Final Report

Identification of effective antifolates against *Toxoplasma gondii* dihydrofolate reductase thymidylate synthase (TgDHFR-TS) and determination of crystal structure of inhibitor–TgDHFR-TS complex

Jarunee Vanichtanankul

30 November 2018

Contract No. TRG5780138

Final Report

Identification of effective antifolates against *Toxoplasma gondii* dihydrofolate reductase thymidylate synthase (TgDHFR-TS) and determination of crystal structure of inhibitor–TgDHFR-TS complex

Jarunee Vanichtanankul

National Center for Genetic Engineering and Biotechnology (BIOTEC)

National Science and Technology Development Agency (NSTDA)

Abstract

_____

Project Code: TRG5780138

Project Title: Identification of effective antifolates against _Toxoplasma gondii_ dihydrofolate reductase thymidylate synthase (TgDHFR-TS) and determination of crystal structure of inhibitor–TgDHFR-TS complex

Investigator: Jarunee Vanichtanankul

E-mail Address: jarunee@biotec.or.th

Project Period: 2 มิถุนายน 2557 – 1 มิถุนายน 2559 (ขยายโครงการถึง 1 ธันวาคม 2561)

Abstract:

Pyrimethamine (Pyr), an inhibitor of the dihydrofolate reductase (DHFR) enzyme, has been used for decades to treat toxoplasmosis caused by _Toxoplasma gondii_ (Tg) infection. However, Pyr treatment is effective only at high doses that causes antifolate associated toxicity to patients. Therefore, safer alternative treatments are required. In this study, we established a simple bacterial surrogate system, TgDKOTolC, and used it for screening BIOTEC antifolates against the bifunctional DHFR-thymidylate synthase target (TgDHFR-TS). We found that more than 50% of compounds tested are more effective against TgDHFR-TS than Pyr. Several of these compounds are also active by enzyme inhibition assay and _T. gondii_ cell-based assay. To understand the binding mode of Pyr and selected inhibitors, high resolution X-ray structures of the full-length TgDHFR-TS-inhibitor complexes were explored. We found that there is conflict interaction of _p_-chlorophenyl of Pyr with Thr83 of TgDHFR which may deteriorate its binding affinity against TgDHFR. For those effective compounds, no such conflict interaction but additional electrostatic and hydrophobic interactions in the active site pocket was observed. A unique kinked crossover helix was identified in the TgDHFR-TS structure which plays an important role in DHFR-TS interdomain interactions. The identification of high-affinity inhibitors of TgDHFR-TS and their modes of binding together with the simple screening tools could aid development of new effective antifolates for toxoplasmosis treatment.

Keywords: _Toxoplasma gondii_, dihydrofolate reductase, antifolate, bacteria surrogate, crystal structure

Acknowledgement

Table of Contents

List of Tables

Page

List of Illustrations

List of Abbreviations

CCCP	Carbonyl cyanide *m*-chlorophenyl hydrazine
Ch	*Cryptosporidium hominis*
cLogP	A computed octanal-water partition coefficient
DHF	Dihydrofolate, natural substrate of DHFR
DHFR	Dihydrofolate reductase
DHPS	Dihydropteroate synthase
DTT	Dithiothreitol
dUMP	Deoxyuridine monophosphate
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
EPIs	Efflux pump inhibitors
HFFs	Human Foreskin Fibroblasts
Hs	Homo sapiens
IC_{50}	50% Inhibitory concentration
IPTG	Isopropyl-ß-D-thiogalactopyranoside
KP	Potassium phosphate
MM	Minimal media
MTX	Methotrexate
NADPH	Nicotinamide Adenine Dinucleotide Phosphate reduced
PAßN	Phenylalanine arginine ß-naphthylamide
PEcΔTF	*E. coli* BL21 (DE3) strain whose *thyA* and *folA* genes were knocked out
PEG	Polyethylene glycol
Pf	*Plasmodium falciparum*
Pyr	Pyrimethamine
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
TES	N-[Tris(hydroxymethyl)methyl]-2-aminoethanesulfonic acid
Tg	*Toxoplasma gondii*
TgDKO	PEcΔTF cell expressing TgDHFR-TS
TgDKO-TolC	PEcΔTFTolC cell expressing TgDHFR-TS
Tmp	Trimethoprim
TS	Thymidylate synthase

