</b>		-1.51	-	<b>B8</b>		3.40	8.4
<b>B3</b>		-0.53	-	<b>B9</b>		0.91	-
<b>B4</b>		-0.07	-	<b>B10</b>		0.86	-
<b>B5</b>		0.81	-				

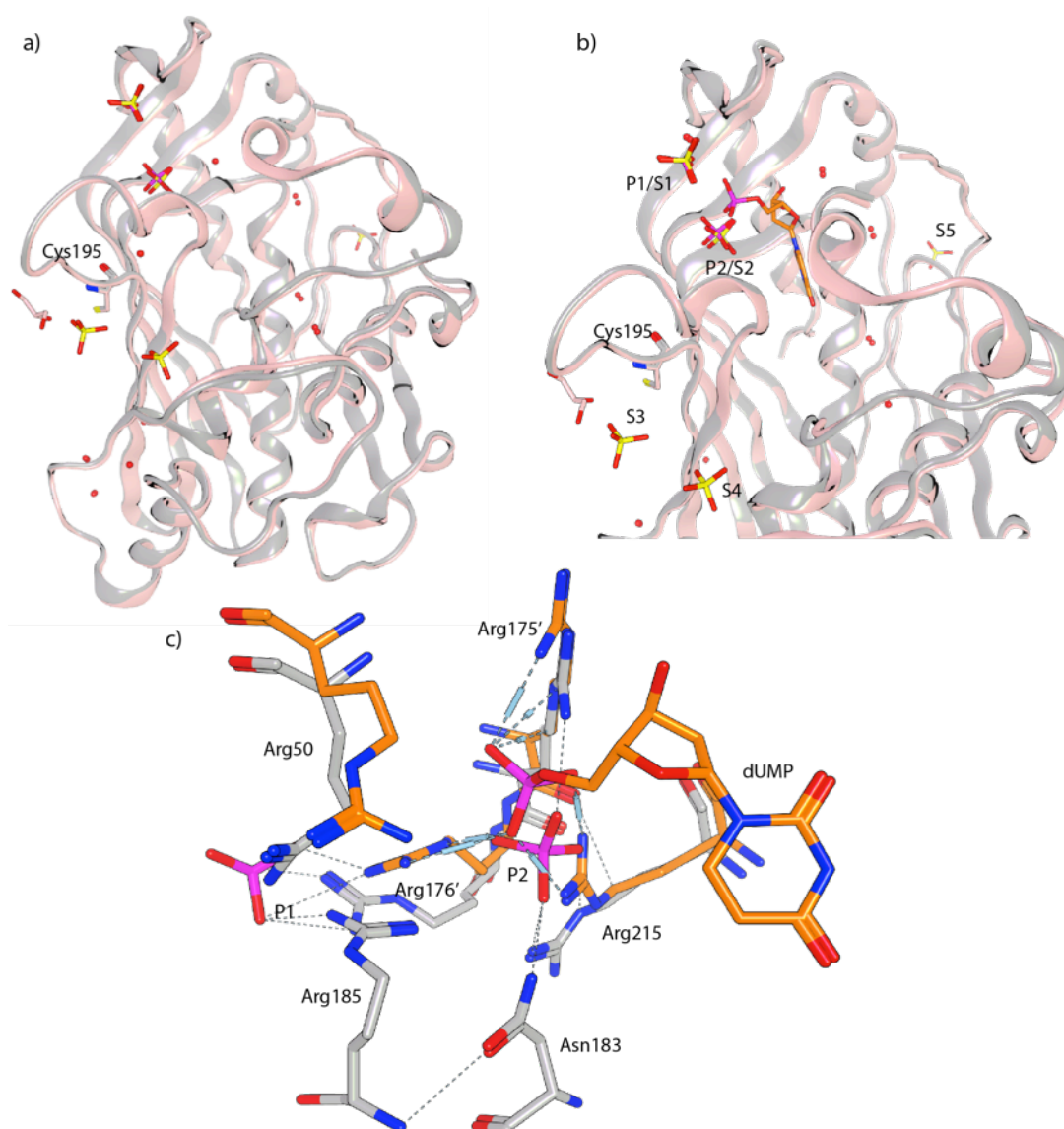
### **3.4 Characterization of inactive conformers of hTS**

#### **3.4.1 Crystal structures of mutant hTS in the inactive conformation**

As mentioned earlier, induced expression of hTS enzyme has been implicated as one of the drug resistance mechanisms to hTS inhibitors (Peters et al., 1995; Saga et al., 2003). Previous studies have shown that acute induction of hTS could be attributed to disruption of the auto-regulatory feedback process, which is mediated via the translational repression effect resulting from the interaction between hTS protein and its mRNA (Chu et al., 1991). Meanwhile, an inactive conformation of hTS has been found in the crystal structure obtained in the presence of high concentration of ammonium sulfate (Phan et al., 2001), in which the catalytic cysteine Cys195 is rotated away from the active site. Building off the translational repression model of hTS protein-hTS mRNA complex, it is possible that hTS functions as an RNA binding protein in its inactive conformation and catalyzes dTMP synthesis in the active conformation. hTS inhibitors, which stabilize the active conformation, might then prevent hTS-mRNA complex formation in its inactive state and thus release the translational repression, leading to overexpression of hTS protein. Herein we present the work on mutant hTS proteins that are in the inactive conformation, setting out to investigate the biological relevance (as an RNA binding protein in particular) of the inactive conformer of hTS protein.

Based on the study from Lovelace et al., two hTS mutants M190K and A191K in the inactive conformation are explored in this work (Lovelace et al., 2009). So far all the crystal structures of the inactive conformer of hTS published are obtained from a condition containing high concentration of ammonium sulfate (Cardinale et al., 2011; Lovelace et al., 2005; Lovelace et al., 2009; Phan et al., 2001). It is reasonable to question the relevance of these structures and the inactive conformations instead could be a crystallographic artifact due to the abundant presence of sulfate as several sulfate ions are observed in the structures. To further study the inactive conformation, we performed crystallization screening for the two mutants to establish new crystal forms. Unsurprisingly, most of the crystals obtained were from conditions containing various amount of ammonium sulfate. However, we managed to get crystals of A191K from a sulfate and phosphate-free condition (Table 3.7). The structures were subsequently solved and revealed that not only the two mutants crystallized from

sulfate-containing conditions but also A191K from the sulfate-free condition were in the inactive conformations.

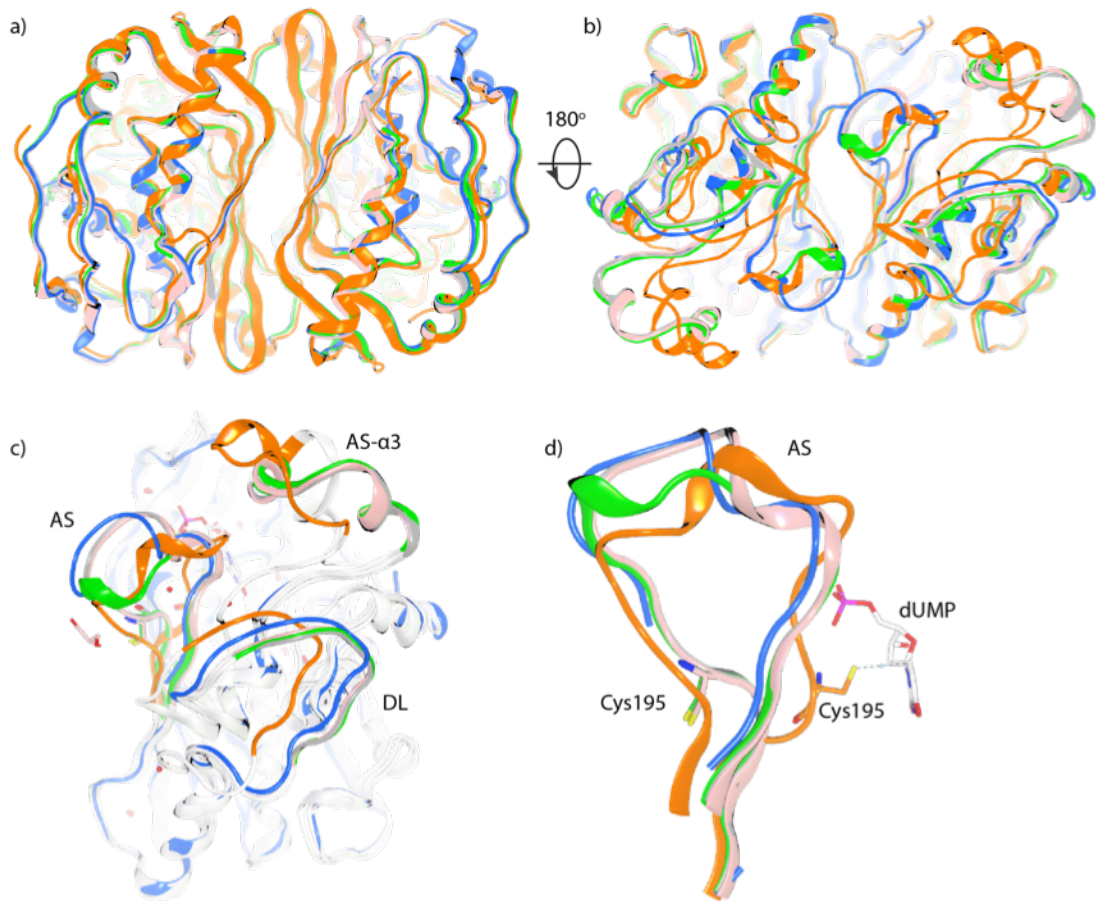


**Figure 3.20. Crystal structures of A191K-sf (sulfate-free) and A191K-s (sulfate containing).** a) Structural alignment of A191K-sf (grey) and A191K-s (pink). Both are in the inactive conformation with the catalytic cysteine Cys195 sitting outside the active site. b) Two phosphate ions bind in the structure of A191K-sf, labeled as P1 and P2. Five sulfate ions are found in the structure of A191K-s, labeled as S1-5. In addition, one glycerol molecule binds in the dimer interface, near to S3. c) Close view of the binding site for two conserved phosphate/sulfate ions. P2 is near to but different from the phosphate moiety of dUMP (orange). The four arginine residues (Arg50, Arg215, Arg175' and Arg176') that coordinate the phosphate moiety of dUMP (orange) undergo positional changes to coordinate the two phosphate ions in the inactive conformation (grey). Upon the conformational changes in the inactive conformation, two distal residues (Arg185 and Asn183) in the active conformation move close to the active site and interact with the phosphate ions.

Three mutant structures are used for following analysis: A191K crystallizing in the sulfate-free condition as A191K-sf; A191K crystallizing in a sulfate-containing condition as A191K-s; and M190K crystallizing in a sulfate-containing condition as M190K. Crystallographic refinement statistics are summarized in Table 3.7. The structures of A191K-sf and A191K-s were very similar, both adopting an inactive conformation with the catalytic cysteine Cys195 sitting outside of the active site (Figure 3.20a). This confirms that the inactive conformation is not due to high concentration of sulfate ions in the crystallization solutions. In the structure of A191K-sf, there were two extra tetrahedral densities near the dUMP binding site, in which we fitted with two phosphate ions. As illustrated in Figure 3.20c, one of the phosphate ions sat close to but different from the position of the phosphate moiety of dUMP. The highly conserved residues accommodating the phosphate moiety of dUMP (Arg50, Arg215, Ser216, Arg175' and Arg176') interacted extensively with the two phosphate ions in the inactive conformer as well, but all in a different position. Notably, two additional residues Asn183 and Arg185, which were about 18 Å away from dUMP in the active conformer, came in and interacted with the two phosphate ions in the inactive conformer. It was clear from the structure that the two phosphate ions played an essential role in stabilizing the inactive conformer. Since there was no phosphate added during protein purification or crystallization, the phosphate ions in the structure were picked up in the host cell and co-purified with the protein. In the structure of A191K-s monomer, five sulfate ions were identified, one binding in the protein surface and four in the vicinity of the active site. Of note was that two of the sulfate ions overlapped the two phosphate ions in A191K-sf and another two bound in two undiscovered sites, between the active site (AS) loop and DL loop (Figure 3.20b). Moreover, there was one glycerol molecule from the cryo-solvent bound near to one sulfate ion and buried in the dimer interface cavity, which was absent in the structures of A191K-sf, M190K and other inactive conformers of hTS.

Superposition of the three inactive mutant dimers with inactive wild-type hTS (PDB 1HW3) and the active dimer of hTS-dUMP showed that the dimer interface  $\beta$ -sheet core and the region near N-termini were relatively conserved between the active and inactive states (Figure 3.21a). The structural differences were mainly located on the opposite site, with the major and most apparent differences in the active site (AS)

loop, DL loop and AS- $\alpha$ 3 as highlighted in Figure 3.21b and c. An interesting finding was that the structures of A191K, M190K and the inactive hTS were divergent in these three protein segments, though all adopted an inactive conformation (Figure 3.21d). It appeared that the AS loop first twisted the ends of the loop and relocated to the inactive state observed in the structure of 1HW3 together with the “back off” of the DL loop to allow subsequent loop rotation. After that the AS loop underwent further conformational refinement by shifting away from the dimer interface to the inactive state of A191K and then rewound residues Leu187-Leu192 to the inactive form of M190K. All the conformational rearrangements in the AS loop were coupled with the adjustment in DL loop and AS- $\alpha$ 3 region by moving away from their original positions that would block the inactive conformation of the AS loop sterically. As a consequence, the inactive dimers of hTS were less compact and the dimer interface became more solvent accessible. It should be noted that the AS- $\alpha$ 3 and DL segments contained the variable insert region I and II as compared to other TS enzymes, respectively. The AS- $\alpha$ 3 was disordered in the inactive state of 1HW3 while it was clearly defined in a new conformation in A191K and M190K. The region (Lys107-Ser114) connecting the helix remained disordered in all structures of inactive conformers, but it constituted a helix (AS- $\alpha$ 2) in the active conformer and was involved in the conformational changes upon folate binding. The C-terminal loop was relatively defined in spite of the lack of density for the last two residues Ala312 and Val313. Similar to the active state of hTS, the N-terminal region was disordered in the inactive states.



**Figure 3.21. Structural comparison of the active and inactive conformations of hTS.**  
 a) Front view of the structural alignment of A191K-sf (grey), A191K-s (pink), M190K (green), hTS-dUMP (orange) and the inactive crystal structure of hTS (PDB: 1HW3; blue).  
 b) Back view of the structural alignment. Significant segmental shifts are observed at this side.  
 c) Alignment of the inactive and active conformers with regions undergoing significant conformational changes highlighted. AS, active-site loop, residues 181-197; DL, dimer interface loop, residues 146-153; AS-α3, active-site helix 3, residues 117-128.  
 d) Structural comparison of the AS loop between the active and inactive states of hTS.

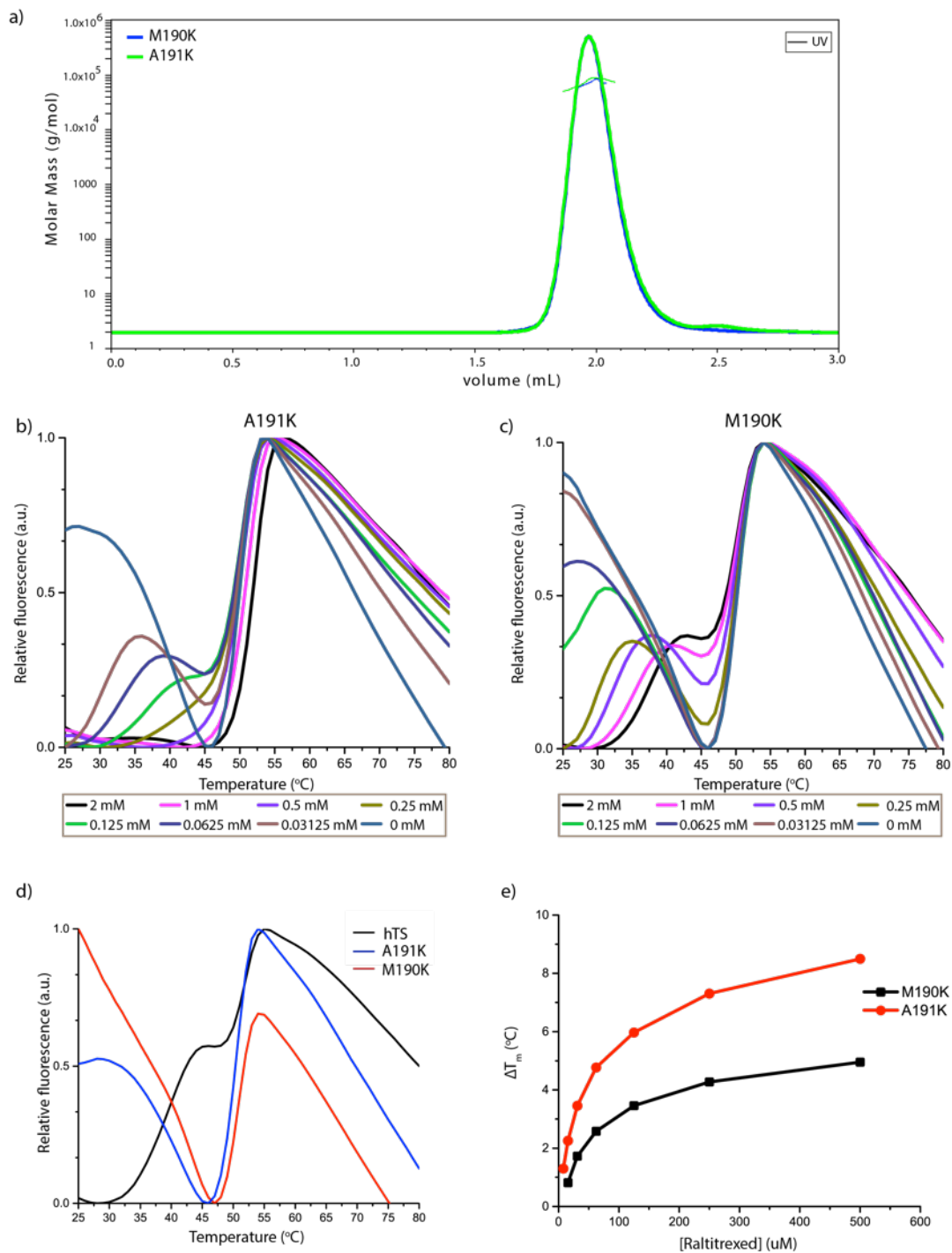
**Table 3.7 Data collection and refinement statistics of structures of A191K and M190K**

Dataset	A191K-sf	A191K-s	M190K
Crystallization condition	100 mM HEPES, pH 7.0, 10% PEG 6000	100 mM Tris, pH 8.5, 0.2 M lithium sulfate, 1.26 M ammonium sulfate	100 mM MES, pH 6.5, 1.6 M ammonium sulfate, 10% 1,4-Dixane
Data collection			
X-ray source	AS-MX1	AS-MX1	AS-MX1
Wavelength, Å	0.9537	0.9537	0.9537
Space group	P3 <sub>1</sub> 21	P3 <sub>1</sub> 21	P3 <sub>1</sub> 21
Cell parameters			
a, b, c in Å	95.62, 95.62, 79.68	95.56, 95.56, 81.13	95.96, 95.96, 80.90
α, β, γ in degree	90, 90, 120	90, 90, 120	90, 90, 120
Resolution range, Å	30 – 1.98 (2.05 – 1.98)	30 – 2.31 (2.39 – 2.31)	30 – 2.10 (2.18 – 2.10)
No. observed reflections	155 049	128 528	312 234
No. unique reflections	27 905	18 438	25 623
Completeness, %	94.1 (68.2)	95.5 (73.7)	99.8 (100.0)
<i>R</i> <sub>merge</sub> , %	8.5 (41.6)	8.5 (35.1)	6.4 (58.3)
Average I/σ	17.4 (2.7)	18.4 (4.3)	40.2 (4.5)
Refinement statistics			
Resolution range, Å	30 – 1.98	30 – 2.31	30 – 2.10
<i>R</i> <sub>work</sub> , %	16.90	17.29	18.74
<i>R</i> <sub>free</sub> , %	20.80	21.71	22.79
Protein/asymmetric unit	1	1	1
Ligand	2 phosphate ions	5 sulfate ions and 1 glycerol molecule	5 sulfate ions
Water molecules	200	125	222
Root mean square deviations			
Bond lengths, Å	0.008	0.008	0.008
Bond angles, degree	1.085	1.068	1.104

### 3.4.2 Biophysical characterization of mutant hTS

The asymmetric units in all structures of inactive hTS conformers were monomers. According to our MALS data, the two mutants M190K and A191K were dimers as wild-type hTS (Figure 3.22a). When examined using DSF, both M190K and A191K had high initial fluorescence, indicating that the mutants folded with large area of hydrophobic patches exposed. This was in agreement with the crystal structures, where the two inactive conformers opened the DL loop that served as a gate of the dimer interface core and thus exposed the hydrophobic  $\beta$ -sheets. However, it was apparent that A191K gave a lower initial fluorescence than M190K (Figure 3.22d), which could be correlated to the conformational differences between the two mutants reflected in the crystal structures.

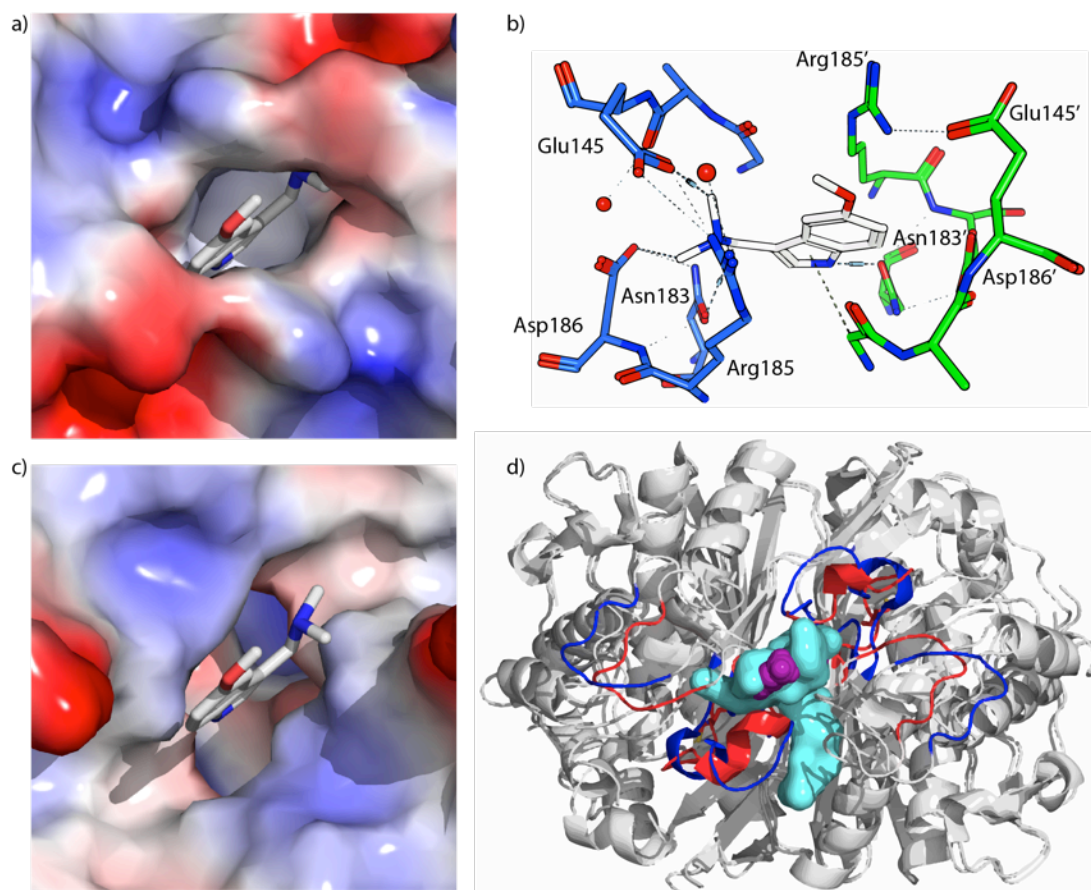
It was also notable that in a dUMP dose-response experiment the melting profiles of the mutants were gradually transformed to biphasic curves as observed for wild-type hTS (Figure 3.22b and c), implying that the mutants were reversed to the active state. A191K exhibited a monophasic profile when the concentration of dUMP reached 250  $\mu$ M, but M190K failed to achieve that even at 2 mM of dUMP. Similarly, raltitrexed invoked higher  $T_m$  shifts on A191K than M190K when tested in the same conditions. The DSF data suggested that M190K had significantly lower affinity for dUMP, which in turn affected the binding of raltitrexed.



**Figure 3.22. Biophysical characterization of M190K and A191K.** a) MALS data for M190K and A191K. The mean molar masses determined for M190K is 72.2 kDa and for MA191K 73.5 kDa, corresponding to the dimer sizes. b) dUMP dose-response data on A191K. c) dUMP dose-response data on M190K. d) Melting profiles of hTS, A191K and M190K. High initial fluorescence is observed for M190K and A191K. e) Raltitrexed dose-response data on A191K and M190K in the presence of 1 mM dUMP.

### 3.4.3 Evaluation of type IV hits in mutant hTS

During fragment screening, we identified two type IV fragments (F13 and C1440) bound in the dimer interface of the active hTS conformer. As shown in the structure of hTS in complex with F13 (Figure 3.23a and b), type IV hits bound to a fairly polar cleft in the dimer interface formed by Glu145 (in the DL loop) together with Asn183, Arg185 and Asp186 (in the AS loop) from each hTS monomers. Nevertheless, this polar cleft disappeared and was replaced with a larger and relatively hydrophobic pocket in the inactive conformers due to conformational changes (Figure 3.23c). There have been several ligands reported in the dimer interface pocket of inactive hTS conformers (PDB: 3N5E, 4GD7, 4G2O and 4G6W). However, our type IV hits were the first ligands identified to bind in the dimer interface of the active hTS conformer. Structural superposition revealed that the binding position of F13 partially overlapped the positions of ligands in the inactive conformers (Figure 3.23d), indicating that the dimer interface of hTS was able to accommodate different types of ligands in the active and inactive states. We followed up to examine type IV hits on the two mutants. The DSF data showed that F13 led to a concentration-dependent reduction of the initial fluorescence intensity on M190K and A191K (Figure 3.24), indicative of closing up of the dimer interface. This suggested that F13, which bound to the dimer interface of active hTS, could shift the conformational equilibrium of hTS to the active state. Extrapolating from the structural and DFS data, it is possible that the novel dimer interface cavity serves as an allosteric site of the enzyme and binding of specific types of ligands at this site can shift the equilibrium from the active to the inactive state of hTS.



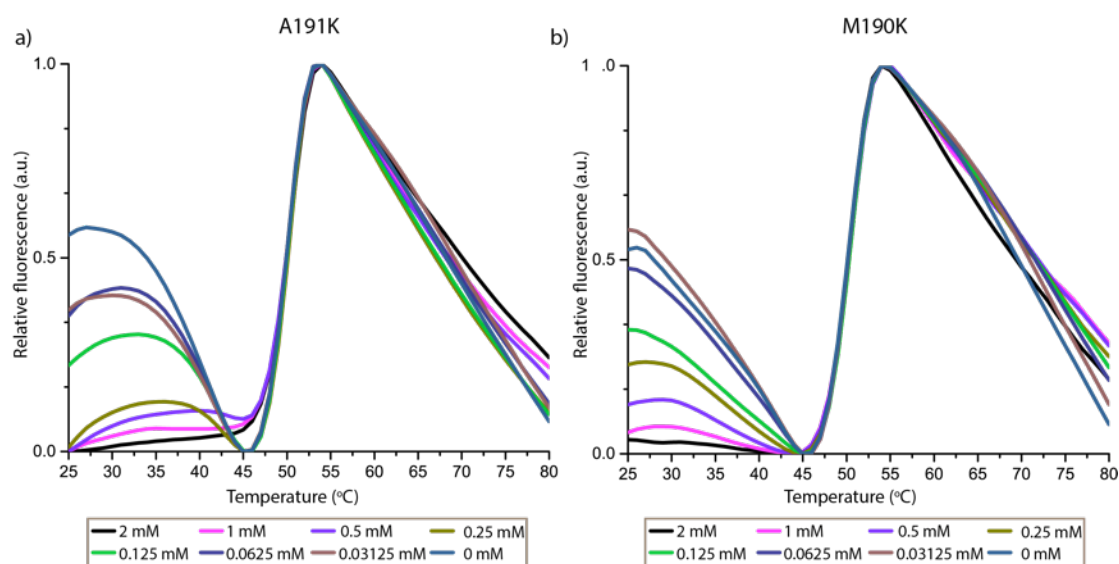
**Figure 3.23. Crystal structures of the dimer interface binders in the active and inactive states of hTS.** a) Electrostatic potential surface for the binding region of F13 in the dimer interface. b) Interactions of F13 with protein residues. c) Electrostatic potential surface of the F13 binding site in the inactive conformation. d) Structural alignment of hTS-F13 and hTS-peptide (PDB: 3N5E). The peptide (shown as cyan surface) binds in the dimer interface of the inactive conformation of hTS (AS and DL loop shown in blue). F13 binds (shown as purple surface) binds in the dimer interface of the active conformation of hTS (AS and DL loop shown in red).

#### 3.4.4 Investigation of RNA binding to hTS

In previous investigations, the 5'-upstream sequence containing the AUG start codon of hTS-mRNA has been predicted to form a stable stem-loop secondary structure (Figure 3.25a) and shown to interact with hTS protein to repress its own translation (Brunn et al., 2014; Lin et al., 2000). In this study, the interaction of this RNA sequence TS33 was studied on the mutant hTS proteins to investigate the possible roles of inactive hTS conformers in RNA binding. When examined using DSF, TS33 resulted in negative  $T_m$  shifts in the two mutants as well as the wild-type hTS, while the random RNA sequences did not trigger significant difference on the proteins (Figure 3.25b, c and d). It was evident that the shift affected only the second

phase of wild-type hTS. Also, there was no significant effect on the initial fluorescence of the melting profiles. The data suggested the formation of interaction between TS33 and both wild-type and mutant hTS, but there was no evidence showing the conformational switching between the active and inactive states upon the interaction.

The interaction study was subsequently followed up with native PAGE analysis. Even though small difference was observed for RNA mobility when TS33 was incubated with A191K, there was no significant mobility shift detected for the protein A191K after incubation with RNA (Figure 3.25e and f). Hence, the results were not conclusive. To gain direct insight into the molecular interaction between RNA and hTS, we attempted co-crystallization of the mutant and wild-type hTS with TS33 but had so far not obtained any diffracting crystals. Soaking with TS33 was also applied to hTS, M190K and A191K crystals. All the crystals cracked in a few minutes after TS33 was added, which is often an indication of interaction though there were no significant conformational changes and ligand density observed in the structures with short-time soaking.



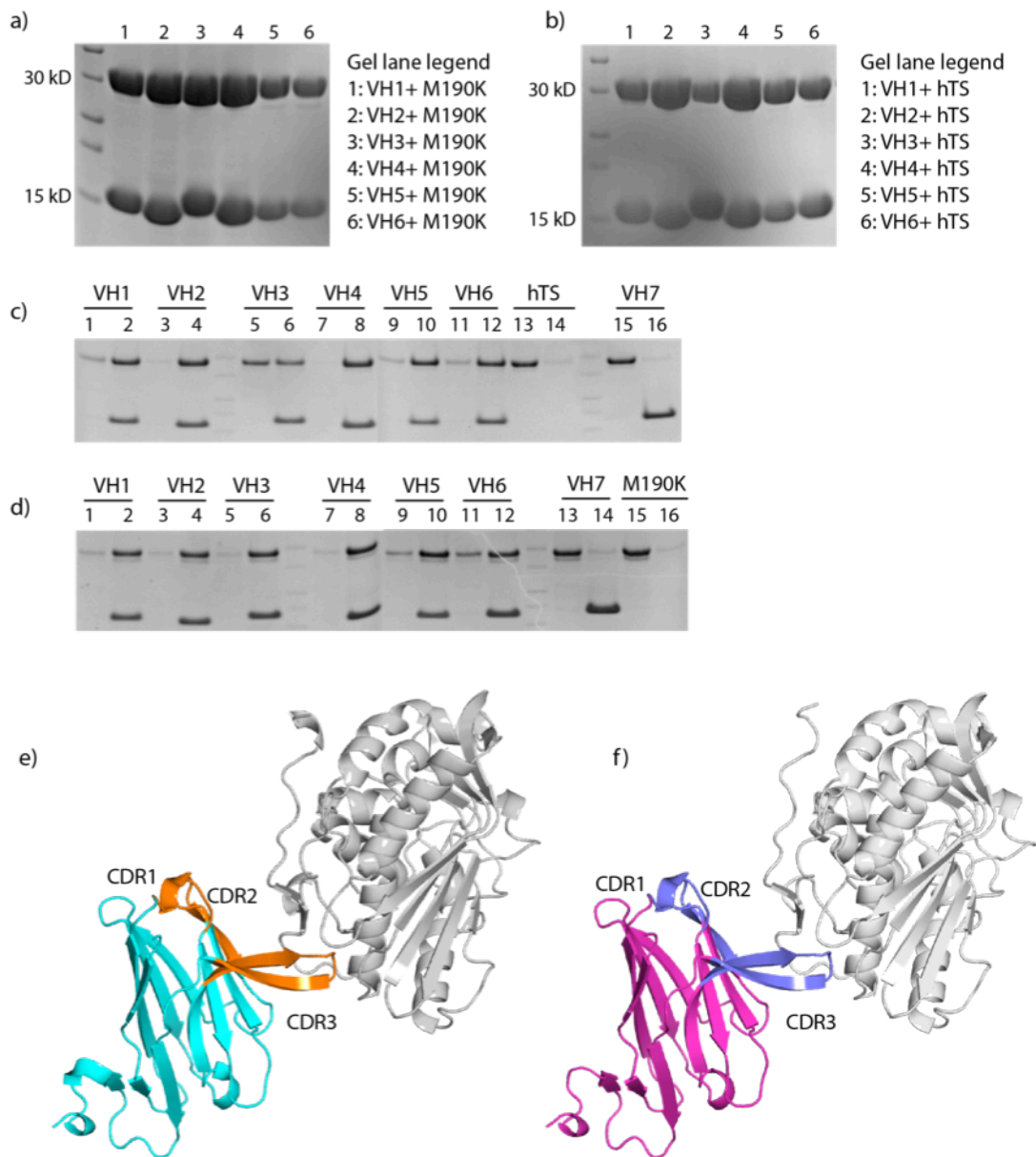
**Figure 3.24. DSF data of F13 on M190K and A191K.** a) F13 dose-response data on A191K. b) F13 dose-response data on M190K.



### 3.4.5 Identification of hTS binders from VH phage-display library

Having demonstrated that there was a conformational equilibrium between the active and inactive state in purified recombinant hTS protein, we were interested in developing antibodies that recognize the inactive conformer specifically. Inactive conformer-specific antibodies could be useful tools to detect the existence of the inactive conformer *in vivo* and identify ligands disturbing the conformational equilibrium. Phage display strategy was applied to the inactive hTS mutant M190K using a library of single-heavy chain variable-fragment antibody VH developed in our laboratory. After four rounds of panning, six VH clones showing stronger binding to M190K than hTS by phage ELISA were expressed in *E. coli* cells and purified for further characterization.

To further evaluate the binding of purified VH binders (VH1-6), pull-down and co-immunoprecipitation assay were performed with both M190K and hTS. The results showed that all the six purified VH binders could bind both M190K and hTS, but VH3 had apparent higher affinity/specificity to the mutant M190K (Figure 3.26). Following up the positive interaction results, we proceeded with crystallization screening of the VH binders with both M190K and hTS. At present, we have obtained crystals and solved the structures of hTS-VH2 and hTS-VH4. Attempts to crystallize M190K with VH binders (VH3 particularly) have not been successful so far. As shown in Figure 3.26e and f, VH2 and VH4 recognized hTS in a similar manner: the complementarity determining region (CDR) 3 of the VH binders bound in a cleft near to the central long helix of hTS. This region remained invariant between the active and inactive states of hTS and it explained why VH2 and VH4 showed binding to both M190K and wild-type hTS in the interaction assays.



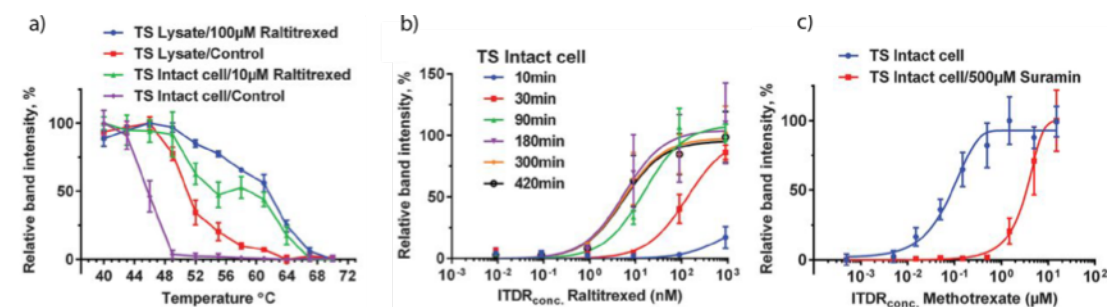
**Figure 3.26. Interaction assays and crystal structures of VH binders with hTS.** a) CoIP assay data for the six VH binders with M190K. b) CoIP assay data for the six VH binders with wild-type hTS. c) Pull down assay data for VH binders with wild-type hTS. Lane 1, 3, 5, 7, 9, 11, 13 and 15 are the flow-through samples and lane 2, 4, 6, 8, 10, 12, 14 and 16 are the elution samples. VH7 is an unrelated VH binder used as a negative control. d) Pull down assay data for VH binders with M190K. e) The crystal structure of VH2 (cyan) in complex with hTS (grey). The three CDRs are colored in orange. f) The crystal structure of VH4 (magenta) in complex with hTS (grey). The three CDRs are colored in purple.

### 3.5 Collaborative work

Apart from the work presented above, I have been involved in two collaborative projects on hTS using cell-based assays, one of which has been published (Molina et al., 2013) and the other submitted (out for review in Nature Chemical Biology).

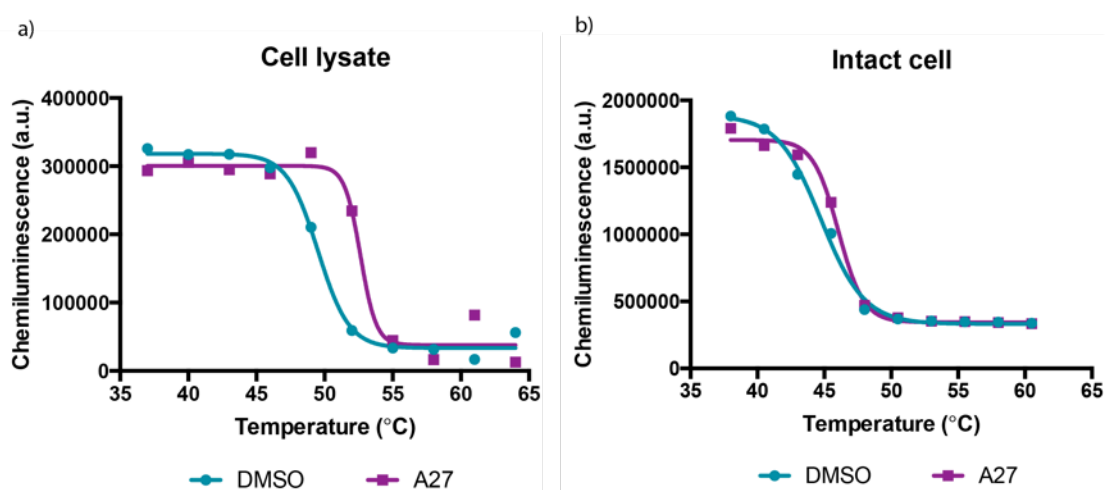
#### 3.5.1 Cellular thermal shift assay (CETSA)

A key parameter to optimize throughout medicinal chemistry efforts is to achieve optimal drug target engagement at the intended site of action inside cells. Our group has developed a novel label-free method for evaluating target engagement directly in cells and tissues. Similar to thermal shift assays on purified proteins, this technique is based on the biophysical principles of thermal stabilization of target proteins induced by ligand binding, and thus named the cellular thermal shift assay (CETSA). CETSA can provide direct information on cellular target engagement and constitutes a methodology that is likely to be valuable throughout the drug discovery process. Using this approach, a range of critical factors that are important for drug development against hTS can be monitored, including drug transport, drug activation and drug resistance (Figure 3.27).



**Figure 3.27. CETSA experiments on hTS.** a) Melting curves in intact cells versus cell lysate for hTS with raltitrexed. b) Time-dependent isothermal dose response fingerprint (ITDRF<sub>CETSA</sub>) of raltitrexed on hTS in whole cell experiments. Time dependence of response suggests a highly active transport of raltitrexed into cells. c) ITDRF<sub>CETSA</sub> of raltitrexed on hTS in cells treated with and without folic acid synthesis inhibitor suramin. Blocking of folic acid synthesis with suramin inhibits the activation of methotrexate, reflected by the shift of ITDRF<sub>CETSA</sub>. Figure retrieved from (Molina et al., 2013).