1. Abstract

Pyrimethamine (Pyr), an inhibitor of the dihydrofolate reductase (DHFR) enzyme, has been used for decades to treat toxoplasmosis caused by *Toxoplasma gondii* (Tg) infection. However, Pyr treatment is effective only at high doses that causes antifolate associated toxicity to patients. Therefore, safer alternative treatments are required. In this study, we established a simple bacterial surrogate system, TgDKOTolC, and used it for screening BIOTEC antifolates against the bifunctional DHFR-thymidylate synthase target (TgDHFR-TS). We found that more than 50% of compounds tested are more effective against TgDHFR-TS than Pyr. Several of these compounds are also active by enzyme inhibition assay and *T. gondii* cell-based assay. To understand the binding mode of Pyr and selected inhibitors, high resolution X-ray structures of the full-length TgDHFR-TS-inhibitor complexes were explored. We found that there is conflict interaction of *p*-chlorophenyl of Pyr with Thr83 of TgDHFR which may deteriorate its binding affinity against TgDHFR. For those effective compounds, no such conflict interaction but additional electrostatic and hydrophobic interactions in the active site pocket was observed. A unique kinked crossover helix was identified in the TgDHFR-TS structure which plays an important role in DHFR-TS interdomain interactions. The identification of high-affinity inhibitors of TgDHFR-TS and their modes of binding together with the simple screening tools could aid development of new effective antifolates for toxoplasmosis treatment.

2. Executive summary

Toxoplasmosis is one of five neglected parasitic diseases recognized by Centers for Disease Control and Prevention (CDC). It is a major foodborne disease responsible for deaths and hospitalizations. Toxoplasmosis is caused by infection with *Toxoplasma gondii*, an intracellular protozoan parasite. Standard treatment for acute toxoplasmosis is a combination of pyrimethamine (Pyr) plus a sulfa drug, usually sulfadiazine, which are inhibitors of the target enzymes dihydrofolate reductase (DHFR) and dihydropteroate synthase (DHPS) respectively. Although Pyr/sulfa combination is still effective for toxoplasmosis, high doses are needed. Thus, unavoidable side effects such as blood abnormality, anorexia, convulsion and vomiting are observed. Therefore, alternative effective drugs with lower toxicity are required.

The purpose of this project is to screen for new effective antifolates against TgDHFR-TS using a bacterial surrogate system and to identify their binding modes. At the beginning of the project, we employed PEcΔTF-TgDT (TgDKO), an *E. coli* BL21 (DE3) strain whose *thyA* and *folA* genes were knocked out and reconstituted by *Tgdhfr-ts* genes, for evaluating inhibitory effect of antifolates against TgDHFR. However, we found that most of the tested inhibitors were inactive against TgDKO and some are not in good agreement with the results from enzyme inhibition assay. We also found that

the discrepancy was partly due to active pumping out of antifolates from the cells, leading to inadequate intracellular concentration of the inhibitors for growth inhibition. Efflux pump inhibitors (EPIs) such as CCCP and PAßN, were found to enhance the inhibitory effect of antifolates, EDTA, probenecid and verapamil were not, presumably by causing accumulation of compounds in the cells by blocking energy dependent RND family efflux pump. This phenomenon is confirmed by employing a PEcΔTF strain whose *TolC* gene was knocked out (PEcΔTFTolC) as the surrogate host. We found that PEcΔTFTolC cells expressing TgDHFR-TS (PEcΔTFTolC-TgDT or TgDKO-TolC) showed more susceptible to most antifolates than TgDKO. These data confirm that TolC-dependent efflux pump plays an important role in antifolate susceptibility of PEcΔTF-TgDT and in line with previous report on TolC-dependent methotrexate resistant *E. coli*. These information prompt us to employ TgDKO-TolC as a new surrogate model for evaluation of antifolate potency against TgDHFR-TS.