The purified protein differs from the target protein in its cellular context in certain properties such as post-translational modifications, metabolites levels and molecular interactions with other partners. To assess the target engagement of compounds from our FBDD in a physiology-relevant setting, CETSA experiments were performed in intact mammalian cells and cell lysate. Compound **A27**, a 19  $\mu\text{M}$  affinity compound in type I ligand series, invoked a thermal shift of about 3  $^{\circ}\text{C}$  on hTS in cell lysate (Figure 3.28a). In the whole cell experiment, the shift was smaller but significant, supporting that **A27** binds to hTS inside cells after three hours of treatment (Figure 3.28b). The CETSA data indicate that this series of compounds is able to target hTS in the cellular context with promising cell penetration properties.

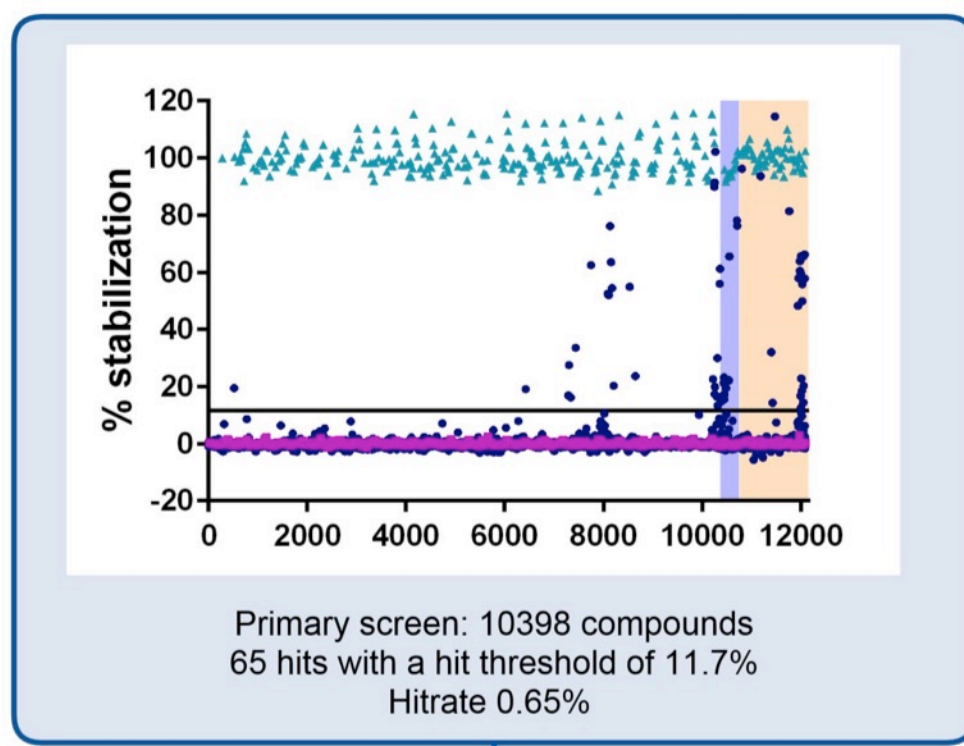


**Figure 3.28. CETSA data of compound A27 on hTS.** a) Melting curves of hTS in cell lysate with 1 mM **A27**, using DMSO as control. b) Melting curves of hTS in intact cells treated with 1 mM **A27**, using DMSO as the treatment control.

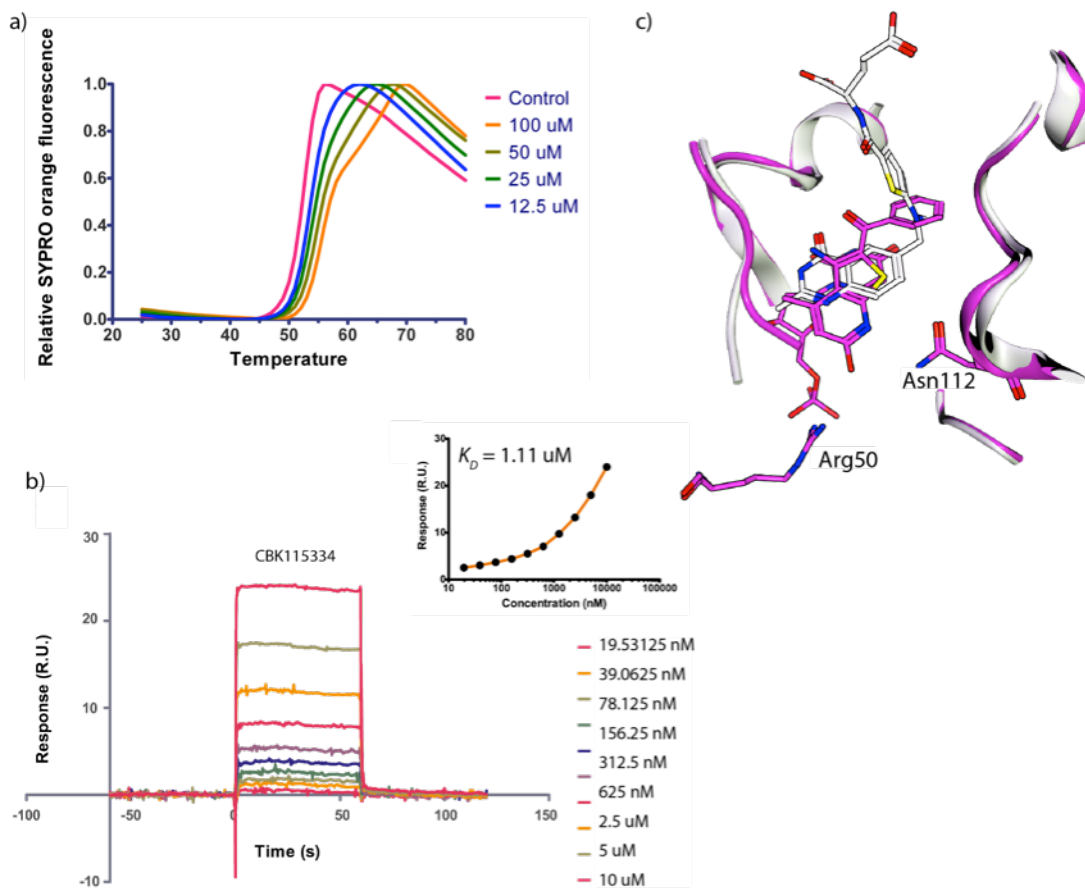
### 3.5.2 Characterization of hits from CETSA based screening

In the second collaborative project, a pilot screen using high throughput CETSA was performed for hTS with 10,398 compounds from the primary screening set at Chemical Biology Consortium Sweden (CBCS; [www.cbcs.se](http://www.cbcs.se)) (Figure 3.29). In addition to the known hTS inhibitors (such as the fluoropyrimidines and antifolates), the screen identified a set of compounds in clinical use but not known to affect hTS. Studies using liquid chromatography coupled to tandem mass spectrometry (LC-MS) indicated that the majority of these compounds were prodrugs that undergo intracellular metabolism before binding to hTS. The CETSA based screen also

identified a group of compounds with novel chemistry. I was involved in the structural and biophysical characterization of these compounds using purified protein. One of the compounds, CBK115334 gave relatively larger DSF  $\Delta T_m$  on purified hTS (Figure 3.30a). Further characterization of this compound using SPR showed an affinity of 1.11  $\mu\text{M}$  to purified hTS (Figure 3.30b). The crystal structure of hTS with CBK115334 was subsequently solved, revealing that CBK115334 bound in the folate binding site of hTS (Figure 3.30c). The binding involved pi-stacking interactions with the substrate dUMP and polar interactions with residues lining the catalytic cavity (Asn112 and Arg50 in particular), which constituted a novel mode of binding as compared to known TS inhibitors. Together this suggests that CBK115334 could be another good start point for generation of cell active hTS ligands.



**Figure 3.29. Pilot screen using CETSA-HTS to measure target engagement of hTS.** Scatter plot illustrating normalized screen data for 50  $\mu\text{M}$  library compounds (●), where 0% corresponds to hTS levels observed in the presence of DMSO only and 100% corresponds to hTS levels observed in the presence of 100  $\mu\text{M}$  raltitrexed (▲). The screening set contains a nucleoside subset (highlighted in blue) and the Prestwick drug set (highlighted in orange).



**Figure 3.30. Biophysical data for CBK115334.** a) DSF melting curves of CBK115334 in the presence of 500 μM dUMP. b) Sensorgrams and steady state affinity data fitting of CBK115334. c) Novel binding mode of CBK115334 (magenta) in the folate binding site of hTS compared with raltitrexed (white).

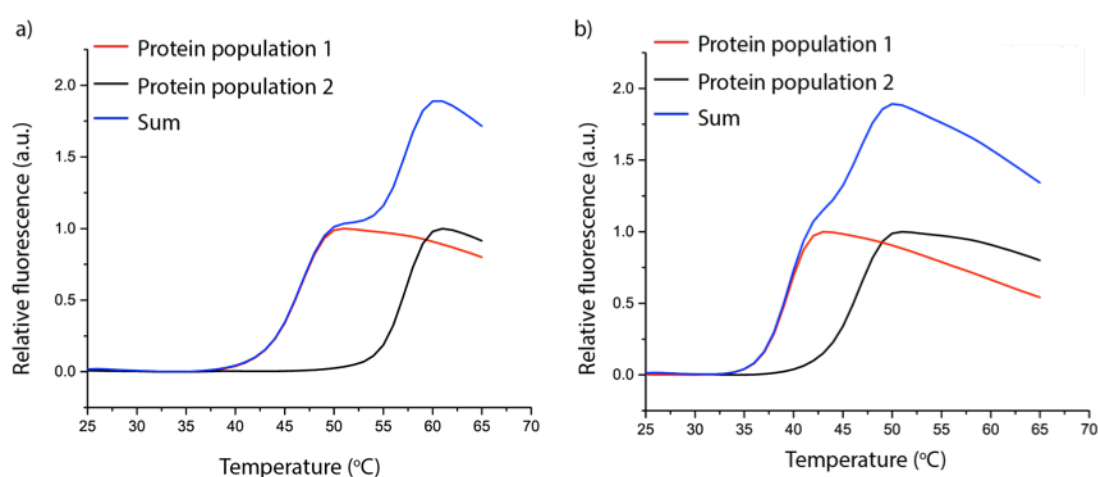
## CHAPTER 4 DISCUSSION

### 4.1 DSF as an informative characterization technique for hTS

Thermal shift assays measure changes in protein thermal stability, which have been used to identify optimal stabilizing buffer or protein formulation for crystallization and to detect ligand binding (Ericsson et al., 2006). Two *in vitro* thermal shift assays have been applied in the current studies: DSF and DSLS. DSF is a fluorescence-based assay and monitors protein thermal unfolding by utilizing an environmentally sensitive dye, which has low fluorescence in aqueous solution but high fluorescence when in hydrophobic environment. Distinct from DSF, DSLS monitors the denaturation and subsequent aggregation of unfolded proteins via light scattering measurement. Protein melting temperature  $T_m$  from DSF and aggregation temperature  $T_{agg}$  from DSLS do not necessarily correlate, since different denaturation processes are measured. In the study of hTS, both DSF and DSLS are able to detect and demonstrate sequential binding of dUMP and antifolate to hTS. However, the DSF provides more complex but better characterization data on hTS.

When evaluated using DSF, hTS exhibits a biphasic melting profile, of which the first phase can be stabilized by the presence of dUMP or phosphate ion and the second phase becomes less apparent by altering the C-terminal residues. On the basis of biophysical and structural data, two models for the biphasic melting profile are proposed. One is that the biphasic profile is the cumulative result of two or more melting curves from different protein populations. Not considering that there could be a subpopulation of hTS in the inactive conformation at this point, the protein examined consists of at least two main populations: unliganded hTS and hTS bound with dUMP. This is demonstrated by the ITC data as well as the presence of dUMP in the crystal structures while no dUMP is added for crystallization. As shown in Figure 4.1a, the ultimate protein melting profile observed is the sum of the two melting curves and becomes a biphasic curve. Addition of dUMP or phosphate ion increases the population of liganded hTS and therefore its contribution in the observed melting profile, which is translated into the concentration-dependent shift of the first phase to the right and ultimately to a monophasic melting curve when all protein are liganded at high concentrations. Building off this model, the biphasic curve in the sole presence of raltitrexed at high concentrations could result from the existence of two protein

complexes: stable ternary complex hTS-dUMP-raltitrexed and less stable binary complex hTS-raltitrexed. The lack or certain alteration of the C-terminal residues (residues 308-310 in particular) yielded melting profiles with less apparent biphasic feature. Meanwhile, the structural data showed that these three C-terminal residues were in close contact with a loop segment 211-218 and could thus play a role in stabilizing this loop and make contributions to the overall stability of protein. Taken together, it is possible that without these residues at the C-terminus, the gap of melting curves between unliganded and liganded protein population is smaller, resulting in vague phase separation after summing up the curves (Figure 4.1b).



**Figure 4.1. Model 1 of the biphasic melting profile of hTS on DSF.** a) Assume there are two different protein populations with a low (population 1, red) and high (population 2, black) melting temperature, the observed melting curve, which is the sum of the two melting curves, is a biphasic curve (blue). b) If the thermal stability difference is smaller due to the lack of C-terminal residues, the resulting observed curve has a less apparent biphasic feature.

A second possible model for the biphasic profile is that the first phase with lower melting temperature is the unfolding of certain protein region with low thermal stability and the second is the global unfolding of protein. In this model, it is possible that the first phase is the dissociation of hTS dimer into monomer. Binding of dUMP or phosphate ion enhances hTS dimerization and thus stabilizes the first phase, which is demonstrated by the electrostatic interactions between phosphate group and the four arginines contributed by each subunit in the crystal structure. Binding of antifolate together with dUMP increases the global thermal stability of hTS monomer. It is also possible that the first phase constitutes to the local unfolding of the active site region.

Because the C-terminal loop, serving as the lid of the active site, is very flexible manifested in the crystal structures, the active site could be of low thermal stability in the unliganded state. Similar to the first model, one possible explanation for profile difference resulting from C-terminal variations is that lacking of the intact C-terminal residues hTS monomer is less stable and starts unfolding globally earlier, leading to the loss of apparent biphasic feature.

At present we cannot distinguish the different possibilities for the biphasic melting profile. The profile could also be a result of the combination of multiple factors. Despite the lack of clear definition for the two phases, one conclusion we can deduce from the proposed models is that the first melting phase reports more on the binding events in the dUMP binding site while the second melting phase reports more on the folate binding pocket and the global thermal stability of the protein.

#### **4.2 Implications of hTS structures**

In the four crystal forms obtained for hTS, the asymmetric unit contains four or six monomers. There are small differences between the monomers in the asymmetric unit, mainly at the C-terminus. It is possible that the differences reflect a mechanistic non-equivalence of the two active sites in TS as previously proposed (Anderson et al., 1999; Danenberg and Danenberg, 1979; Dev et al., 1994; Lovelace et al., 2007), but the difference could also be attributed to the non-equivalent crystal contacts. Further studies are needed to establish whether a subunit non-equivalency is present when protein is in solution.

A surprising finding from the structures is that the seven C-terminal residues could bind in the active cavity of hTS across both dUMP and folate binding pocket. Interestingly, the binding of these residues in the active site triggers extensive conformational changes and results in a protein conformation similar to the covalent ternary complex. This observation could also be an artifact due to the crystal packing forces and further studies are required to examine whether this conformation is present in solution. We have however tested heptapeptide and tetrapeptide of the C-terminal residues against hTS using DSF, but no significant thermal stability shift is observed (data not shown).

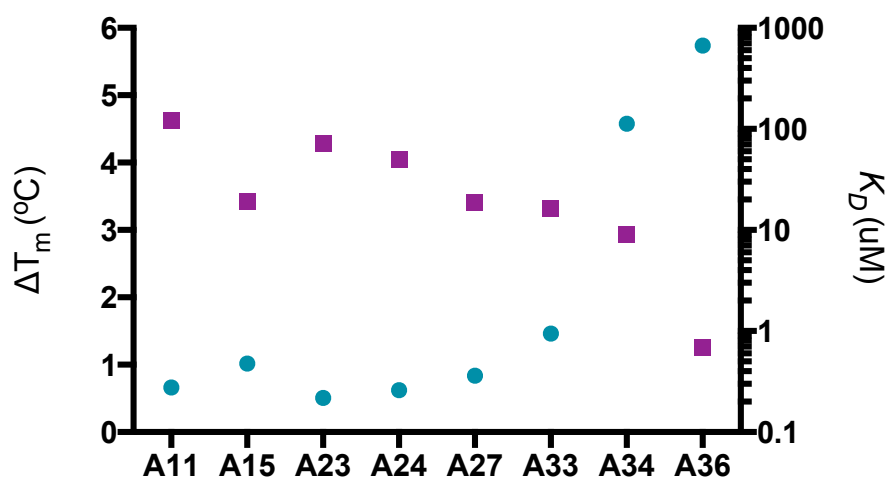
The conformational changes of TS induced by folate/antifolate binding have been well described since the first structure of ternary complexes of *E. coli* TS was determined (Matthews et al., 1990). However, there has been a lack of information on the conformational changes of hTS protein. In the current study, we have established several structures of the functional complexes of hTS and also compared the crystal structures of unliganded hTS with binary and ternary complexes of hTS using one crystal form from soaking approach. Overall, the conformational changes taking place in hTS are similar to those in observed in *E. coli* TS: binding of dUMP leads to a more ordered structure but does not trigger significant protein rearrangement; binding of antifolates induces extensive conformational changes involving mainly the C-terminal loop and protein segments lining the active site moving to close the cavity. Compared with *E. coli* TS, there are two insert regions in hTS, which correspond to the segments AS- $\alpha$ 3 and DL loop (Figure 3.7). Both segments undergo positional shifts with a vector toward the active site upon the binding of raltitrexed and nolatrexed. Of note is that these two insert regions are also key players in the conformational switching between the active and inactive states of hTS. In contrast to the closure of the active site in the ternary complex, they open upon inactivation. It is likely that the insert regions in hTS function to increase conformational flexibility and cooperativity of the enzyme and thus facilitate to overcome the energy barrier associated with conformational changes.

In the structure of hTS-P<sub>i</sub>, the catalytic cavity is featured with a large number of ordered water molecules. Some of these water molecules have been proposed to be involved in the enzyme's catalytic reaction. Aside from this, these water molecules in the active site probably help in accommodating the conformational changes upon ternary complex formation as well as stabilizing the ligand-free state of hTS. The energy costs associated with the conformational shifts could be offset by the entropy gain in releasing the water molecules from the cavity.

### **4.3 Fragment screening and hit translation into crystal structure**

In this work, we have applied a fragment-based strategy to discover novel compounds as compared to known hTS inhibitors. Currently all hTS inhibitors in clinical use are prone to resistance development and significant toxicity (Ackland et

al., 2006; Jarmula, 2010; Walling, 2006). The fragment-based strategy was driven mainly by DSF and guided by crystal structures of protein-ligand complexes. During chemical evolution, ITC was used to determine binding affinity and characterize the thermodynamic profile of ligands. We find that DSF generally produces reliable data for hTS system. Despite inevitable false negatives, DSF-based fragment screening appears to be efficient and leads to identification of four types of initial fragment ligands and even ligands in a novel pocket of hTS. DSF was also used as a primary technique to evaluate optimized compounds and found to be in good correlation with ITC affinity data (Figure 4.2). Orthogonal evaluation using enzymatic assay demonstrated that the ranking of binding affinity data from biophysical techniques correlated well with the potency ranking in inhibiting hTS activity (Figure 3.18).



**Figure 4.2. Good correlation observed between DFS data and  $K_D$  values from ITC.** The  $T_m$  shift data (●,  $\Delta T_m$  at 125  $\mu M$  in the presence of 500  $\mu M$  dUMP) correlate well with ITC affinity data (■,  $K_D$  values) for compounds from the same series (type I compounds here).

In practice, one of the main challenges of FBDD is to translate hits from screening assays into crystal structures. To achieve this, a robust crystallographic system is required: the crystals have to be reproducible and well diffracting; the crystal packing should allow ready access to the binding pockets; and the crystals have to survive soaking with high concentration of compounds as well as DMSO. Fortunately, we managed to obtain such a crystal system and successfully translate some of our screening hits into the crystal structures. However, there are still hit compounds that do not translate well into the crystals, which could be due to the

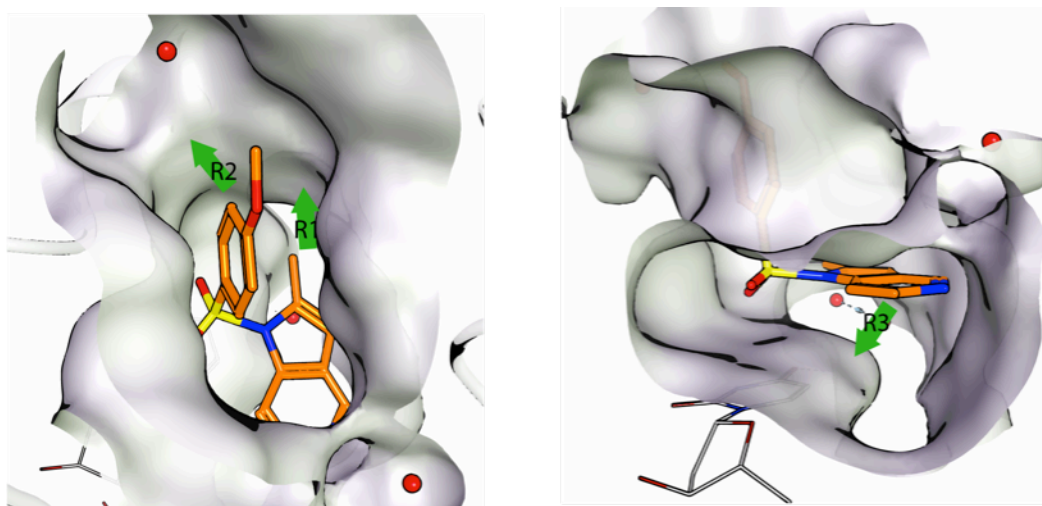
crystal system being less flexible than in solution and partly locked into a particular conformation. In the case of hTS, considering that we have demonstrated the existence of at least three different conformations (one active and two inactive), there is a higher possibility of problems in translating hTS screening hits using one crystal form. On the basis of this, one possible reason for failing to translate certain screening hits (phosphate hits in particular) is that those hits bind to a different hTS conformer from the active conformer in the crystal.

#### **4.4 Novel starting points for hTS inhibitor development**

From our efforts in FBDD, we have identified and translated four types of fragment hits of different binding modes into hTS crystal structures, of which two (type I and II) are novel as compared to existing hTS inhibitors and one is unique in its binding site (type IV). The fragment strategy proves to be efficient and successful in identifying novel start points for hTS inhibitor development. Potentially these novel start points offer an opportunity to circumvent the drug resistance mechanisms via rational drug design. Non-folate like inhibitors will not depend on active transport and polyglutamation and resistance mechanisms in systems affecting these processes should not yield resistance to non-folate like inhibitors.

As described above, we decided to follow up type I hits in folate binding site for fragment evolution because of its novel binding mode in sampling the back pocket of the active site. In fact, during the chemical elaboration of this series of compounds we observed several additional interesting properties that differentiate this series from known hTS inhibitors. One distinguishing feature is that their binding does not trigger significant difference in the protein segments including the most vulnerable C-terminal loop. In the course of fragment evolution of type I hits, the protein conformation is conserved and virtually identical to that of hTS-P<sub>i</sub> and hTS-dUMP (except for **A37**), suggesting that binding mode of type I hits limits the loss of binding energy resulting from the protein conformational entropic costs. Different from previous antifolates such as raltitrexed and nolatrexed, the elaborated compounds of this series are in close contact with the residues in the back pocket and AS- $\alpha$ 2 region instead of the C-terminal residues. Importantly, while the protein segments remain in the initial low-energy conformation, the residues present an optimal orientation of the

functional groups to interact with the compounds. Taking compound **A36** as an example, it fits nicely into the geometric arrangement of the initial unbound state of hTS and forms edge-to-face interactions with Trp109 as well as hydrogen bond with Asn112 (Figure 3.17). **A36**, the most potent compound of this series obtained at this point ( $K_D = 0.69 \mu\text{M}$ ), represents a novel scaffold for drug design targeting hTS. It possesses three potential expansion vectors to further explore the binding pocket and to gain potency and desired drug-like properties (Figure 4.3). Based on experiences in our previous FBDD projects (Larsson et al., 2013), the initial improvements of binding from fragment expansion is the most changing part of the process while affinity down to low nM can be relatively rapid when a sub-nM compound has been established.



**Figure 4.3. Potential vectors of expansion on A36.** Three potential vectors (R1, R2 and R3) can be used for further chemical evolution on compound **A36**. R1 points into the back pocket near Glu87; R2 vectors out to the entrance channel of active site; and R3 directs to the “hot spot” region above dUMP.

The compounds of type II hits in folate binding site are featured with the imidazoline ring connected to a sulfur, which is an efficient binder with a significantly smaller framework and also captures the key interactions with C-terminal residues and Asp218 observed in raltitrexed. Elaboration work on the imidazoline ring has demonstrated that it is the unique electrostatic nature that drives the binding interactions. This binding mode was discovered late during my thesis work and unfortunately, the chemical evolution work on the phenyl scaffold in this series has been hampered by chemical synthesis difficulties and therefore is not so

elaborated at present. However, we have shown that the compounds have a vector ready for affinity development by expanding into the entrance channel of the active site. Even though this work is only in the initial stage, type II hits also hold promise for fragment merging with type I hits. The successful compounds could bind with entropic advantages owing to limited loss of protein conformational entropic costs as well as increased enthalpy gain from the interactions with C-terminal residues.

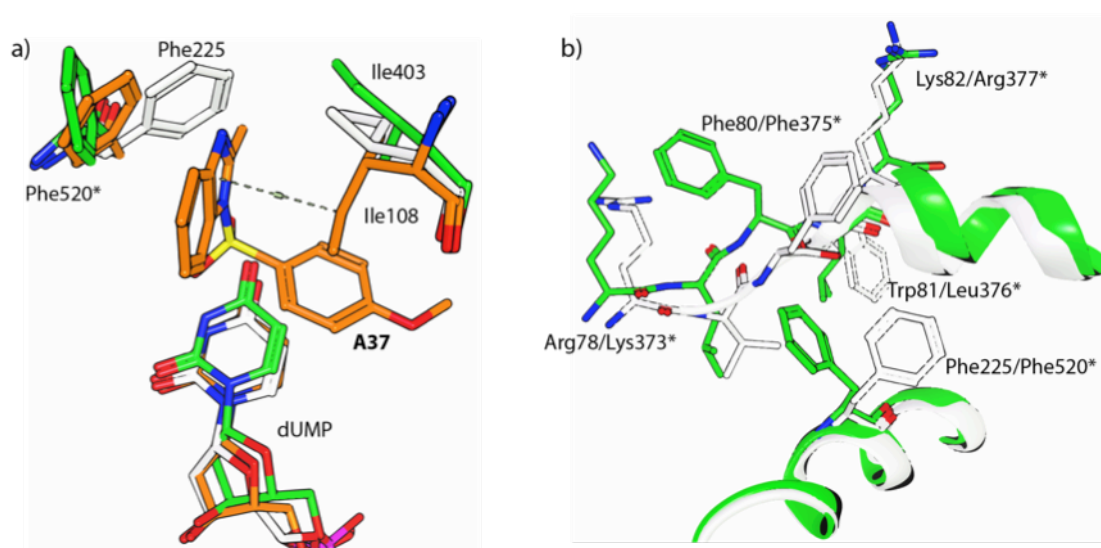
#### **4.5 Fragment merging strategy**

Fragment merging is considered as an elegant fragment evolution strategy as the binding geometry and synthetic route of the core scaffold are usually maintained and the second-site sub-structural component is attached as a new substituent. However, the strategy is difficult in practice due to the significant limitations in maintaining conformational freedom in the merged molecules. In this work, we have attempted twice to apply fragment merging for the elaboration of initial fragment hits. The first attempt was made with **A33** to combine **A15** and **A27**, in which the atomic overlap of the initial fragments was not great enough and both had to sacrifice certain binding interactions, and ultimately failed. The failure of fragment merging is not unexpected in this case, as theoretically the larger the atomic overlap, the higher probability for the elaborated fragments to recapitulate the binding pose and interactions of the initial fragments. However, apart from the atomic overlap, substantial attention should be paid to the conformational preferences of the merged compounds. This is the lesson we learn in our second attempt of fragment merging for type II hits. The parental fragment of type II hits has great geometric overlap with type I hits, while the two combined compounds (**B9** and **B10**) fail to recapitulate the binding pose of their parents. The failure is possibly attributed to the internal strain that would constitute a large conformational energy barrier to adopt the parents' binding pose.

#### **4.6 Implications of the binding mode of A37**

In the chemical elaboration process on type I hits, compound **A37** discloses a pocket that is not present in other hTS structures. However, this pocket is partially visible, if not apparent, in the structure of *P. falciparum* DHFR-TS, but not explored

before. As shown in Figure 4.4a, Phe520 (corresponding to Phe225 in hTS) in the structure of *P. falciparum* DHFR-TS-dUMP is in such a different rotameric state that it keeps the gate to the pocket open in contrast to what is seen in hTS. A more detailed structural analysis reveals that the gatekeeper phenylalanine residue is conserved among almost all TS enzymes including *P. falciparum* TS domain, whereas the residues that stabilize the conformation of this phenylalanine are variable between *P. falciparum* TS domain and other TS enzymes. It appears that the open conformation of Phe520 is stabilized owing to the shift of Phe375 above, which in turn probably results from the sequence difference of its surrounding residues (Figure 4.4b). The subtle differences resulting from differential stabilization of the rotameric states of the gatekeeper phenylalanine opens new avenues for selective inhibitor design against *P. falciparum* DHFR-TS. **A37** could serve as a starting framework for the selective inhibitor design.



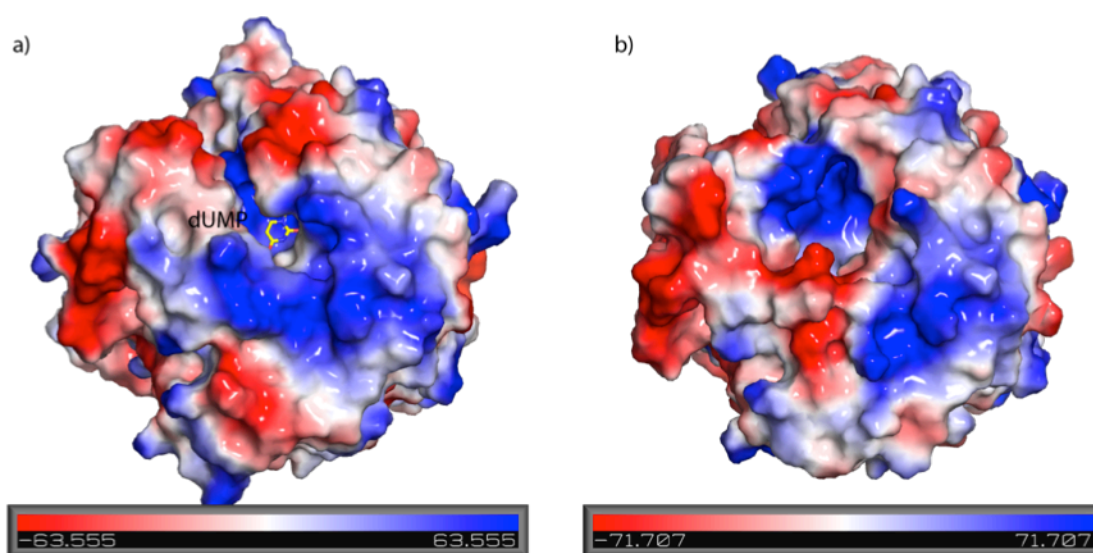
**Figure 4.4. Structural differences between hTS and *P. falciparum* DHFR-TS.** a) Structural alignment of *P. falciparum* DHFR-TS (PDB: 3QGT; green) with hTS-dUMP (white) and hTS-dUMP-A37 (orange). Different from hTS, the residue Phe225 is positioned to open the pocket for **A37** in *P. falciparum* DHFR-TS (\* represents corresponding residues number in *P. falciparum* DHFR-TS). b) The variable residues between hTS and *P. falciparum* DHFR-TS. Phe80 (Phe375\*) also sits in a different position in the two proteins, which is probably due to the variable residues surrounding.

#### 4.7 Insights into the transition between active and inactive states of hTS

In this work, we have investigated the binding of dUMP and raltitrexed to two inactive hTS mutants (M190K and A191K) using DSF and proved that the inactive conformers could be reversed to the active state by active-site ligands binding. Together with the previous studies that crystallize wild-type hTS in an inactive conformation (Phan et al., 2001), the data demonstrate that hTS exists in an equilibrium of active and inactive conformations. The crystal structures obtained in this study provide insights into the mechanism of hTS activation and inactivation. The three different inactive states shown in Figure 3.20 possibly represent snapshots of the conformation transition trajectory: the inactive form obtained from wild-type hTS as the less ordered intermediate state I; A191K as a more ordered intermediate state II; and M190K as the ultimate stable inactive state. This is supported by our DSF data that A191K can be rescued to the active state with a significant lower concentration of dUMP than M190K. However, it is not unlikely that the inactive conformation observed in M190K is not present in nature but a result of mutation artifact as the conformation of AS loop that contains the mutation site is very different from that in the inactive conformer of wild-type hTS.

In the inactive conformer of A191K-sf, there are two phosphate ions bound in the vicinity of the active site when no additional phosphate ions are added. However, in the active conformation, hTS binds only one phosphate ion in a position corresponding to the phosphate moiety of dUMP at 100 mM phosphate concentration (shown in the structure of hTS vc1 and hTS vc5 crystal form 3). It indicates that the conformational switching from the active to inactive state changes the electrostatic properties of the protein. Large areas of positively charged surface are often indicative of RNA binding site. Interestingly, the molecular electrostatic potential shows that the folate binding region exhibits the largest area of positively charged surface in both the active and inactive conformers (Figure 4.5). It is possible that this positively charged region serves as the folate binding site in the active state of hTS while it is involved in the binding of mRNA in the inactive state of hTS. Moreover, the protein segment (Lys107-Ser114) in the folate binding region is ordered in the active conformation but disordered in the inactive conformers. The disorder/mobility of this segment might correlate with mRNA binding and could be the recognition site of mRNA.

As the sole enzyme involved in *de novo* synthesis of dTMP, hTS is supposed to be tightly regulated to meet the metabolic demands and also to maintain the appropriate nucleotides balance. However, unlike other enzymes involved in nucleotide synthesis that are typically regulated via allosteric interactions and feedback inhibition, regulation of hTS activity has been only reported on its transcription and the detailed mechanism for this regulation remains unclear. In this study, a potential allosteric site at the dimer interface has been identified, which might regulate the transition between the active and inactive states of hTS. Integrating the experimental data, a regulatory mechanism of hTS activity is proposed: in the resting state of cells, hTS populates the inactive conformation due to an allosteric effector in the dimer interface cavity and binds to its own mRNA thus repressing hTS protein synthesis; when cells start to divide, another allosteric effector binds to the dimer interface to activate hTS and release it from mRNA, leading to an increase in both hTS protein level and enzyme activity to meet the cellular demands. Even though our screening of a metabolic library did not identify any effectors (data not shown), it is possible that the natural effectors are linked to the cell cycle and synthesis of other nucleotides so that the appropriate balance of the four dNTPs required for DNA synthesis is maintained.



**Figure 4.5. Electrostatic potential surface of the active and inactive states of hTS.** a) The electrostatic potential surface of the structure of hTS-dUMP (active state). b) The electrostatic potential surface of the structure of A191K-sf (inactive state). Large areas of positively charged surface are observed only in the folate binding region in both structures.

#### **4.8 Potential allosteric inhibitors for hTS**

Here, we have discovered a potential allosteric site in regulating the transition between the active and inactive states of hTS, which might in turn regulate the expression level and activity of the enzyme. Taken together, the stabilization of the inactive conformation of hTS might represent a novel strategy to surmount drug resistance associated with elevated target protein expression as well as for enzyme inhibition. The crystal forms for the two mutants (M190K and A191K) reported here enable crystallography studies of inhibitor binding in the inactive conformers of hTS. The two mutants exhibit different properties in that A191K confronts lower energy barrier to switch between the active and inactive states whereas M190K does not equilibrate with the active conformer easily. M190K could be a better target for identifying ligands stabilizing the inactive conformation. In this study, phage display has been applied to M190K in an attempt to obtain conformation-selective binders of hTS. One VH binder showed significant selectivity against M190K compared with wild-type hTS. It would be interesting to know where the VH binder binds in the inactive conformer. If the VH binder stabilizes the inactive state of hTS while does not interfere with its mRNA binding activity, the binder could be an important tool for studying the biological relevance of the inactive conformer and possibly also of therapeutic value to overcome drug resistance associated with hTS overexpression.

## CHAPTER 5 CONCLUSION AND FUTURE PERSPECTIVE

In summary, biophysical studies and fragment-based drug discovery have been performed on hTS using a combination of biophysical techniques, including X-ray crystallography, thermal shift assays (DSF and DSLS) and ITC. Comprehensive structural studies of hTS with the active-site ligands, dUMP and antifolates (methotrexate, raltitrexed and nolatrexed), reveal that binding of substrate dUMP stabilizes the overall structure of hTS though it does not trigger significant conformational changes. Binding of antifolates in the folate binding site induces extensive protein segmental rearrangements, mainly involving the C-terminal loop, the DL loop and the helices (AS- $\alpha$ 1, 2, 3 and 4) lining the active site of hTS. In this work, DSF proves to be a considerably reliable and effective probing technique in the system of hTS and is thus used as the primary method for protein characterization and binding studies. Fragment-based drug discovery strategy is also applied to hTS driven by DSF-based screening. In total, eight fragments of four types of binding modes have been successfully translated into the crystal structures from fragment screening: 1) two L-shaped type I hits binding in the back pocket and the entrance channel of the folate binding site; 2) one type II hit sitting above the ribose ring of dUMP but also overlapping type I hits in the entrance channel; 3) three fragments in type III hits binding in a planar conformation face to face with the pyrimidine ring of dUMP; and 4) two fragments of type IV binding mode locating in a novel pocket in the dimer interface of hTS.

Chemical elaboration guided by crystal structures has been applied to type I and II hits because of the novelty value of their binding modes as compared to existing hTS inhibitors. During the course of fragment elaboration on type I hits, the back pocket of the active site, which has not been explored in known hTS inhibitors, is found to be a potential site to exploit for improved compound activity. Additionally, compound from this series proved to be cell-active in targeting hTS using CETSA. In the final rounds of chemical expansion, one compound with a  $K_D$  of 0.69  $\mu$ M (**A36**) is developed and represents a novel starting point for hTS inhibitor design. Another compound, **A37** discloses a unique pocket by displacing two residues at the binding site. As this pocket constitutes a subtle difference between hTS and *P. falciparum* DHFR-TS, it can be explored for selective inhibitor design against *P.*

*falciparum* DHFR-TS and **A37** serves as a starting framework. Chemical expansion on type II hits has demonstrated effective affinity gain by vectoring out to the entrance channel. Although the strategy is not successful with elaborated compounds at present, fragment merging of type I and type II hits opens new avenues for the development of a novel class of hTS inhibitors. Both type I and type II compounds hold promises for overcoming the drug resistance mechanisms associated with antifolates metabolism via drug transportation and polyglutamation.

In an attempt to understand another drug resistance mechanism associated with the overexpression of hTS enzyme, further studies also focus attention on the inactive conformation of hTS using two mutants M190K and A191K. The crystal structures of the mutants reveal two different inactive conformers of hTS from the inactive conformer observed in wild-type hTS, which possibly represent the different states in the conformational switching between the active and inactive hTS. The two variable insert regions of TS enzymes appear to play an important role in the conformational transition upon activation/inactivation of hTS. Furthermore, the experimental data suggest that the dimer interface cavity could be an allosteric site that regulates the transition between the active and inactive states of hTS. Also, the folate binding region might serve as the mRNA binding site in the inactive states of hTS. Phage display has been performed to identify conformation-selective VH binders for hTS. One VH binder shows selectivity against the inactive conformation in the interaction assays. Further studies are required to understand the molecular interaction between the VH binder and hTS.

Several future directions can be explored for the purpose of drug development as well as understanding and thus overcoming drug resistance:

- Compound **A36** in type I hits possesses three potential vectors for further chemical elaboration. Future attention can be paid to these vectors to improve compound activity and gain drug-like properties as a novel class of hTS inhibitors.
- Further studies can focus on the fragment merging strategy to recapitulate the binding pose of type I and II hits as well. A combination of crystal structure analysis and in silico predictions of

conformational preference could be used as a better guide to overcome the conformational limitations of fragment merging.

- Compound **A37** uncovers a pocket that has been neglected in the development of selective inhibitors against *P. falciparum* DHFR-TS. It is worthwhile to examine this compound on *P. falciparum* DHFR-TS for binding and selectivity activity. If the results are promising, **A37** can be worked as a starting point to develop potent and selective inhibitors against *P. falciparum* DHFR-TS, which might be of therapeutic value for the treatment of malaria.
- On the basis of current experimental data, the inactive conformation of hTS might function as an mRNA binding protein and regulates its own translation level. Further studies, particularly *in vivo* studies, are required to investigate this biological relevance and validate the drug target value of the inactive conformer of hTS. Also, studies are needed to confirm the direct interaction between mRNA and the inactive conformer of hTS and to define the mRNA molecular recognition site in hTS.
- In this work, an allosteric site in the dimer interface is proposed to regulate the conformational switching between the active and inactive states of hTS. It will be important to identify the natural allosteric effectors if the regulatory role of the inactive hTS conformer is confirmed. In addition, this site could be used for allosteric inhibitors design.