We screened a set of 218 compounds including three scaffolds, 182 diaminopyrimidines, 18 diaminodihydrotriazines and 18 diaminoquinazolines, using TgDKO-TolC, the new surrogate system. More than 50% of compounds tested (134 from 218 compounds) are more effective against TgDHFR-TS than Pyr (IC_{50} < 38.1 µM). Further analysis reveals that the TgDKO-TolC surrogate system is suitable for testing compounds having cLogP greater than 3, which there is correlation between the new surrogate system and enzyme inhibition assays (r^2 > 0.6). We found that about 30% of the inhibitors are more effective than Pyr by both assays. Several of these compounds are also active by *T. gondii* inhibition assay.

In order to gain information on binding mode of TgDHFR-TS and inhibitors for drug development, co-complex structures of TgDHFR-TS were attempted. We have successfully optimized the condition for crystallization of full-length TgDHFR-TS with inhibitors. Diffraction data of 15 co-complex protein crystals were collected at NSRRC, Taiwan with high resolution of 2.05 to 2.85 Å. Structure refinement of Pyr–TgDHFR-TS complex revealed conflict interaction of *p*-Cl phenyl of Pyr with Thr83 of TgDHFR, resemble the conflict interaction of Pyr with Ser108Asn found in *Plasmodium falciparum* DHFR resistant to Pyr, thus explaining the high dose requirement for toxoplasmosis treatment. Compounds that can avoid such conflict were with good potency against TgDHFR. With the shared similarity around the active site of Pf- and Tg- DHFR, it is promising to identify new anti-TgDHFR from antifolate library designed for Pyr-resistant PfDHFR.

In conclusion, the tools and knowledge generated from this project should facilitate discovery of novel effective antifolates for toxoplasmosis treatment.

3. Objective

The aim of this project is to identify effective antifolates against TgDHFR-TS from BIOTEC antifolate library using a bacterial surrogate system, and to understand the binding mode of selected antifolates by crystal structure determination.

4. Research methodology

4.1 Compound selection and preparation

A set of antifolate compounds has been selected from BIOTEC antifolate library available in our laboratory (Protein-Ligand Engineering and Molecular Biology Laboratory, Medical Molecular Biology Research Unit, BIOTEC, NSTDA). The library can be categorized into three sub-libraries according to the core structure (S1 subsite) as 2,4-diaminopyrimidine– (Pyrimidine), 2,4-diamino-dihydrotriazine– (Dihydrotriazine) and 2,4-diaminoquinazoline– (Quinazoline) derivatives with about 500, 600 and 19 compounds in size, respectively (**Figure 1**). These compounds can be further classified, based on their S2 and S3 subsites, as rigid, semi-rigid and flexible types. Pyrimethamine (Pyr) and trimethoprim (Tmp) are rigid and semi-rigid types of pyrimidine scaffolds, while cycloguanil (Cyc) and WR99210 are rigid and flexible types of dihydrotriazine scaffolds