## REFERENCES

- Ackland, S.P., Clarke, S.J., Beale, P., Peters, G.J., 2006. Thymidylate synthase inhibitors. *Update on Cancer Therapeutics* 1, 403-427.
- Anderson, A.C., O'Neil, R.H., DeLano, W.L., Stroud, R.M., 1999. The structural mechanism for half-the-sites reactivity in an enzyme, thymidylate synthase, involves a relay of changes between subunits. *Biochemistry* 38, 13829-13836.
- Arner, E.S., Eriksson, S., 1995. Mammalian deoxyribonucleoside kinases. *Pharmacology & Therapeutics* 67, 155-186.
- Assaraf, Y.G., 2006. The role of multidrug resistance efflux transporters in antifolate resistance and folate homeostasis. *Drug Resistance Updates : reviews and commentaries in antimicrobial and anticancer chemotherapy* 9, 227-246.
- Baddiley, J., Hughes, N.A., 1960. The synthesis of nucleotide coenzymes, advances in enzymology and related areas of molecular biology. John Wiley & Sons, Inc., pp. 157-203.
- Baker, M., 2013. Fragment-based lead discovery grows up. *Nature Reviews Drug Discovery* 12, 5-7.
- Barbour, K.W., Berger, S.H., Berger, F.G., 1990. Single amino acid substitution defines a naturally occurring genetic variant of human thymidylate synthase. *Molecular Pharmacology* 37, 515-518.
- Blakley, R.L., Vitols, E., 1968. The control of nucleotide biosynthesis. *Annual Review of Biochemistry* 37, 201-224.
- Bohacek, R.S., McMartin, C., Guida, W.C., 1996. The art and practice of structure-based drug design: A molecular modeling perspective. *Medicinal Research Reviews* 16, 3-50.
- Brunn, N.D., Dibrov, S.M., Kao, M.B., Ghassemian, M., Hermann, T., 2014. Analysis of mRNA recognition by human thymidylate synthase. *Bioscience Reports* 34, e00168.
- Canman, C.E., Lawrence, T.S., Shewach, D.S., Tang, H.-Y., Maybaum, J., 1993. Resistance to fluorodeoxyuridine-induced DNA damage and cytotoxicity correlates with an elevation of deoxyuridine triphosphatase activity and failure to accumulate deoxyuridine triphosphate. *Cancer Research* 53, 5219-5224.
- Cardinale, D., Guaitoli, G., Tondi, D., Luciani, R., Henrich, S., Salo-Ahen, O.M., Ferrari, S., Marverti, G., Guerrieri, D., Ligabue, A., Frassinetti, C., Pozzi, C., Mangani, S., Fessas, D., Guerrini, R., Ponterini, G., Wade, R.C., Costi, M.P., 2011. Protein-protein interface-binding peptides inhibit the cancer therapy target human thymidylate synthase. *Proceedings of the National Academy of Sciences of the United States of America* 108, E542-549.

- Carreras, C.W., Santi, D.V., 1995. The catalytic mechanism and structure of thymidylate synthase. *Annual Review of Biochemistry* 64, 721-762.
- Castiglioni, A., Krumbhaar, E.B., Alfred, A., 1947. *A history of medicine*. Knopf New York.
- Chabes, A., Thelander, L., 2000. Controlled protein degradation regulates ribonucleotide reductase activity in proliferating mammalian cells during the normal cell cycle and in response to DNA damage and replication blocks. *Journal of Biological Chemistry* 275, 17747-17753.
- Chabner, B.A., Roberts, T.G., Jr., 2005. Timeline: Chemotherapy and the war on cancer. *Nature Reviews Cancer* 5, 65-72.
- Chen, D., Frezza, M., Schmitt, S., Kanwar, J., Dou, Q.P., 2011. Bortezomib as the first proteasome inhibitor anticancer drug: current status and future perspectives. *Current Cancer Drug Targets* 11, 239-253.
- Cheng, N., Payne, R.C., Traut, T.W., 1986. Regulation of uridine kinase. Evidence for a regulatory site. *The Journal of Biological Chemistry* 261, 13006-13012.
- Christopherson, R.I., Lyons, S.D., Wilson, P.K., 2002. Inhibitors of de novo nucleotide biosynthesis as drugs. *Accounts of Chemical Research* 35, 961-971.
- Chu, E., Koeller, D.M., Casey, J.L., Drake, J.C., Chabner, B.A., Elwood, P.C., Zinn, S., Allegra, C.J., 1991. Autoregulation of human thymidylate synthase messenger RNA translation by thymidylate synthase. *Proceedings of the National Academy of Sciences of the United States of America* 88, 8977-8981.
- Cohen, M.H., Cortazar, P., Justice, R., Pazdur, R., 2010. Approval summary: pemetrexed maintenance therapy of advanced/metastatic nonsquamous, non-small cell lung cancer (NSCLC). *Oncologist* 15, 1352-1358.
- Connolly, G.P., Duley, J.A., 1999. Uridine and its nucleotides: biological actions, therapeutic potentials. *Trends in Pharmacological Sciences* 20, 218-225.
- Conroy, T., Desseigne, F., Ychou, M., Bouche, O., Guimbaud, R., Becouarn, Y., Adenis, A., Raoul, J.L., Gourgou-Bourgade, S., de la Fouchardiere, C., Bennouna, J., Bacht, J.B., Khemissa-Akouz, F., Pere-Verge, D., Delbaldo, C., Assenat, E., Chauffert, B., Michel, P., Montoto-Grillot, C., Ducreux, M., 2011. FOLFIRINOX versus gemcitabine for metastatic pancreatic cancer. *The New England journal of medicine* 364, 1817-1825.
- Cosgrove, M., 1998. Nucleotides. *Nutrition* 14, 748-751.
- Curtin, N.J., Hughes, A.N., 2001. Pemetrexed disodium, a novel antifolate with multiple targets. *The Lancet Oncology* 2, 298-306.

- Danenberg, K.D., Danenberg, P.V., 1979. Evidence for a sequential interaction of the subunits of thymidylate synthetase. *The Journal of Biological Chemistry* 254, 4345-4348.
- Davis, S.T., Berger, S.H., 1993. Variation in human thymidylate synthase is associated with resistance to 5-fluoro-2'-deoxyuridine. *Molecular Pharmacology* 43, 702-708.
- De Souza, R., Zahedi, P., Badame, R.M., Allen, C., Piquette-Miller, M., 2011. Chemotherapy dosing schedule influences drug resistance development in ovarian cancer. *Molecular Cancer Therapeutics* 10, 1289-1299.
- Delaney, G., Jacob, S., Featherstone, C., Barton, M., 2005. The role of radiotherapy in cancer treatment. *Cancer* 104, 1129-1137.
- Dev, I.K., Dallas, W.S., Ferone, R., Hanlon, M., McKee, D.D., Yates, B.B., 1994. Mode of binding of folate analogs to thymidylate synthase. Evidence for two asymmetric but interactive substrate binding sites. *The Journal of Biological Chemistry* 269, 1873-1882.
- DeVita, V.T., Chu, E., 2008. A history of cancer chemotherapy. *Cancer Research* 68, 8643-8653.
- Di Tommaso, P., Moretti, S., Xenarios, I., Orobittg, M., Montanyola, A., Chang, J.M., Taly, J.F., Notredame, C., 2011. T-Coffee: a web server for the multiple sequence alignment of protein and RNA sequences using structural information and homology extension. *Nucleic Acids Research* 39, W13-17.
- Easterby, J.S., 1989. The analysis of metabolite channelling in multienzyme complexes and multifunctional proteins. *The Biochemical Journal* 264, 605-607.
- Elcock, A.H., Potter, M.J., Matthews, D.A., Knighton, D.R., McCammon, J.A., 1996. Electrostatic channeling in the bifunctional enzyme dihydrofolate reductase-thymidylate synthase. *Journal of Molecular Biology* 262, 370-374.
- Emsley, P., Cowtan, K., 2004. Coot: model-building tools for molecular graphics. *Acta crystallographica. Section D, Biological Crystallography* 60, 2126-2132.
- Emura, T., Suzuki, N., Yamaguchi, M., Ohshimo, H., Fukushima, M., 2004. A novel combination antimetabolite, TAS-102, exhibits antitumor activity in FU-resistant human cancer cells through a mechanism involving FTD incorporation in DNA. *International Journal of Oncology* 25, 571-578.
- Ericsson, U.B., Hallberg, B.M., DeTitta, G.T., Dekker, N., Nordlund, P., 2006. Thermofluor-based high-throughput stability optimization of proteins for structural studies. *Analytical Biochemistry* 357, 289-298.
- Farber, S., Diamond, L.K., Mercer, R.D., Sylvester Jr, R.F., Wolff, J.A., 1948. Temporary remissions in acute leukemia in children produced by folic acid antagonist, 4-aminopteroyl-glutamic acid (aminopterin). *New England Journal of Medicine* 238, 787-793.

- Fauman, E.B., Rutenber, E.E., Maley, G.F., Maley, F., Stroud, R.M., 1994. Water-mediated substrate/product discrimination: The product complex of thymidylate synthase at 1.83 Å. *Biochemistry* 33, 1502-1511.
- Ferlay, J., Soerjomataram, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., Parkin, D.M., Forman, D., Bray, F., 2015. Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. *International Journal of Cancer* 136, E359-E386.
- Finer-Moore, J., Fauman, E.B., Foster, P.G., Perry, K.M., Santi, D.V., Stroud, R.M., 1993. Refined structures of substrate-bound and phosphate-bound thymidylate synthase from *Lactobacillus casei*. *Journal of Molecular Biology* 232, 1101-1116.
- Finer-Moore, J.S., Montfort, W.R., Stroud, R.M., 1990. Pairwise specificity and sequential binding in enzyme catalysis: thymidylate synthase. *Biochemistry* 29, 6977-6986.
- Finer-Moore, J.S., Santi, D.V., Stroud, R.M., 2003. Lessons and conclusions from dissecting the mechanism of a bisubstrate enzyme: thymidylate synthase mutagenesis, function, and structure. *Biochemistry* 42, 248-256.
- Fojo, T., Bates, S., 2003. Strategies for reversing drug resistance. *Oncogene* 22, 7512-7523.
- Fox, I.H., 1978. Degradation of purine nucleotides. Springer Berlin Heidelberg, pp. 93-124.
- Freireich, E.J., 1967. The management of acute leukemia. *Canadian Medical Association Journal* 96, 1605-1610.
- Fritz, T.A., Liu, L., Finer-Moore, J.S., Stroud, R.M., 2002. Tryptophan 80 and leucine 143 are critical for the hydride transfer step of thymidylate synthase by controlling active site access. *Biochemistry* 41, 7021-7029.
- Gattermann, N., Dadak, M., Hofhaus, G., Wulfert, M., Berneburg, M., Loeffler, M.L., Simmonds, H.A., 2004. Severe impairment of nucleotide synthesis through inhibition of mitochondrial respiration. *Nucleosides, Nucleotides & Nucleic Acids* 23, 1275-1279.
- Gilman, A., 1963. The initial clinical trial of nitrogen mustard. *The American Journal of Surgery* 105, 574-578.
- Gottesman, M.M., 2002. Mechanisms of cancer drug resistance. *Annual Review of Medicine* 53, 615-627.
- Goulian, M., Bleile, B.M., Dickey, L.M., Grafstrom, R.H., Ingraham, H.A., Neynaber, S.A., Peterson, M.S., Tseng, B.Y., 1986. Mechanism of thymineless death. *Advances in Experimental Medicine and Biology* 195 Pt B, 89-95.

- Graslund, S., Nordlund, P., Weigelt, J., Hallberg, B.M., Bray, J., Gileadi, O., Knapp, S., Oppermann, U., Arrowsmith, C., Hui, R., Ming, J., dhe-Paganon, S., Park, H.W., Savchenko, A., Yee, A., Edwards, A., Vincentelli, R., Cambillau, C., Kim, R., Kim, S.H., Rao, Z., Shi, Y., Terwilliger, T.C., Kim, C.Y., Hung, L.W., Waldo, G.S., Peleg, Y., Albeck, S., Unger, T., Dym, O., Prilusky, J., Sussman, J.L., Stevens, R.C., Lesley, S.A., Wilson, I.A., Joachimiak, A., Collart, F., Dementieva, I., Donnelly, M.I., Eschenfeldt, W.H., Kim, Y., Stols, L., Wu, R., Zhou, M., Burley, S.K., Emtage, J.S., Sauder, J.M., Thompson, D., Bain, K., Luz, J., Gheyi, T., Zhang, F., Atwell, S., Almo, S.C., Bonanno, J.B., Fiser, A., Swaminathan, S., Studier, F.W., Chance, M.R., Sali, A., Acton, T.B., Xiao, R., Zhao, L., Ma, L.C., Hunt, J.F., Tong, L., Cunningham, K., Inouye, M., Anderson, S., Janjua, H., Shastry, R., Ho, C.K., Wang, D., Wang, H., Jiang, M., Montelione, G.T., Stuart, D.I., Owens, R.J., Daenke, S., Schutz, A., Heinemann, U., Yokoyama, S., Bussow, K., Gunsalus, K.C., 2008. Protein production and purification. *Nature Methods* 5, 135-146.
- Graves, L.M., Guy, H.I., Kozlowski, P., Huang, M., Lazarowski, E., Pope, R.M., Collins, M.A., Dahlstrand, E.N., Earp, H.S., 3rd, Evans, D.R., 2000. Regulation of carbamoyl phosphate synthetase by MAP kinase. *Nature* 403, 328-332.
- Hajduk, P.J., Greer, J., 2007. A decade of fragment-based drug design: strategic advances and lessons learned. *Nature Reviews Drug Discovery* 6, 211-219.
- Hanahan, D., Weinberg, R.A., 2000. The hallmarks of cancer. *Cell* 100, 57-70.
- Hann, M.M., Leach, A.R., Harper, G., 2001. Molecular complexity and its impact on the probability of finding leads for drug discovery. *Journal of Chemical Information and Computer Sciences* 41, 856-864.
- Hardy, L.W., Graves, K.L., Nalivaika, E., 1995. Electrostatic guidance of catalysis by a conserved glutamic acid in *Escherichia coli* dTMP synthase and bacteriophage T4 dCMP hydroxymethylase. *Biochemistry* 34, 8422-8432.
- Hardy, L.W., Finer-Moore, J.S., Montfort, W.R., Jones, M.O., Santi, D.V., Stroud, R.M., 1987. Atomic structure of thymidylate synthase: target for rational drug design. *Science* 235, 448-455.
- Harrap, K.R., Jackman, A.L., Newell, D.R., Taylor, G.A., Hughes, L.R., Calvert, A.H., 1989. Thymidylate synthase: A target for anticancer drug design. *Advances in Enzyme Regulation* 29, 161-179.
- Hartman, S.C., 1963. Phosphoribosyl Pyrophosphate Amidotransferase: purification and general catalytic properties. *Journal of Biological Chemistry* 238, 3024-3035.
- Havla, J., Kumpfel, T., Hohlfeld, R., 2015. Immunotherapies for multiple sclerosis : review and update. *Der Internist* 56, 432-445.
- Hazarika, M., White, R.M., Johnson, J.R., Pazdur, R., 2004. FDA drug approval summaries: pemetrexed (Alimta). *Oncologist* 9, 482-488.

- Heidelberger, C., Chaudhuri, N.K., Danneberg, P., Mooren, D., Griesbach, L., Duschinsky, R., Schnitzer, R.J., Plevin, E., Scheiner, J., 1957. Fluorinated pyrimidines, a new class of tumour-inhibitory compounds. *Nature* 179, 663-666.
- Hennenfent, K.L., Govindan, R., 2006. Novel formulations of taxanes: a review. Old wine in a new bottle? *Annals of Oncology* 17, 735-749.
- Holohan, C., Van Schaeybroeck, S., Longley, D.B., Johnston, P.G., 2013. Cancer drug resistance: an evolving paradigm. *Nature Reviews Cancer* 13, 714-726.
- Houghton, J.A., Harwood, F.G., Tillman, D.M., 1997. Thymineless death in colon carcinoma cells is mediated via fas signaling. *Proceedings of the National Academy of Sciences of the United States of America* 94, 8144-8149.
- Hu, C.M., Zhang, L., 2012. Nanoparticle-based combination therapy toward overcoming drug resistance in cancer. *Biochemical Pharmacology* 83, 1104-1111.
- Huang, X., Gibson, L.M., Bell, B.J., Lovelace, L.L., Peña, M.M.O., Berger, F.G., Berger, S.H., Lebioda, L., 2010. Replacement of Val3 in human thymidylate synthase affects its kinetic properties and intracellular stability. *Biochemistry* 49, 2475-2482.
- Hyatt, D.C., Maley, F., Montfort, W.R., 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.
- Institute, N.C., 2009. Common Terminology Criteria for Adverse Events v4.0
- Ishikawa, T., Utoh, M., Sawada, N., Nishida, M., Fukase, Y., Sekiguchi, F., Ishitsuka, H., 1998. Tumor selective delivery of 5-fluorouracil by capecitabine, a new oral fluoropyrimidine carbamate, in human cancer xenografts. *Biochemical Pharmacology* 55, 1091-1097.
- Ivanetich, K.M., Santi, D.V., 1990. Bifunctional thymidylate synthase-dihydrofolate reductase in protozoa. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* 4, 1591-1597.
- Izquierdo, M., 2005. Short interfering RNAs as a tool for cancer gene therapy. *Cancer gene therapy* 12, 217-227.
- Jackman, A.L., Jodrell, D.I., Gibson, W., Stephens, T.C., 1991. ICI D1694, an inhibitor of thymidylate synthase for clinical study. *Advances in Experimental Medicine and Biology* 309a, 19-23.
- Jackman, A.L., Gibson, W., Brown, M., Kimbell, R., Boyle, F.T., 1993. The role of the reduced-folate carrier and metabolism to intracellular polyglutamates for the activity of ICI D1694. *Advances in Experimental Medicine and Biology* 339, 265-276.
- Jackson, R.C., Weber, G., Morris, H.P., 1975. IMP dehydrogenase, an enzyme linked with proliferation and malignancy. *Nature* 256, 331-333.

- Jarmula, A., 2010. Antifolate inhibitors of thymidylate synthase as anticancer drugs. *Mini Reviews in Medicinal Chemistry* 10, 1211-1222.
- Jencks, W.P., 1981. On the attribution and additivity of binding energies. *Proceedings of the National Academy of Sciences of the United States of America* 78, 4046-4050.
- Ju, J., Pedersen-Lane, J., Maley, F., Chu, E., 1999. Regulation of p53 expression by thymidylate synthase. *Proceedings of the National Academy of Sciences* 96, 3769-3774.
- June, C.H., 2007. Adoptive T cell therapy for cancer in the clinic. *Journal of Clinical Investigation* 117, 1466.
- Kamb, A., Finer-Moore, J.S., Stroud, R.M., 1992a. Cofactor triggers the conformational change in thymidylate synthase: implications for an ordered binding mechanism. *Biochemistry* 31, 12876-12884.
- Kamb, A., Finer-Moore, J., Calvert, A.H., Stroud, R.M., 1992b. Structural basis for recognition of polyglutamyl folates by thymidylate synthase. *Biochemistry* 31, 9883-9890.
- Kelley Bentley, J., Beavo, J.A., 1992. Regulation and function of cyclic nucleotides. *Current Opinion in Cell Biology* 4, 233-240.
- Kelley, R.E., Andersson, H.C., 2014. Chapter 55 - Disorders of purines and pyrimidines. In: José, B., José, M.F. (Eds.), *Handbook of Clinical Neurology*. Elsevier, pp. 827-838.
- Klebe, G., 2000. Recent developments in structure-based drug design. *Journal of Molecular Medicine* 78, 269-281.
- Knowles, J.R., 1980. Enzyme-catalyzed phosphoryl transfer reactions. *Annual Review of Biochemistry* 49, 877-919.
- Kobayashi, H., Takemura, Y., Ohnuma, T., 1998. Variable expression of RFC1 in human leukemia cell lines resistant to antifolates. *Cancer Letters* 124, 135-142.
- Kolberg, M., Strand, K.R., Graff, P., Andersson, K.K., 2004. Structure, function, and mechanism of ribonucleotide reductases. *Biochimica et biophysica acta* 1699, 1-34.
- Kufe, D.W., Major, P.P., 1981. 5-Fluorouracil incorporation into human breast carcinoma RNA correlates with cytotoxicity. *The Journal of Biological Chemistry* 256, 9802-9805.
- Landau, M.J., 2013. Structural and mechanistic studies of a bifunctional enzyme from pathogenic protozoa: probing interdomain communication and selective inhibition of *T. gondii* thymidylate synthase-dihydrofolate reductase. Yale University, Ann Arbor, pp. 146.

- Lane, A.N., Fan, T.W.-M., 2015. Regulation of mammalian nucleotide metabolism and biosynthesis. *Nucleic Acids Research*.
- Langer, G., Cohen, S.X., Lamzin, V.S., Perrakis, A., 2008. Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nature Protocols* 3, 1171-1179.
- Larsson, E.A., Jansson, A., Ng, F.M., Then, S.W., Panicker, R., Liu, B., Sangthongpitag, K., Pendharkar, V., Tai, S.J., Hill, J., Dan, C., Ho, S.Y., Cheong, W.W., Poulsen, A., Blanchard, S., Lin, G.R., Alam, J., Keller, T.H., Nordlund, P., 2013. Fragment-based ligand design of novel potent inhibitors of tankyrases. *Journal of Medicinal Chemistry* 56, 4497-4508.
- Li, W.W., Waltham, M., Tong, W., Schweitzer, B.I., Bertino, J.R., 1993. Increased activity of gamma-glutamyl hydrolase in human sarcoma cell lines: a novel mechanism of intrinsic resistance to methotrexate (MTX). *Advances in Experimental Medicine and Biology* 338, 635-638.
- Liani, E., Rothem, L., Bunni, M.A., Smith, C.A., Jansen, G., Assaraf, Y.G., 2003. Loss of folylpoly-gamma-glutamate synthetase activity is a dominant mechanism of resistance to polyglutamylation-dependent novel antifolates in multiple human leukemia sublines. *International Journal of Cancer* 103, 587-599.
- Lieberman, I., 1956. Enzymatic amination of uridine triphosphate to cytidine triphosphate. *The Journal of Biological Chemistry* 222, 765-775.
- Lim, S.M., Chen, D., Teo, H., Roos, A., Jansson, A.E., Nyman, T., Trésaugues, L., Pervushin, K., Nordlund, P., 2013. Structural and dynamic insights into substrate binding and catalysis of human lipocalin prostaglandin D synthase. *Journal of Lipid Research* 54, 1630-1643.
- Lin, X., Mizunuma, N., Chen, T., Copur, S.M., Maley, G.F., Liu, J., Maley, F., Chu, E., 2000. In vitro selection of an RNA sequence that interacts with high affinity with thymidylate synthase. *Nucleic Acids Research* 28, 4266-4274.
- Liu, Y., Barrett, J.E., Schultz, P.G., Santi, D.V., 1999. Tyrosine 146 of thymidylate synthase assists proton abstraction from the 5-position of 2'-deoxyuridine 5'-monophosphate. *Biochemistry* 38, 848-852.
- Liu, Y., Yin, T.J., Zhou, R., Zhou, S., Fan, L., Zhang, R.G., 2013. Expression of thymidylate synthase predicts clinical outcomes of pemetrexed-containing chemotherapy for non-small-cell lung cancer: a systemic review and meta-analysis. *Cancer Chemotherapy and Pharmacology* 72, 1125-1132.
- Löffler, M., Zameitat, E., 2013. Pyrimidine biosynthesis and degradation (catabolism). *Encyclopedia of Biological Chemistry*. Academic Press, Waltham, pp. 712-718.
- Löffler, M., Fairbanks, L.D., Zameitat, E., Marinaki, A.M., Simmonds, H.A., 2005. Pyrimidine pathways in health and disease. *Trends in Molecular Medicine* 11, 430-437.

- Longley, D.B., Boyer, J., Allen, W.L., Latif, T., Ferguson, P.R., Maxwell, P.J., McDermott, U., Lynch, M., Harkin, D.P., Johnston, P.G., 2002. The role of thymidylate synthase induction in modulating p53-regulated gene expression in response to 5-fluorouracil and antifolates. *Cancer Research* 62, 2644-2649.
- Lovelace, L.L., Minor, W., Lebioda, L., 2005. Structure of human thymidylate synthase under low-salt conditions. *Acta Crystallographica. Section D, Biological Crystallography* 61, 622-627.
- Lovelace, L.L., Gibson, L.M., Lebioda, L., 2007. Cooperative inhibition of human thymidylate synthase by mixtures of active site binding and allosteric inhibitors. *Biochemistry* 46, 2823-2830.
- Lovelace, L.L., Johnson, S.R., Gibson, L.M., Bell, B.J., Berger, S.H., Lebioda, L., 2009. Variants of human thymidylate synthase with loop 181-197 stabilized in the inactive conformation. *Protein Science* 18, 1628-1636.
- Makin, G., 2014. Principles of chemotherapy. *Paediatrics and Child Health* 24, 161-165.
- Maley, G.F., Maley, F., 1963. Feedback control of purified deoxycytidylate deaminase. *Science* 141, 1278-1279.
- Mamot, C., Drummond, D.C., Hong, K., Kirpotin, D.B., Park, J.W., 2003. Liposome-based approaches to overcome anticancer drug resistance. *Drug Resistance Updates* 6, 271-279.
- Martin, G.S., 2003. Cell signaling and cancer. *Cancer Cell* 4, 167-174.
- Matthews, D.A., Villafranca, J.E., Janson, C.A., Smith, W.W., Welsh, K., Freer, S., 1990. Stereochemical mechanism of action for thymidylate synthase based on the X-ray structure of the covalent inhibitory ternary complex with 5-fluoro-2'-deoxyuridylate and 5,10-methylenetetrahydrofolate. *Journal of Molecular Biology* 214, 937-948.
- Mayer, L.D., Janoff, A.S., 2007. Optimizing combination chemotherapy by controlling drug ratios. *Molecular Interventions* 7, 216-223.
- McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., Read, R.J., 2007. Phaser crystallographic software. *Journal of Applied Crystallography* 40, 658-674.
- McCredie, K.B., Ho, D.H.W., Freireich, E.J., 1973. L-asparaginase for the treatment of cancer. *Cancer Journal for Clinicians* 23, 220-227.
- McNeill, D.R., Lam, W., DeWeese, T.L., Cheng, Y.C., Wilson, D.M., 3rd, 2009. Impairment of APE1 function enhances cellular sensitivity to clinically relevant alkylators and antimetabolites. *Molecular Cancer Research* 7, 897-906.
- Mellor, H.R., Snelling, S., Hall, M.D., Modok, S., Jaffar, M., Hambley, T.W., Callaghan, R., 2005. The influence of tumour microenvironmental factors on the

efficacy of cisplatin and novel platinum(IV) complexes. *Biochemical Pharmacology* 70, 1137-1146.

- Miller, E., 1971. The metabolism and pharmacology of 5-fluorouracil. *Journal of Surgical Oncology* 3, 309-315.
- Moffatt, B.A., Ashihara, H., 2002. Purine and pyrimidine nucleotide synthesis and metabolism. *The Arabidopsis Book / American Society of Plant Biologists* 1, e0018.
- Molina, D.M., Jafari, R., Ignatushchenko, M., Seki, T., Larsson, E.A., Dan, C., Sreekumar, L., Cao, Y., Nordlund, P., 2013. Monitoring drug target engagement in cells and tissues using the cellular thermal shift assay. *Science* 341, 84-87.
- Montfort, W.R., Perry, K.M., Fauman, E.B., Finer-Moore, J.S., Maley, G.F., Hardy, L., Maley, F., Stroud, R.M., 1990. Structure, multiple site binding, and segmental accommodation in thymidylate synthase on binding dUMP and an anti-folate. *Biochemistry* 29, 6964-6977.
- Moriarty, N.W., Grosse-Kunstleve, R.W., Adams, P.D., 2009. Electronic ligand builder and optimization workbench (eLBOW): a tool for ligand coordinate and restraint generation. *Acta Crystallographica Section D* 65, 1074-1080.
- Moscow, J.A., Connolly, T., Myers, T.G., Cheng, C.C., Paull, K., Cowan, K.H., 1997. Reduced folate carrier gene (RFC1) expression and anti-folate resistance in transfected and non-selected cell lines. *International Journal of Cancer* 72, 184-190.
- Munch-Petersen, B., Cloos, L., Jensen, H.K., Tyrsted, G., 1995. Human thymidine kinase 1. Regulation in normal and malignant cells. *Advances in Enzyme Regulation* 35, 69-89.
- Munoz-Pinedo, C., El Mjiyad, N., Ricci, J.E., 2012. Cancer metabolism: current perspectives and future directions. *Cell Death and Disease* 3, e248.
- Murshudov, G.N., Skubák, P., Lebedev, A.A., Pannu, N.S., Steiner, R.A., Nicholls, R.A., Winn, M.D., Long, F., Vagin, A.A., 2011. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallographica Section D: Biological Crystallography* 67, 355-367.
- Noordhuis, P., Holwerda, U., Van der Wilt, C.L., Van Groeningen, C.J., Smid, K., Meijer, S., Pinedo, H.M., Peters, G.J., 2004. 5-Fluorouracil incorporation into RNA and DNA in relation to thymidylate synthase inhibition of human colorectal cancers. *Annals of Oncology* 15, 1025-1032.
- Noree, C., Monfort, E., Shiau, A.K., Wilhelm, J.E., 2014. Common regulatory control of CTP synthase enzyme activity and filament formation. *Molecular Biology of the Cell* 25, 2282-2290.
- Nyhan, W.L., 2001. Nucleotide synthesis via salvage pathway, eLS. John Wiley & Sons, Ltd.

- Otwinowski, Z., Minor, W., 1997. Processing of X-ray diffraction data collected in oscillation mode. Elsevier, pp. 307-326.
- Pardoll, D.M., 2012. The blockade of immune checkpoints in cancer immunotherapy. *Nature Reviews Cancer* 12, 252-264.
- Peña, Maria Marjorette O., Xing, Yang Y., Koli, S., Berger, Franklin G., 2006. Role of N-terminal residues in the ubiquitin-independent degradation of human thymidylate synthase. *Biochemical Journal* 394, 355-363.
- Peña, M.M.O., Melo, S.P., Xing, Y.-Y., White, K., Barbour, K.W., Berger, F.G., 2009. The intrinsically disordered N-terminal domain of thymidylate synthase targets the enzyme to the ubiquitin-independent proteasomal degradation pathway. *Journal of Biological Chemistry* 284, 31597-31607.
- Pereira, D.A., Williams, J.A., 2007. Origin and evolution of high throughput screening. *British Journal of Pharmacology* 152, 53-61.
- Perry, K.M., Fauman, E.B., Finer-Moore, J.S., Montfort, W.R., Maley, G.F., Maley, F., Stroud, R.M., 1990. Plastic adaptation toward mutations in proteins: structural comparison of thymidylate synthases. *Proteins* 8, 315-333.
- Persson, R., Cedergren-Zeppezauer, E.S., Wilson, K.S., 2001. Homotrimeric dUTPases; structural solutions for specific recognition and hydrolysis of dUTP. *Current Protein & Peptide Science* 2, 287-300.
- Peters, G.J., van der Wilt, C.L., van Triest, B., Codacci-Pisanelli, G., Johnston, P.G., van Groeningen, C.J., Pinedo, H.M., 1995. Thymidylate synthase and drug resistance. *European Journal of Cancer* 31, 1299-1305.
- Peters, G.J., Backus, H.H.J., Freemantle, S., van Triest, B., Codacci-Pisanelli, G., van der Wilt, C.L., Smid, K., Lunec, J., Calvert, A.H., Marsh, S., McLeod, H.L., Bloemena, E., Meijer, S., Jansen, G., van Groeningen, C.J., Pinedo, H.M., 2002. Induction of thymidylate synthase as a 5-fluorouracil resistance mechanism. *Biochimica et Biophysica Acta - Molecular Basis of Disease* 1587, 194-205.
- Phan, J., Steadman, D.J., Koli, S., Ding, W.C., Minor, W., Dunlap, R.B., Berger, S.H., Lebioda, L., 2001. Structure of human thymidylate synthase suggests advantages of chemotherapy with noncompetitive inhibitors. *The Journal of Biological Chemistry* 276, 14170-14177.
- Piche, A., Rancourt, C., 2001. Gene therapy to overcome drug resistance in cancer: targeting key regulators of the apoptotic pathway. *Current Gene Therapy* 1, 317-324.
- Pizzorno, G., Chang, Y.M., McGuire, J.J., Bertino, J.R., 1989. Inherent resistance of human squamous carcinoma cell lines to methotrexate as a result of decreased polyglutamylation of this drug. *Cancer Research* 49, 5275-5280.

- Pogolotti, A.L., Weill, C., Santi, D.V., 1979. Thymidylate synthetase catalyzed exchange of tritium from [5-3H]-2'-deoxyuridylate for protons of water. *Biochemistry* 18, 2794-2798.
- Qiu, L.X., Tang, Q.Y., Bai, J.L., Qian, X.P., Li, R.T., Liu, B.R., Zheng, M.H., 2008. Predictive value of thymidylate synthase expression in advanced colorectal cancer patients receiving fluoropyrimidine-based chemotherapy: evidence from 24 studies. *International journal of cancer. Journal international du cancer* 123, 2384-2389.
- Rahman, L., Voeller, D., Rahman, M., Lipkowitz, S., Allegra, C., Barrett, J.C., Kaye, F.J., Zajac-Kaye, M., 2004. Thymidylate synthase as an oncogene: a novel role for an essential DNA synthesis enzyme. *Cancer Cell* 5, 341-351.
- Ramachandran, G.N., Ramakrishnan, C., Sasisekharan, V., 1963. Stereochemistry of polypeptide chain configurations. *Journal of Molecular Biology* 7, 95-99.
- Ramsay, E.C., Dos Santos, N., Dragowska, W.H., Laskin, J.J., Bally, M.B., 2005. The formulation of lipid-based nanotechnologies for the delivery of fixed dose anticancer drug combinations. *Current Drug Delivery* 2, 341-351.
- Rang, H.P., 2006. The receptor concept: pharmacology's big idea. *British Journal of Pharmacology* 147, S9-S16.
- Raviña Rubira, E., 2011. *The evolution of drug discovery : from traditional medicines to modern drugs*. Weinheim : Wiley-VCH, c2011.
- Rees, D.C., Congreve, M., Murray, C.W., Carr, R., 2004. Fragment-based lead discovery. *Nature Reviews Drug Discovery* 3, 660-672.
- Reichard, P., 1985. Ribonucleotide reductase and deoxyribonucleotide pools. *Basic Life Sciences* 31, 33-45.
- Rothem, L., Aronheim, A., Assaraf, Y.G., 2003. Alterations in the expression of transcription factors and the reduced folate carrier as a novel mechanism of antifolate resistance in human leukemia cells. *The Journal of Biological Chemistry* 278, 8935-8941.
- Rowe, P.B., Lewis, G.P., 1973. Mammalian folate metabolism. Regulation of folate interconversion enzymes. *Biochemistry* 12, 1962-1969.
- Rowinsky, E.K., Donehower, R.C., 1991. The clinical pharmacology and use of antimicrotubule agents in cancer chemotherapeutics. *Pharmacology & Therapeutics* 52, 35-84.
- Roy, S., 1999. Multifunctional enzymes and evolution of biosynthetic pathways: retro-evolution by jumps. *Proteins* 37, 303-309.
- Saga, Y., Suzuki, M., Mizukami, H., Kohno, T., Takei, Y., Fukushima, M., Ozawa, K., 2003. Overexpression of thymidylate synthase mediates

desensitization for 5-fluorouracil of tumor cells. *International Journal of Cancer* 106, 324-326.

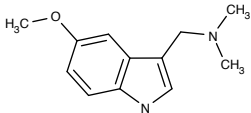
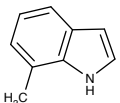
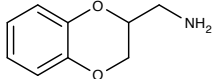
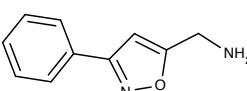
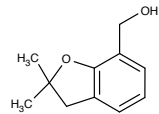
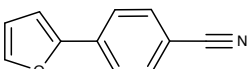
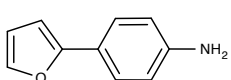
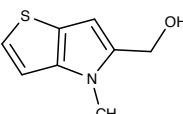
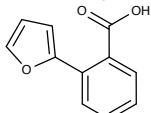
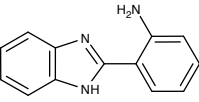
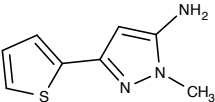
- Saif, M.W., Syrigos, K.N., Katirtzoglou, N.A., 2009. S-1: a promising new oral fluoropyrimidine derivative. *Expert Opinion on Investigational Drugs* 18, 335-348.
- Santi, D.V., McHenry, C.S., Sommer, H., 1974. Mechanism of interaction of thymidylate synthetase with 5-fluorodeoxyuridylate. *Biochemistry* 13, 471-481.
- Schiffer, C.A., Clifton, I.J., Davisson, V.J., Santi, D.V., Stroud, R.M., 1995. Crystal structure of human thymidylate synthase: a structural mechanism for guiding substrates into the active site. *Biochemistry* 34, 16279-16287.
- Schuffenhauer, A., Ruedisser, S., Marzinik, A.L., Jahnke, W., Blommers, M., Selzer, P., Jacoby, E., 2005. Library design for fragment based screening. *Current Topics in Medicinal Chemistry* 5, 751-762.
- Scott, D.E., Coyne, A.G., Hudson, S.A., Abell, C., 2012. Fragment-based approaches in drug discovery and chemical biology. *Biochemistry* 51, 4990-5003.
- Sethi, T., Rintoul, R.C., Moore, S.M., MacKinnon, A.C., Salter, D., Choo, C., Chilvers, E.R., Dransfield, I., Donnelly, S.C., Strieter, R., Haslett, C., 1999. Extracellular matrix proteins protect small cell lung cancer cells against apoptosis: A mechanism for small cell lung cancer growth and drug resistance in vivo. *Nature Medicine* 5, 662-668.
- Shane, B., 1989. Folylpolyglutamate synthesis and role in the regulation of one-carbon metabolism. *Vitamins and Hormones* 45, 263-335.
- Shewach, D.S., Reynolds, K.K., Hertel, L., 1992. Nucleotide specificity of human deoxycytidine kinase. *Molecular Pharmacology* 42, 518-524.
- Showalter, S.L., Showalter, T.N., Witkiewicz, A., Havens, R., Kennedy, E.P., Hucl, T., Kern, S.E., Yeo, C.J., Brody, J.R., 2008. Evaluating the drug-target relationship between thymidylate synthase expression and tumor response to 5-fluorouracil. Is it time to move forward? *Cancer Biology & Therapy* 7, 986-994.
- Shuker, S.B., Hajduk, P.J., Meadows, R.P., Fesik, S.W., 1996. Discovering high-affinity ligands for proteins: SAR by NMR. *Science* 274, 1531-1534.
- Skipper, H.E., Schabel, F.M., Jr., Wilcox, W.S., 1964. Experimental evaluation of potential anticancer agents. XIII. on the criteria and kinetics associated with "curability" of experimental leukemia. *Cancer Chemotherapy Reports. Part 1* 35, 1-111.
- Sotelo-Mundo, R.R., Changchien, L., Maley, F., Montfort, W.R., 2006. Crystal structures of thymidylate synthase mutant R166Q: structural basis for the nearly complete loss of catalytic activity. *Journal of Biochemical and Molecular Toxicology* 20, 88-92.

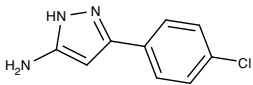
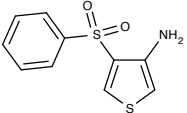
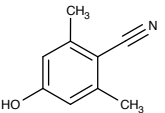
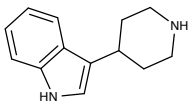
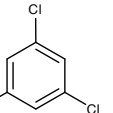
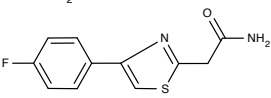
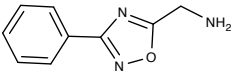
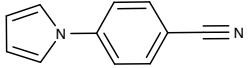
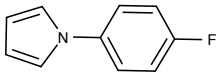
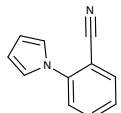
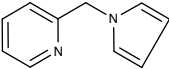
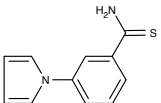
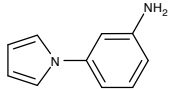
- Spsychala, J., Madrid-Marina, V., Nowak, P.J., Fox, I.H., 1989. AMP and IMP dephosphorylation by soluble high-and low-Km 5'-nucleotidases. *American Journal of Physiology-Endocrinology and Metabolism* 256, E386-E391.
- Stanley McKnight, G., 1991. Cyclic AMP second messenger systems. *Current Opinion in Cell Biology* 3, 213-217.
- Stayton, M.M., Rudolph, F.B., Fromm, H.J., 1983. Regulation, genetics, and properties of adenylosuccinate synthetase: a review. *Current Topics in Cellular Regulation* 22, 103-141.
- Steeg, P.S., Theodorescu, D., 2008. Metastasis: a therapeutic target for cancer. *Nature Clinical Practice. Oncology* 5, 206-219.
- Sudhakar, A., 2009. History of cancer, ancient and modern treatment methods. *Journal of Cancer Science & Therapy* 1, 1-4.
- Szalma, S., Koka, V., Khasanova, T., Perakslis, E.D., 2010. Effective knowledge management in translational medicine. *Journal of Translational Medicine* 8, 68.
- Takeuchi, T., 1995. Antitumor antibiotics discovered and studied at the Institute of Microbial Chemistry. *Journal of Cancer Research and Clinical Oncology* 121, 505-510.
- Thompson, D.K., Garbers, D.L., 1990. Guanylyl cyclase in cell signalling. *Current Opinion in Cell Biology* 2, 206-211.
- Tillman, D.M., Petak, I., Houghton, J.A., 1999. A Fas-dependent component in 5-fluorouracil/leucovorin-induced cytotoxicity in colon carcinoma cells. *Clinical Cancer Research* 5, 425-430.
- Toga, W., Kondo, M., Tokoro, A., 2003. Preclinical and clinical profile of imatinib mesilate, a potent protein-tyrosine kinase inhibitor for CML therapy. *Nihon yakurigaku zasshi. Folia pharmacologica Japonica* 121, 119-128.
- Tong, Y., Liu-Chen, X., Ercikan-Abali, E.A., Capiiaux, G.M., Zhao, S.C., Banerjee, D., Bertino, J.R., 1998. 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. *The Journal of Biological Chemistry* 273, 11611-11618.
- Tsai, J., Lee, J.T., Wang, W., Zhang, J., Cho, H., Mamo, S., Bremer, R., Gillette, S., Kong, J., Haass, N.K., Sproesser, K., Li, L., Smalley, K.S., Fong, D., Zhu, Y.L., Marimuthu, A., Nguyen, H., Lam, B., Liu, J., Cheung, I., Rice, J., Suzuki, Y., Luu, C., Settachatgul, C., Shellooe, R., Cantwell, J., Kim, S.H., Schlessinger, J., Zhang, K.Y., West, B.L., Powell, B., Habets, G., Zhang, C., Ibrahim, P.N., Hirth, P., Artis, D.R., Herlyn, M., Bollag, G., 2008. Discovery of a selective inhibitor of oncogenic B-Raf kinase with potent antimelanoma activity. *Proceedings of the National Academy of Sciences of the United States of America* 105, 3041-3046.