A set of 218 compounds was selected from pyrimidines (182 analogs), dihydrotriazines (18 analogs) and quinazolines (18 analogs) (**Table 1**). Details of chemical structures of these selected compounds are as shown in Appendix 8.1. All compounds were prepared as stock solution of 10, 25, 50 and 100 mM in dimethyl sulfoxide (DMSO).
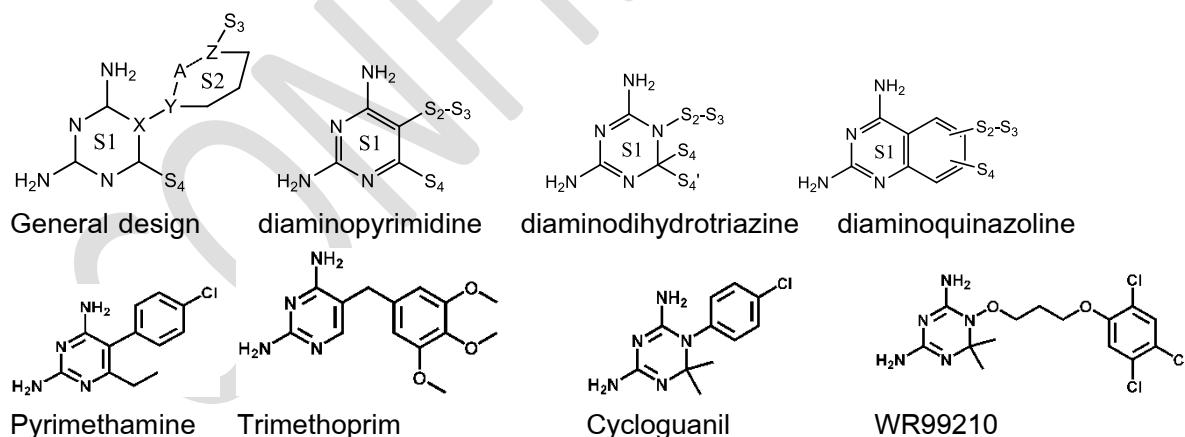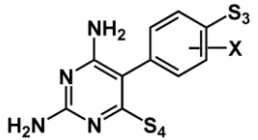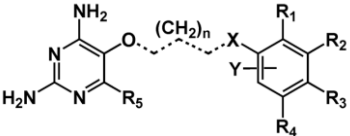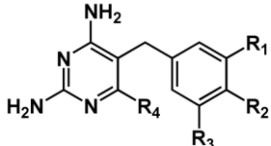
Figure 1. Chemical structure of antifolates.
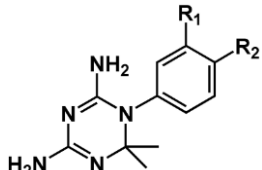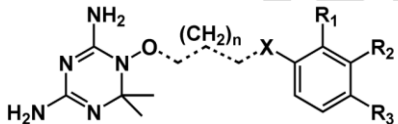
General structure of DHFR inhibitors in BIOTEC antifolates collection is with 4 subsites, S_1, S_2, S_3 and S_4. Three main scaffolds of antifolates (S_1) are pyrimidines, dihydrotriazines and quinazolines with diamino substitution. Substitutions at S2-S3 defines the classification of antifolates into rigid (pyrimethamine- and cycloguanil- like compounds), semi-rigid (trimethoprim-like compounds) and flexible (WR99210- like compounds) inhibitors.

Table 1. Details structure of 218 compounds in this study
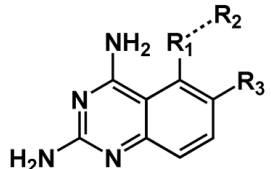
Pyrimidine scaffold (182 compounds)

Rigid compounds	42 compounds (Pyr included) X = Phenyl, Triazole S3 = halide, carboxylic, etc S4 = H, Me, Et, Ph, Pr, iPr, *n*-hexyl, *n*-nonyl, OMe, etc
Flexible compounds	85 compounds (P218 included) n = 1-3, X = C, O or S, Y = Ph, Quinoline R1, R2, R3, R4, R5 = halide, carboxylic, etc
Semi-rigid compounds	55 compounds (Tmp included) R1 = H, OMe, OEt, OPr, OBu, OHep, etc R2 = H, OMe, OHep, OCH_2Ph, OC_3H_6Ph, etc R3 = H or OMe R4 = H, Me or Et

Dihydrotriazine scaffold (18 compounds)

Rigid compounds	3 compounds (Cyc included) R1, R2 = H or Cl
Flexible compounds	15 compounds n = 1-3 X = O or S R1, R2 = H or Cl R3 = H, Cl, Br, Me, NO_2 or COOH

Quinazoline scaffold (18 compounds)

	18 compounds R1 = H or O R2 = Me, Et, Bt, tert-butyl, etc R3 = H or O-$(CH_2)_3$-O-(2,4,5-Cl-Ph)

4.2 Bacterial surrogate model and growth inhibitory assay

Two bacterial cells, PEcΔTF-TgDT (TgDKO) and PEcΔTFTolC-TgDT (TgDKO-TolC), were employed in this study as surrogate hosts for evaluation of inhibitory effect of antifolates. TgDKO is *E. coli* BL21 (DE3) whose indigenous genes encoding for TS and DHFR enzymes, *thyA* and *folA*, respectively, were deleted (PEcΔTF) and reconstituted by Tgdhfr-ts (PEcΔTF was obtained from Nattida Suwanakitti, BIOTEC). TgDKO-TolC is TgDKO with whose TolC encoded gene has been knocked out (PEcΔTFTolC was from Yuwadee Talawanich, BIOTEC). Bacterial cells were overnight cultured in Luria-Bertani (LB) broth at 37°C and diluted to OD_{600} = 0.05 in LB, minimal media (MM) plus 0.2% and 0.4% casamino acid for system optimization. The growth inhibition assay was performed in a 96–well plate with 2–fold serial dilution of inhibitors at the final concentration of 1-2% DMSO in the selected media. The plate was incubated at 37°C, 800 rpm in a microplate incubator shaker (PST-60HL-4, Plate Shaker-Thermostat, Biosan). The cell growth was monitored every hour for about 6–7 hrs incubation time at the wavelength 600 nm using Multiskan™ GO Microplate Spectrophotometer. For about 6–7 hrs incubation time. At this point, cell growth reached the stationary phase. At this point, the data were used to calculate the half maximal inhibitory concentration (IC_{50}) by dose-response curve of growth inhibitory. Each drug was tested in triplicate in for at least three replications separate experiments.

4.3 *In vitro T. gondii* testing

Parasitic protozoa, *T. gondii*, RH-GFP (NR-743) were requested from Malaria Research and Reference Reagent Resource Center (MR4) through the Biodefense and Emerging Infections Research Repository (BEI Resources, USA). The parasite was cultured in Human Foreskin Fibroblasts (HFFs) were cultured in Minimum Essential Media (MEM) supplemented with 2.2 g/L $NaHCO_3$, 1% non-essential amino acid, and 2% FBS (Fetal Bovine Serum). For drug testing, 4000 HFFs were seeded into 96 well plates containing MEM with 10% FBS and incubated to form monolayer for 6 hrs. After that, the media were changed into MEM with 2% FBS and incubated further for 16 hrs. These well-formed HFFs cells were infected with 10,000 cells of the GFP-expressing *T. gondii* and then treated with 5 concentrations of in the presence of various concentration of antifolates drugs in 0.1% DMSO. The fluorescence signals were detected at Excitation 485 nm and Emission 515 nm after 4 days incubation. The results were expressed as IC_{50} from the dose-response curves.

4.4 Cloning of Tgdhfrts in expression vector

Gene coding *Tgdhfrts* was subcloned from pL0017 (Chris Janse and Andy Waters, Leiden University Medical Center, Netherlands) to a modified pET17b vector and expressed in *E. coli* BL21 (DE3) under T7 promoter (Thai petty patent No. 13652). In this study, we managed to remove the non-

essential loops of TgDHFR-TS (residues deleted: 50–72 and 202–218, located on loop 1 and loop 3) by PCR using 5-phosphorylated primers: TgLoop1-For: 5'- AG ATT CAA CGC CGT TGT CAT GGG A -3' and TgLoop1-Rev: 5'- CA GGC GAC TGG CTT CTT CGG -3' for residue 50–72 deletion, TgLoop3-For: 5'- AG AGA GAA GGA CAA TGA AGC GAC G -3' and TgLoop3-Rev: 5'- CT GCG CAG CAG TTG ATT TGT TTG AAA -3' for residue 202–218 deletion. The PCR products were ligated and transformed into *E. coli* DH5α and plated on LB agar plates containing 100 µg mL^{-1} ampicillin. Well separated colonies were picked and cultured. DNA sequencing has confirmed the success in construction of recombinant plasmid.

4.5 Expression and purification

E. coli BL21(DE3) carrying pET17b-Tgdhfr-ts were grown in LB at 37°C until OD$_{600}$ =1 and then induced by 0.4 mM IPTG at 20°C for 20 hrs. Pellet cell was harvested by centrifugation and resuspended in lysis buffer (20 mM potassium phosphate buffer, pH 7.0, 50 mM KCl, 0.1 mM EDTA, 10 mM DTT, and 20% (v/v) glycerol). Following steps of cell lysis by French Pressure Cell Press and centrifugation at 27,000×g for 1 hour, supernatant containing soluble form of TgDHFR-TS was obtained. The enzyme was purified using methotrexate sepharose chromatography and Q-sepharose chromatography as previously described (Chitnumsub *et al.*, 2004). Eluted enzyme was concentrated using ultrafiltration (Amicon 50 kDa MWCO). The concentrated enzyme (>10 mg mL^{-1}) was kept at −80°C in the lysis buffer.