- Van Cutsem, E., Kohne, C.H., Lang, I., Folprecht, G., Nowacki, M.P., Cascinu, S., Shchepotin, I., Maurel, J., Cunningham, D., Tejpar, S., Schlichting, M., Zobel, A., Celik, I., Rougier, P., Ciardiello, F., 2011. Cetuximab plus irinotecan, fluorouracil, and leucovorin as first-line treatment for metastatic colorectal cancer: updated analysis of overall survival according to tumor KRAS and BRAF mutation status. *Journal of Clinical Oncology* 29, 2011-2019.
- Vogelstein, B., Kinzler, K.W., 2004. Cancer genes and the pathways they control. *Nature Medicine* 10, 789-799.
- Walling, J., 2006. From methotrexate to pemetrexed and beyond. A review of the pharmacodynamic and clinical properties of antifolates. *Investigational New Drugs* 24, 37-77.
- Warwick, G.P., 1963. The mechanism of action of alkylating agents. *Cancer Research* 23, 1315-1333.
- Watts, R.W.E., Watts, J.E.M., Seegmiller, J.E., 1965. Xanthine oxidase activity in human tissues and its inhibition by allopurinol (4-hydroxypyrazolo[3,4-d]pyrimidine). *The Journal of Laboratory and Clinical Medicine* 66, 688-697.
- Weeks, L.D., Fu, P., Gerson, S.L., 2013. Uracil-DNA glycosylase expression determines human lung cancer cell sensitivity to pemetrexed. *Molecular Cancer Therapeutics* 12, 2248-2260.
- Wilson, P.M., Danenberg, P.V., Johnston, P.G., Lenz, H.J., Ladner, R.D., 2014. Standing the test of time: targeting thymidylate biosynthesis in cancer therapy. *Nature Reviews. Clinical Oncology* 11, 282-298.
- Yardley, D.A., 2013. Drug resistance and the role of combination chemotherapy in improving patient outcomes. *International Journal of Breast Cancer* 2013, 15.
- Zahreddine, H., Borden, K., 2013. Mechanisms and insights into drug resistance in cancer. *Frontiers in Pharmacology* 4.
- Zalkin, H., Dixon, J.E., 1992. De novo purine nucleotide biosynthesis. *Progress in Nucleic Acid Research and Molecular Biology*. Academic Press, pp. 259-287.
- Zeng, X., Morgenstern, R., Nyström, A.M., 2014. Nanoparticle-directed sub-cellular localization of doxorubicin and the sensitization breast cancer cells by circumventing GST-Mediated drug resistance. *Biomaterials* 35, 1227-1239.
- Zhang, L., Yu, J., Park, B.H., Kinzler, K.W., Vogelstein, B., 2000. Role of BAX in the apoptotic response to anticancer agents. *Science* 290, 989-992.
- Zhao, Y., Adjei, A.A., 2015. Targeting angiogenesis in cancer therapy: moving beyond vascular endothelial growth factor. *The Oncologist* 20, 660-673.

## APPENDIX

Table 1 DSF-based screening hit list from the 1<sup>st</sup> fragment library of 500 compounds

Cpd ID	Structure	MW (Da)	Formula	$\Delta T_m$ (°C) at 1 mM in phosphate screen	$\Delta T_m$ (°C) at 1 mM in dUMP screen	Catalog number
<b>Common hits</b>						
F13		204.3	C <sub>12</sub> H <sub>16</sub> N <sub>2</sub> O	1.05	1.24	AC12586R3
F14		131.2	C <sub>9</sub> H <sub>9</sub> N	1.32	0.89	AC12704R3
F90		165.2	C <sub>9</sub> H <sub>11</sub> NO <sub>2</sub>	1.19	1.13	CC01513R3
F95		174.2	C <sub>10</sub> H <sub>10</sub> N <sub>2</sub> O	1.23	1.34	CC64513R3
F151		178.2	C <sub>11</sub> H <sub>14</sub> O <sub>2</sub>	0.86	0.82	CC27209R3
F168		169.2	C <sub>11</sub> H <sub>7</sub> NO	1.42	0.63	CC34416R3
F169		159.2	C <sub>10</sub> H <sub>9</sub> NO	2.13	2.08	CC34414R3
F185		167.2	C <sub>8</sub> H <sub>9</sub> NOS	2.03	2.38	CC40909R3
F206		188.2	C <sub>11</sub> H <sub>8</sub> O <sub>3</sub>	1.17	0.78	CC46201R3
F261		209.3	C <sub>13</sub> H <sub>11</sub> N <sub>3</sub>	1.17	0.92	DP01095R3
F274		179.2	C <sub>8</sub> H <sub>9</sub> N <sub>3</sub> S	0.65	0.74	KM00835R3

Cpd ID	Structure	MW (Da)	Formula	$\Delta T_m$ ( $^{\circ}\text{C}$ ) at 1 mM in phosphate screen	$\Delta T_m$ ( $^{\circ}\text{C}$ ) at 1 mM in dUMP screen	Catalog number
F277		193.6	C <sub>9</sub> H <sub>8</sub> ClN <sub>3</sub>	1.11	0.65	KM01548R3
F278		239.3	C <sub>10</sub> H <sub>9</sub> NO <sub>2</sub> S <sub>2</sub>	1.27	1.37	KM01765R3
F344		147.2	C <sub>9</sub> H <sub>9</sub> NO	1.18	0.79	RDP00221R3
F356		200.3	C <sub>13</sub> H <sub>16</sub> N <sub>2</sub>	1.71	0.95	RH02036R3
F363		162.0	C <sub>6</sub> H <sub>5</sub> Cl <sub>2</sub> N	0.57	0.54	SB00590R3
F396		236.3	C <sub>11</sub> H <sub>9</sub> FN <sub>2</sub> OS	0.92	0.89	SB02112R3
F489		175.2	C <sub>9</sub> H <sub>9</sub> N <sub>3</sub> O	1.57	1.79	CC76913ZZ
F493		168.2	C <sub>11</sub> H <sub>8</sub> N <sub>2</sub>	3.75	1.21	CC25516ZZ
<b>Phosphate hits</b>						
F240		161.2	C <sub>10</sub> H <sub>8</sub> FN	0.84	-0.44	CD02331R3
F242		168.2	C <sub>11</sub> H <sub>8</sub> N <sub>2</sub>	1.98	-0.01	CD03418R3
F244		158.2	C <sub>10</sub> H <sub>10</sub> N <sub>2</sub>	2.84	-0.22	CD04786R3
F415		202.3	C <sub>11</sub> H <sub>10</sub> N <sub>2</sub> S	1.22	-1.6	SPB01074R3
F461		158.2	C <sub>10</sub> H <sub>10</sub> N <sub>2</sub>	0.76	-0.13	CC21914R3

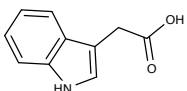
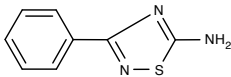
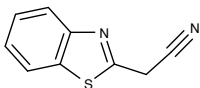
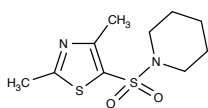
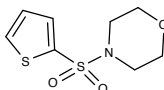
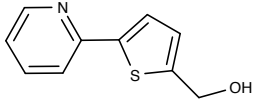
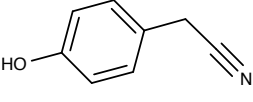
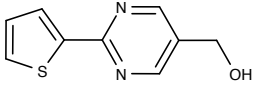
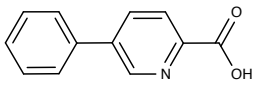
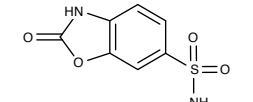
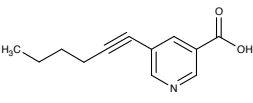
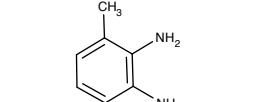
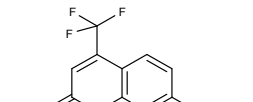
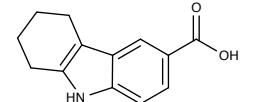
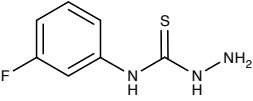
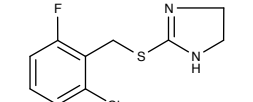
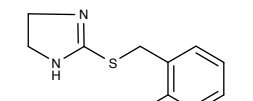
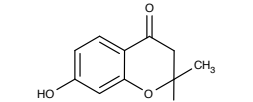
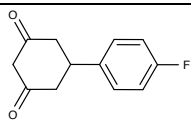
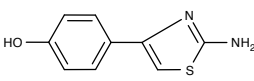
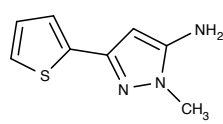
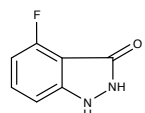
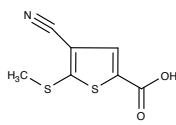
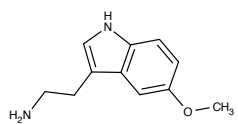
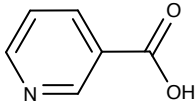
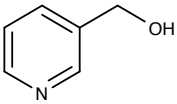
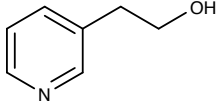
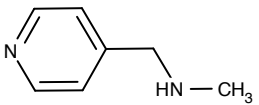
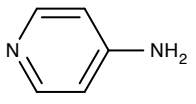
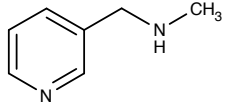
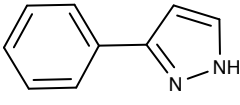
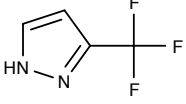
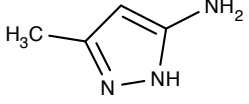
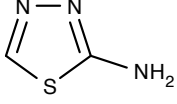
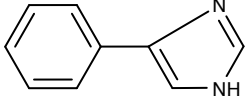
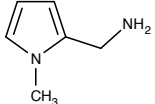
Cpd ID	Structure	MW (Da)	Formula	$\Delta T_m$ ( $^{\circ}\text{C}$ ) at 1 mM in phosphate screen	$\Delta T_m$ ( $^{\circ}\text{C}$ ) at 1 mM in dUMP screen	Catalog number
F470		175.2	C <sub>10</sub> H <sub>9</sub> NO <sub>2</sub>	0.71	0.38	RH01882R3
<b>dUMP hits</b>						
F70		177.2	C <sub>8</sub> H <sub>7</sub> N <sub>3</sub> S	-0.85	0.57	BTB13395R3
F262		174.2	C <sub>9</sub> H <sub>6</sub> N <sub>2</sub> S	0.64	0.92	GK00309R3
F331		260.4	C <sub>10</sub> H <sub>16</sub> N <sub>2</sub> O <sub>2</sub> S <sub>2</sub>	-0.35	0.96	MO08140R3
F335		233.3	C <sub>8</sub> H <sub>11</sub> NO <sub>3</sub> S <sub>2</sub>	-0.1	0.81	MO08155R3

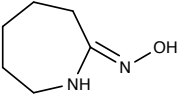
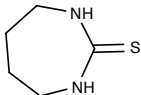
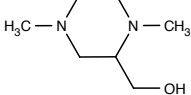
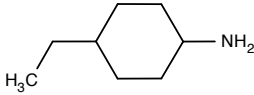
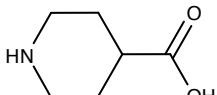
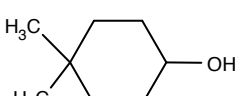
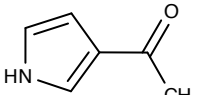
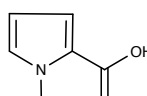
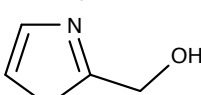
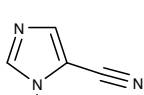
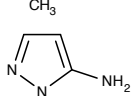
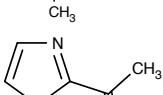
Table 2 DSF-based screening hit list from the 2<sup>nd</sup> fragment library of 1500 compounds

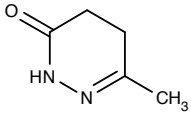
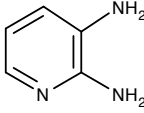
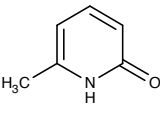
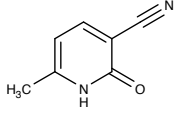
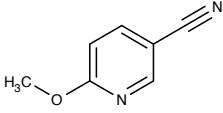
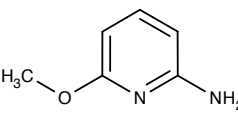
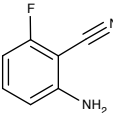
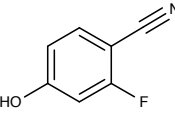
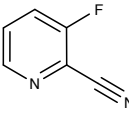
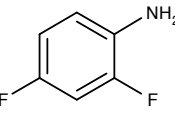
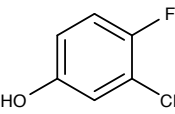
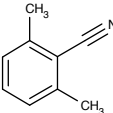
Cpd ID	Structure	MW (Da)	Formula	$\Delta T_m$ (°C) at 1 mM in dUMP screen	Catalog number
C26		191.3	C <sub>10</sub> H <sub>9</sub> NOS	0.57	CC02209R3
C206		133.1	C <sub>8</sub> H <sub>7</sub> NO	0.53	SB00735R3
C335		192.2	C <sub>9</sub> H <sub>8</sub> N <sub>2</sub> OS	0.56	CC66509R3
C361		199.2	C <sub>12</sub> H <sub>9</sub> NO <sub>2</sub>	0.60	CC55001R3
C469		214.2	C <sub>7</sub> H <sub>6</sub> N <sub>2</sub> O <sub>4</sub> S	0.55	CC75719R3
C571		203.2	C <sub>12</sub> H <sub>13</sub> NO <sub>2</sub>	0.63	KM09434R3
C716		122.2	C <sub>7</sub> H <sub>10</sub> N <sub>2</sub>	0.79	BTB07447R3
C858		230.1	C <sub>10</sub> H <sub>5</sub> F <sub>3</sub> O <sub>3</sub>	0.58	RJF00002R3
C874		215.2	C <sub>13</sub> H <sub>13</sub> NO <sub>2</sub>	0.50	MO07745R3
C964		185.2	C <sub>7</sub> H <sub>8</sub> FN <sub>3</sub> S	1.10	RJF00203R3
C993		325.6	C <sub>10</sub> H <sub>11</sub> BrClFN <sub>2</sub> S	2.02	CD05341R3
C1097		263.2	C <sub>10</sub> H <sub>12</sub> Cl <sub>2</sub> N <sub>2</sub> S	1.45	DP01049R3
C1114		192.2	C <sub>11</sub> H <sub>12</sub> O <sub>3</sub>	0.50	S13733R3

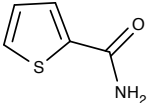
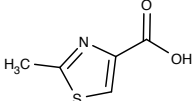
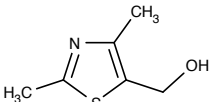
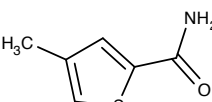
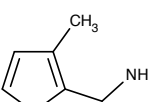
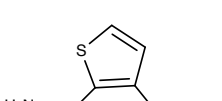
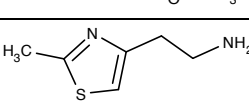
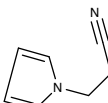
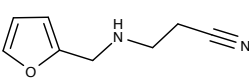
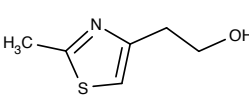
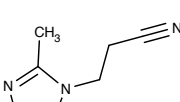
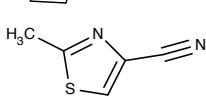
Cpd ID	Structure	MW (Da)	Formula	$\Delta T_m$ ( $^{\circ}\text{C}$ ) at 1 mM in dUMP screen	Catalog number
C1143		206.2	C <sub>12</sub> H <sub>11</sub> FO <sub>2</sub>	0.50	CD11363R3
C1153		192.2	C <sub>9</sub> H <sub>8</sub> N <sub>2</sub> OS	0.81	BTB07444R3
C1238		179.2	C <sub>8</sub> H <sub>9</sub> N <sub>3</sub> S	0.52	KM00835R3
C1242		152.1	C <sub>7</sub> H <sub>5</sub> FN <sub>2</sub> O	0.58	HTS01520R3
C1293		199.3	C <sub>7</sub> H <sub>5</sub> NO <sub>2</sub> S <sub>2</sub>	0.51	MAY00266R3
C1440		190.2	C <sub>11</sub> H <sub>14</sub> N <sub>2</sub> O	1.23	AC39858R3

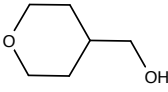
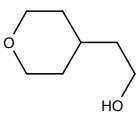
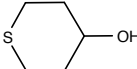
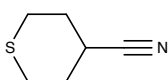
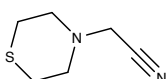
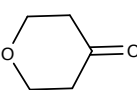
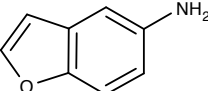
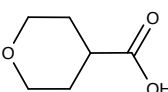
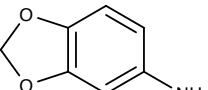
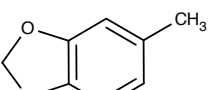
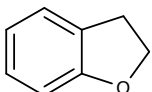
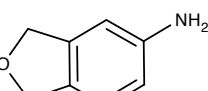
**Table 3 Fragment cocktails**

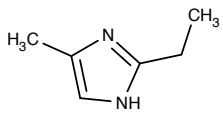
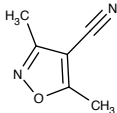
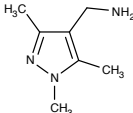
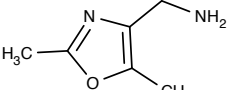
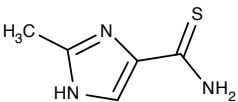
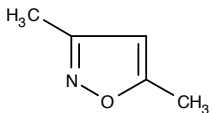
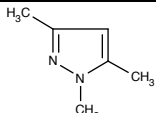
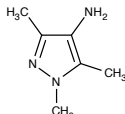
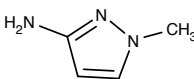
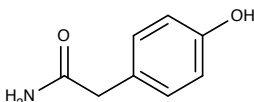
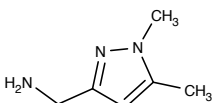
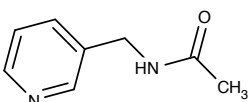
Cocktail number	Cpd ID	Structure	MW (Da)	Formula	Catalog number
1	F100		123.1	C6H5NO2	CC04101R3
	F221		109.1	C6H7NO	CC04109R3
	F295		123.2	C7H9NO	KM09875R3
	F304		122.2	C7H10N2	MO00914R3
	F380		94.1	C5H6N2	SB01321R3
	F437		122.2	C7H10N2	TL00368R3
2	F18		144.2	C9H8N2	AC29603R3
	F268		136.1	C4H3F3N2	HAN00203R3
	F273		97.1	C4H7N3	KM00084R3
	F287		101.1	C2H3N3S	CC08114R3
	F411		144.2	C9H8N2	SEW05612R3
	F466		110.2	C6H10N2	CC02713R3