4.6 Kinetic properties analysis and enzyme inhibition assay

DHFR activity was determined spectrophotometrically by measuring the rate of NADPH depletion at 340 nm using ε_{340} of 12300 M^{-1}cm^{-1}. Kinetics studies were performed using 6–10 mU of purified TgDHFR-TS enzyme in the reaction (200 µL) of 1×DHFR buffer (50 mM TES, pH 5.0, 75 mM ß-mercaptoethanol and 1 mg mL^{-1} BSA) containing 50 µM each of DHF and NADPH. Enzyme inhibition assay was performed under the same condition in the presence of varying concentrations of antifolates, the *IC$_{50}$* value was read from dose-response curve of DHFR inhibition.

4.7 Protein crystallization

The purified protein in 100 mM KP buffer, pH 7.0 (10 mg mL^{-1}) was mixed with 2.5 mM each of NADPH, dUMP and inhibitor, then incubated on ice for 1 hour. Initially, micro-batch technique was performed for crystal screening by adding 1 µL of varying crystallization solution to 1 µL of the enzyme-ligand mixture under oil on a 60-well plate (well diameter 1 mm). When suitable crystallization solutions were identified. Hanging drops technique was sequentially performed to improve size and quality of the crystals. Optimization was done by varying concentration of crystallization solution (type of buffer, salt and precipitant) and ratio of protein-ligand mixture to crystallization solution. Crystallization plates were incubated at 15, 24 and 30°C and regularly monitored under microscope.

4.8 Data collection and structure determination

Single crystal was transferred in a cryo-protectant (crystallization solution plus 20% glycerol) and flash-frozen in liquid nitrogen or taken directly from mother liquor and flash-cooled. Data were collected at beamline BL13B1 at NSRRC (Taiwan, ROC.) with ADSC Q315r CCD. Data were processed using HKL-2000 (Otwinowski & Minor, 1997).

Crystal structures were determined by molecular replacement with PHASER (McCoy *et al.*, 2007) or MOLREP (Vagin & Teplyakov, 2010) in the CCP4 suite (Winn *et al.*, 2011) using structure PDB ID 4EIL as a search model. The model was built using program COOT (Emsley *et al.*, 2010) and refined by REFMAC (Murshudov *et al.*, 2011) in CCP4. Finally, models were validated using RAMPAGE (Lovell *et al.*, 2003) as listed in **Table 2**. Figures were prepared using PyMOL (Schrödinger, 2010).

Table 2. Data collection and model refinement statistics.

	PYR	P39	P40	P65	P218
Condition	LN1/4K+Ad69 (2%BZ)	LN1/4K+Ad43	LN1/4K+2.5% BZ	LN1/4K+100mMGHCl	LN1/4K+SB32 (1:2:1)
Days	40	10	19	6	15
Data collection					
Space group	P_1	P_1	P_1	P_1	P_1
Molecules/ASU	4	4	4	8	4
Solvent content (%)	55.74	54.46	54.35	54.99	53.90
V_m ($Å^3$/Da)	2.78	2.70	2.69	2.73	2.67
Cell dimensions					
a, b, c (Å)	71.282, 101.121, 107.208	68.404, 102.536, 106.198	70.125, 100.713, 106.036	107.148, 117.778, 118.809	69.168, 93.339, 116.989
α, β, γ (deg)	85.447, 83.702, 84.254	89.950, 87.737, 84.689	85.540, 83.518, 84.673	89.990, 90.027, 89.941	83.705, 80.672, 80.246
Resolution (Å)	30–2.15 (2.23–2.15)	30–2.05 (2.12–2.05)	50–2.7 (2.8–2.7)	50–2.6 (2.69–2.60)	50–2.5 (2.59–2.50)
Completeness (%)	96.3 (88.3)	95.6 (92.3)	96.6 (94.5)	94.6 (94.9)	94.2 (94.3)
Redundancy	2.2 (2.0)	2.1 (2.0)	3.1 (3.0)	3.0 (2.9)	2.1 (2.1)
Total number of reflections	336,393	350,677	237,708	517,071	195,753
Unique reflections	152,902	170,447	77,348	173,039	91,928
R_{sym}(%)	3.1 (24.3)	4.8 (40.7)	7.0 (39.2)	6.4 (42.7)	5.5 (31.6)
I/σ	25.3 (3.0)	16.4 (2.2)	14.3 (2.6)	15.3 (2.8)	13.1 (2.3)
Refinement					
R_{work}/R_{free}	17.92 (22.37)	18.57 (22.53)	19.72 (26.77)	20.54 (28.02)	19.77 (25.57)
R.m.s. deviations					
Bond length (Å)	0.0099	0.0092	0.0093	0.0127	0.0084
Bond angle (deg)	1.4399	1.3915	1.4238	1.6926	1.3451

Table 2. Data collection and model refinement statistics. (continued)

	B3	B126	B160	B253	MTX
Condition	LN1/4K+100mMGHCl	LN1/4K+2.5% BZ	LN1/4K+100mMGHCl	LN1/4K+2% BZ	LN1/4K+Ad43
Days	8	7	7	15	42
Data collection					
Space group	P_1	P_1	P_1	P_1	P_1
Molecules/ASU	4	8	4	4	4
Solvent content (%)	55.50	55.31	55.50	58.06	54.61
V_m (Å^3/Da)	2.76	2.75	2.76	2.93	2.71
Cell dimensions					
a, b, c (Å)	71.276, 101.015, 106.730	106.915, 119.831, 117.868	71.072, 100.959, 107.185	71.408, 100.619, 117.670	77.766, 101.590, 105.050
α, β, γ (deg)	85.491, 83.593, 84.484	89.955, 90.099, 90.008	85.266, 83.380, 84.308	73.211, 91.149, 84.392	101.558, 109.830, 98.569
Resolution (Å)	50–2.50 (2.59–2.50)	50–2.05 (2.12–2.05)	50–2.70 (2.80–2.70)	50–2.60 (2.69–2.60)	30–2.05 (2.12–2.05)
Completeness (%)	91.3 (92.0)	96.7 (96.3)	95.1 (94.6)	96.6 (96.9)	96.8 (95.1)
Redundancy	2.3 (2.3)	2.7 (2.6)	2.5 (2.4)	2.5 (2.5)	1.9 (1.8)
Total number of reflections	215,688	960,679	192,294	228,970	333,426
Unique reflections	91,785	352,230	77,032	91,764	173,196
R_{sym}(%)	6.4 (34.5)	5.6 (32.9)	8.6 (39.8)	5.9 (33.3)	4.4 (26.2)
I/σ	11.5 (2.0)	16.4 (3.1)	11.0 (2.8)	15.2 (3.1)	18.3 (2.2)
Refinement					
R_{work}/R_{free}	19.41 (25.94)	20.67 (24.89)	19.88 (27.38)	19.77 (26.27)	19.71 (24.32)
R.m.s. deviations					
Bond length (Å)	0.0106	0.0120	0.0127	0.0116	0.0151
Bond angle (deg)	1.5466	1.6136	1.7538	1.5592	1.7811

Table 2. Data collection and model refinement statistics. (continued)