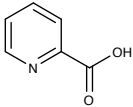
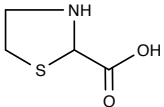
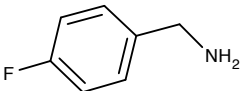
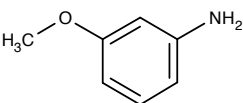
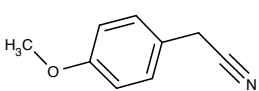
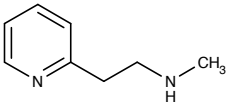
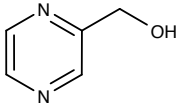
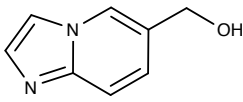
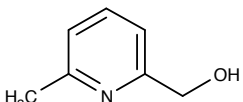
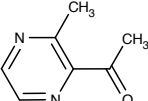
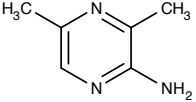
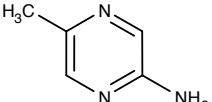
Cocktail number	Cpd ID	Structure	MW (Da)	Formula	Catalog number
3	F253		128.2	C6H12N2O	CD11194R3
	F260		130.2	C5H10N2S	DP01080R3
	F288		144.2	C7H16N2O	KM07851R3
	F313		127.2	C8H17N	MO01181R3
	F366		129.2	C6H11NO2	SB00736R3
	F472		128.2	C8H16O	BTBG00026ZZ
4	F41		109.1	C6H7NO	BTB07157R3
	F96		125.1	C6H7NO2	CC02701R3
	F114		115.2	C4H5NOS	CC09509R3
	F141		107.1	C5H5N3	CC23616R3
	F267		97.1	C4H7N3	GK03066R3
	F368		127.2	C5H5NOS	SB00889R3

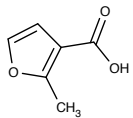
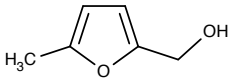
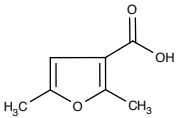
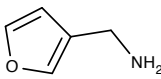
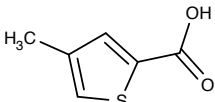
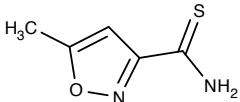
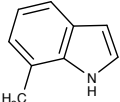
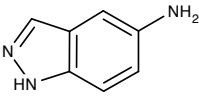
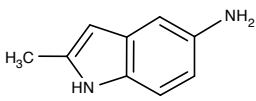
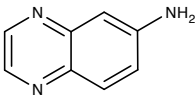
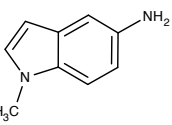
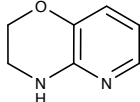
Cocktail number	Cpd ID	Structure	MW (Da)	Formula	Catalog number
5	F37		112.1	C5H8N2O	BTB06143R3
	F53		109.1	C5H7N3	BTB09780R3
	F392		109.1	C6H7NO	SB01977R3
	F422		134.1	C7H6N2O	SPB04732R3
	F423		134.1	C7H6N2O	SPB04785R3
	F488		124.1	C6H8N2O	MO08464ZZ
6	F42		136.1	C7H5FN2	BTB07341R3
	F298		137.1	C7H4FNO	KM10855R3
	F324		122.1	C6H3FN2	MO07483R3
	F362		129.1	C6H5F2N	SB00555R3
	F372		146.5	C6H4ClFO	SB01097R3
	F432		131.2	C9H9N	TL00285R3

Cocktail number	Cpd ID	Structure	MW (Da)	Formula	Catalog number
7	F44		127.2	C5H5NOS	BTB07541R3
	F106		143.2	C5H5NO2S	CC05701R3
	F107		143.2	C6H9NOS	CC05809R3
	F160		141.2	C6H7NOS	CC30617R3
	F414		127.2	C6H9NS	SP01443R3
	F486		143.2	C6H9NOS	CC68613ZZ
8	F229		142.2	C6H10N2S	CC54613R3
	F239		120.2	C7H8N2	CD01949R3
	F292		150.2	C8H10N2O	KM09270R3
	F345		143.2	C6H9NOS	CC54609R3
	F381		135.2	C7H9N3	SB01342R3
	F417		124.2	C5H4N2S	SPB02631R3

Cocktail number	Cpd ID	Structure	MW (Da)	Formula	Catalog number
9	F158		116.2	C6H12O2	CC29909R3
	F187		130.2	C7H14O2	CC67409R3
	F237		118.2	C5H10OS	CC61915R3
	F238		127.2	C6H9NS	CC61916R3
	F247		142.2	C6H10N2S	CD08254R3
	F385		100.1	C5H8O2	SB01650R3
10	F112		133.1	C8H7NO	CC07114R3
	F157		130.1	C6H10O3	CC29901R3
	F361		137.1	C7H7NO2	SB00117R3
	F369		136.1	C8H8O2	SB00943R3
	F383		120.1	C8H8O	SB01482R3
	F491		135.2	C8H9NO	CC77214ZZ

Cocktail number	Cpd ID	Structure	MW (Da)	Formula	Catalog number
11	F11		110.2	C6H10N2	AC11871R3
	F81		122.1	C6H6N2O	CC00616R3
	F128		139.2	C7H13N3	CC14713R3
	F149		126.2	C6H10N2O	CC26813R3
	F397		141.2	C5H7N3S	SB02117R3
	F441		97.1	C5H7NO	TL00573R3
12	F246		110.2	C6H10N2	CD07268R3
	F254		125.2	C6H11N3	CD11242R3
	F263		97.1	C4H7N3	GK00321R3
	F386		151.2	C8H9NO2	SB01670R3
	F409		125.2	C6H11N3	CC12613R3
	F450		150.2	C8H10N2O	TL00908R3

Cocktail number	Cpd ID	Structure	MW (Da)	Formula	Catalog number
13	F99		123.1	C6H5NO2	CC04001R3
	F284		133.2	C4H7NO2S	KM05539R3
	F373		125.1	C7H8FN	SB01159R3
	F378		123.2	C7H9NO	SB01292R3
	F379		147.2	C9H9NO	SB01293R3
	F494		136.2	C8H12N2	BTB11191ZZ
14	F97		110.1	C5H6N2O	CC03409R3
	F162		148.2	C8H8N2O	CC33109R3
	F360		123.2	C7H9NO	RJC03579R3
	F471		136.2	C7H8N2O	SB01015R3
	F473		123.2	C6H9N3	BTB07911ZZ
	F492		109.1	C5H7N3	MO08458ZZ

Cocktail number	Cpd ID	Structure	MW (Da)	Formula	Catalog number
15	F21		126.1	C6H6O3	AC33015R3
	F23		112.1	C6H8O2	AC35851R3
	F85		140.1	C7H8O3	CC01101R3
	F93		97.1	C5H7NO	CC02113R3
	F265		142.2	C6H6O2S	GK01984R3
	F421		142.2	C5H6N2OS	SPB04532R3
16	F14		131.2	C9H9N	AC12704R3
	F20		133.2	C7H7N3	AC30454R3
	F22		146.2	C9H10N2	AC34875R3
	F118		145.2	C8H7N3	CC11714R3
	F189		146.2	C9H10N2	CC41414R3
	F236		136.2	C7H8N2O	CC60830R3

**Table 4 Data collection and refinement statistics of hTS in complex with screening hits**

Dataset	hTS-dUMP-F331	hTS-dUMP-F335	hTS-dUMP-C993	hTS-dUMP-F95
Data collection				
X-ray source	AS-MX1	AS-MX1	AS-MX1	AS-MX1
Wavelength, Å	0.9537	0.9537	0.9537	0.9537
Space group	P <sub>4</sub> <sub>3</sub> -2 <sub>1</sub> -2	P <sub>4</sub> <sub>3</sub> -2 <sub>1</sub> -2	P <sub>4</sub> <sub>3</sub> -2 <sub>1</sub> -2	P <sub>4</sub> <sub>3</sub> -2 <sub>1</sub> -2
Cell parameters				
a, b, c in Å	109.22, 109.22, 317.40	107.90, 107.90, 315.67	110.12, 110.12, 316.30	109.59, 109.59, 317.47
α, β, γ in degree	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Resolution range, Å	30 – 2.90 (3.00 – 2.90)	48.25 – 2.43 (2.48 – 2.43)	30 – 2.40 (2.49 – 2.40)	30 – 2.41 (2.50 – 2.41)
No. observed	327 718	1 149 734	613 018	599 566
No. unique reflections	42 537	71 419	65 735	71 227
Completeness, %	97.2 (97.8)	99.8 (97.4)	84.7 (74.3)	93.9 (93.7)
<i>R</i> <sub>merge</sub> , %	14.5 (59.0)	19.1 (87.3)	13.8 (51.7)	11.8 (55.6)
Average I/σ	11.1 (2.0)	9.9 (2.5)	14.3 (3.8)	14.55 (3.0)
Refinement statistics				
Resolution range, Å	30 – 2.90	48.25 – 2.43	30 – 2.40	30 – 2.41
<i>R</i> <sub>work</sub> , %	25.59	18.63	19.05	18.99
<i>R</i> <sub>free</sub> , %	36.40	26.64	25.03	25.94
Protein/asymmetric unit	6	6	6	6
Ligand	6 dUMP and 1 F331	6 dUMP and 5 F335	6 dUMP and 6 C993	6 dUMP and 1 F95
Water molecules	161	463	476	748
Root mean square deviations				
Bond lengths, Å	0.010	0.008	0.011	0.009
Bond angles, degree	1.328	1.081	1.106	1.102

**Table 4 Data collection and refinement statistics of hTS in complex with screening hits**

Dataset	hTS-dUMP-F489	hTS-dUMP-C1153	hTS-dUMP-F13	hTS-dUMP-C1440
<b>Data collection</b>				
X-ray source	AS-MX1	AS-MX2	NSRRC-13B1	AS-MX1
Wavelength, Å	0.9537	1.0130	0.9789	0.9537
Space group	P <sub>4</sub> -2 <sub>1</sub> -2	P <sub>4</sub> -2 <sub>1</sub> -2	P <sub>4</sub> -2 <sub>1</sub> -2	P <sub>4</sub> -2 <sub>1</sub> -2
<b>Cell parameters</b>				
a, b, c in Å	108.24, 108.24, 317.74	109.45, 109.45, 317.41	109.35, 109.35, 315.38	108.83, 108.83, 317.85
α, β, γ in degree	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Resolution range, Å	30 – 2.85 (2.95 – 2.85)	49.08 – 2.77 (2.86 – 2.77)	30 – 2.70 (2.80 – 2.70)	30 – 2.57 (2.66 – 2.57)
No. observed	426 221	802 157	569 527	1 299 616
No. unique reflections	42 230	50 128	53 294	61 335
Completeness, %	93.5 (89.2)	99.7 (96.9)	98.4 (91.1)	98.8 (96.1)
<i>R</i> <sub>merge</sub> , %	11.6 (55.5)	17.2 (63.5)	11.4 (42.9)	11.3 (47.1)
Average I/σ	19.4 (4.0)	11.0 (3.7)	20.6 (2.4)	26.3 (6.4)
<b>Refinement statistics</b>				
Resolution range, Å	30 – 2.85	49.08 – 2.77	30 – 2.70	30 – 2.57
<i>R</i> <sub>work</sub> , %	21.98	20.78	18.91	19.09
<i>R</i> <sub>free</sub> , %	27.39	29.81	25.46	25.57
Protein/asymmetric unit	6	6	6	6
Ligand	6 dUMP and 1 F489	6 dUMP and 6 C1153	6 P <sub>i</sub> and 2 F13	6 dUMP and 2 C1440
Water molecules	114	186	205	331
<b>Root mean square deviations</b>				
Bond lengths, Å	0.017	0.009	0.009	0.008
Bond angles, degree	1.297	1.206	1.198	1.131

**Table 5 Data collection and refinement statistics of hTS in complex with elaborated compounds**

Dataset	hTS-dUMP-A5	hTS-dUMP-A8	hTS-dUMP-A11	hTS-dUMP-A15
Data collection				
X-ray source	AS-MX2	AS-MX2	AS-MX1	AS-MX1
Wavelength, Å	1.0130	1.0130	0.9537	0.9537
Space group	P4 <sub>1</sub> 2 <sub>1</sub> 2	P4 <sub>1</sub> 2 <sub>1</sub> 2	P4 <sub>1</sub> 2 <sub>1</sub> 2	P4 <sub>1</sub> 2 <sub>1</sub> 2
Cell parameters				
a, b, c in Å	109.30, 109.30, 105.46	108.70, 108.70, 105.29	109.10, 109.10, 105.45	108.55, 108.55, 105.04
α, β, γ in degree	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Resolution range, Å	30 – 2.22 (2.30 – 2.22)	48.61 – 3.07 (3.28 – 3.07)	30 – 2.57 (2.66 – 2.57)	30 – 2.70 (2.80 – 2.70)
No. observed	408 372	188 768	164 456	213 155
No. unique reflections	32 201	12 344	20 752	17 824
Completeness, %	99.9 (100.0)	99.7 (98.2)	99.8 (100.0)	99.8 (100.0)
<i>R</i> <sub>merge</sub> , %	7.8 (24.6)	17.3 (60.5)	5.5 (23.9)	10.9 (60.9)
Average I/σ	30.5 (10.5)	14.8 (5.4)	36.4 (8.3)	23.3 (4.2)
Refinement statistics				
Resolution range, Å	30 – 2.22	48.61 – 3.07	30 – 2.57	30 – 2.70
<i>R</i> <sub>work</sub> , %	17.16	15.62	16.77	16.65
<i>R</i> <sub>free</sub> , %	22.73	25.02	25.64	24.60
Protein/asymmetric unit	2	2	2	2
Ligand	2 dUMP and 2 <b>A5</b>	2 dUMP and 2 <b>A8</b>	2 dUMP and 2 <b>A11</b>	2 dUMP and 2 <b>A15</b>
Water molecules	247	11	190	93
Root mean square deviations				
Bond lengths, Å	0.009	0.010	0.008	0.009
Bond angles, degree	1.210	1.191	1.139	1.132

**Table 5 Data collection and refinement statistics of hTS in complex with elaborated compounds**

Dataset	hTS-dUMP-A27	hTS-dUMP-A33	hTS-dUMP-A34	hTS-dUMP-A36
Data collection				
X-ray source	AS-MX1	AS-MX2	AS-MX2	AS-MX2
Wavelength, Å	0.9537	1.0130	1.0130	1.0130
Space group	P4 <sub>1</sub> 2 <sub>1</sub> 2	P4 <sub>1</sub> 2 <sub>1</sub> 2	P4 <sub>1</sub> 2 <sub>1</sub> 2	P4 <sub>3</sub> 2 <sub>1</sub> 2
Cell parameters				
a, b, c in Å	108.52, 108.52, 105.26	109.19, 109.19, 105.54	109.59, 109.59, 105.66	109.26, 109.26, 315.81
α, β, γ in degree	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Resolution range, Å	30 – 2.40 (2.49 – 2.40)	30 – 2.10 (2.18 – 2.10)	30 – 2.20 (2.28 – 2.20)	30 – 2.09 (2.16 – 2.09)
No. observed	364 178	261 383	247 396	1 280 911
No. unique reflections	25 213	38 425	33 332	114 140
Completeness, %	99.9 (100.0)	99.7 (100.0)	99.8 (100.0)	99.8 (98.7)
<i>R</i> <sub>merge</sub> , %	8.6 (52.8)	8.6 (36.3)	9.2 (63.2)	10.1 (54.7)
Average I/σ	35.8 (6.0)	18.9 (3.8)	19.9 (3.3)	23.1 (4.0)
Refinement statistics				
Resolution range, Å	30 – 2.40	30 – 2.10	30 – 2.20	30 – 2.09
<i>R</i> <sub>work</sub> , %	17.69	19.28	18.35	18.11
<i>R</i> <sub>free</sub> , %	25.19	23.88	24.69	22.08
Protein/asymmetric unit	2	2	2	6
Ligand	2 dUMP and 2 <b>A27</b>	2 dUMP and 2 <b>A33</b>	2 dUMP and 2 <b>A34</b>	6 dUMP and 6 <b>A36</b>
Water molecules	194	468	395	1323
Root mean square deviations				
Bond lengths, Å	0.008	0.008	0.008	0.008
Bond angles, degree	1.173	1.082	1.104	1.083

**Table 5 Data collection and refinement statistics of hTS in complex with elaborated compounds**

Dataset	hTS-dUMP-A37	hTS-dUMP-B8	hTS-dUMP-B9	hTS-dUMP-B10
Data collection				
X-ray source	AS-MX1	AS-MX1	AS-MX1	AS-MX1
Wavelength, Å	0.9537	0.9537	0.9537	0.9537
Space group	P <sub>4</sub> -2 <sub>1</sub> -2	P <sub>4</sub> -2 <sub>1</sub> -2	P <sub>2</sub> -2 <sub>1</sub> -2 <sub>1</sub>	P <sub>2</sub> -2 <sub>1</sub> -2 <sub>1</sub>
Cell parameters				
a, b, c in Å	109.72, 109.72, 316.75	109.50, 109.50, 315.92	101.57, 107.90, 114.298	101.68, 107.79, 113.79
α, β, γ in degree	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Resolution range, Å	30 – 2.15 (2.23 – 2.15)	30 – 2.30 (2.38 – 2.30)	30 – 1.90 (1.97 – 1.90)	30 – 1.82 (1.89 – 1.82)
No. observed	665 529	587 823	748 106	681 506
No. unique reflections	100 449	86 162	99 056	111 608
Completeness, %	94.5 (93.2)	99.6 (100.0)	99.8 (99.9)	99.3 (97.5)
R <sub>merge</sub> , %	8.0 (50.7)	11.8 (54.2)	8.9 (40.3)	5.7 (39.8)
Average I/σ	18.8 (2.6)	17 (3.8)	21.7 (4.7)	27.3 (3.6)
Refinement statistics				
Resolution range, Å	30 – 2.15	30 – 2.30	30 – 1.90	30 – 1.82
R <sub>work</sub> , %	19.82	19.07	15.78	15.77
R <sub>free</sub> , %	25.31	24.40	20.12	19.42
Protein/asymmetric unit	6	6	4	4
Ligand	6 dUMP and 5 A37	6 dUMP and 6 B8	4 dUMP and 4 B9	4 dUMP and 4 B10
Water molecules	1134	1005	1562	1600
Root mean square deviations				
Bond lengths, Å	0.008	0.008	0.007	0.007
Bond angles, degree	1.122	1.116	1.033	1.080

**Table 6 Data collection and refinement statistics of VH-hTS complexes**

Dataset	VH2-hTS	VH4-hTS
Data collection		
X-ray source	AS-MX2	AS-MX2
Wavelength, Å	1.0130	1.0130
Space group	P6 <sub>3</sub> 22	P6 <sub>3</sub> 22
Cell parameters		
a, b, c in Å	145.10, 145.10, 104.28	145.20, 145.20, 104.75
α, β, γ in degree	90, 90, 120	90, 90, 120
Resolution range, Å	30 – 2.45 (2.54 – 2.45)	30 – 2.40 (2.49 – 2.40)
No. observed reflections	476 026	521 642
No. unique reflections	24 320	26 017
Completeness, %	99.9 (100.0)	99.9 (100.0)
<i>R</i> <sub>merge</sub> , %	8.8 (44.6)	9.0 (44.8)
Average I/σ	40.0 (10.2)	39.6 (11.0)
Refinement statistics		
Resolution range, Å	30 – 2.45	30 – 2.40
<i>R</i> <sub>work</sub> , %	18.17	17.22
<i>R</i> <sub>free</sub> , %	23.08	21.87
Protein/asymmetric unit	1 VH2 and 1 hTS	1VH4 and 1 hTS
Ligand	0	0
Water molecules	175	193
Root mean square deviations		
Bond lengths, Å	0.010	0.010
Bond angles, degree	1.324	1.276

**Table 7 Data collection and refinement statistics of hTS-dUMP-CBK115334**

Dataset	hTS-dUMP-CBK115334
Data collection	
X-ray source	AS-MX1
Wavelength, Å	0.9537
Space group	P4 <sub>3</sub> 2 <sub>1</sub> 2
Cell parameters	
a, b, c in Å	108.20, 108.20, 313.93
α, β, γ in degree	90, 90, 90
Resolution range, Å	30 – 3.10 (3.21 – 3.10)
No. observed reflections	413 249
No. unique reflections	34 749
Completeness, %	99.9 (99.9)
<i>R</i> <sub>merge</sub> , %	14.7 (61.9)
Average I/σ	18.8 (5.0)
Refinement statistics	
Resolution range, Å	30 – 3.10
<i>R</i> <sub>work</sub> , %	17.97
<i>R</i> <sub>free</sub> , %	26.44
Protein/asymmetric unit	6
Ligand	6 dUMP and 3 CBK115334
Water molecules	22
Root mean square deviations	
Bond lengths, Å	0.013
Bond angles, degree	1.092