	TMP	P161	P195	B1	B5
Condition	LN1/4K+1.5% BZ	LN1/4K+100mMGHCl	LN1/4K+2.5% BZ	LN1/4K+1.5% BZ	LN1/4K+2.5% BZ
Days	10	19	7	7	8
Data collection					
Space group	P_1	P_1	P_1	P_1	P_1
Molecules/ASU	8	4	8	8	8
Solvent content (%)	55.48	55.80	55.1	55.71	54.93
V_m (Å^3/Da)	2.76	2.78	2.74	2.78	2.73
Cell dimensions					
a, b, c (Å)	107.146, 117.936, 119.955	71.275, 101.282, 107.263	106.822, 117.949, 119.305	106.771, 117.784, 129.205	106.904, 119.017, 117.680
α, β, γ (deg)	90.132, 90.029, 89.990	85.341, 83.247, 84.148	90.064, 90.087, 90.098	69.959, 86.721, 89.990	90.010, 89.985, 90.027
Resolution (Å)	30–2.85 (2.95–2.85)	50–2.45 (2.54–2.45)	50–2.30 (2.38–2.30)	50–2.35 (2.43–2.35)	50–2.6 (2.69–2.6)
Completeness (%)	96.8 (95.1)	96.3 (91.8)	92.1 (91.9)	95.7 (90.1)	95.6 (96.7)
Redundancy	2.7 (2.7)	3.0 (3.0)	2.8 (2.6)	2.9 (2.5)	2.2 (2.2)
Total number of reflections	342,864	320,255	654,418	670,604	367,668
Unique reflections	126,173	105,214	237,572	234,166	169,820
R_{sym}(%)	10.1 (44.7)	7.0 (35.2)	7.0 (37.2)	7.9 (41.4)	5.6 (25.2)
I/σ	9.5 (2.3)	14.2 (2.5)	16.4 (2.2)	12.1 (2.2)	12.6 (3.1)
Refinement					
R_{work}/R_{free}	26.98 (37.77)	22.52 (28.04)	25.33 (32.26)	30.67 (36.40)	22.39 (29.47)
R.m.s. deviations					
Bond length (Å)	0.0122	0.0112	0.013	0.012	0.0116
Bond angle (deg)	1.738	1.5715	1.7532	1.6704	1.6333

Values in parentheses are for the highest resolution shells.[b]$R_{sym} = \Sigma_{hkl}\Sigma_i|I_i(hkl) - \langle I(hkl)\rangle|/\Sigma_{hkl}\Sigma_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of an individual reflection and $\langle I(hkl)\rangle$ is the mean intensity of symmetry-equivalent reflections. [c]$R_{work} = \Sigma_{hkl}||F_{obs}| - |F_{calc}||/\Sigma_{hkl}|F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure-factor amplitudes, respectively. R_{free} was calculated in the same manner as R_{work} but using only a 10%unrefined subset of the reflection data.

5. Results

5.1 To establish a bacterial surrogate model for anti-TgDHFR screening

5.1.1 Enzyme inhibition of antifolates against TgDHFR

Before establishing a surrogate model, a set of compounds classified as rigid (Pyr and Cyc), semi-rigid (Tmp) and flexible (P65, P218 and WR99210) antifolates has been tested first by enzyme inhibition assay. We found that Pyr is 5 times more effective against TgDHFR than Cyc and 200 times better than Tmp (**Table 3**). P65 and WR99210 slightly more effective than Pyr. Intersetingly, P218 is the most effective one with IC_{50} value 4 folds lower than that of Pyr (0.059 vs 0.23 uM). These results suggest that it is promising to explore flexible antifolate library against TgDHFR.

Table 3. Inhibitory effect of antifolates using TgDHFR-TS enzyme assay

Antifolates		TgDHFR inhibition assay IC_{50} (µM)
Rigid type	Pyr	0.23 ± 0.03
	Cyc	1.49 ± 0.18
Semi-rigid type	Tmp	50.4 ± 13.7
Flexible type	P65	0.125 ± 0.02
	P218	0.059 ± 0.0008
	WR99210	0.163 ± 0.01

5.2 Optimization of bacterial surrogate model

5.2.1 Condition for bacterial growth

The TgDKO cell line has been used for the first time as a model to screen compounds against TgDHFR-TS. This bacterial line is *E. coli* BL21 (DE3) whose *thyA* and *folA* genes have been knocked out and the genes function is complemented by TgDHFR-TS from a TgDHFR-TS expressing plasmid. The cell line can survive without thymidine supplement in the media.

To find out which culture media is suitable for use in antifolate screening against TgDHFR, three types of culture media (Luria bertani (LB), minimal media (MM) plus 0.2% casamino acid (CA) and MM plus 0.4% CA) were tested. The growth rate of the TgDKO cells in LB broth and MM+0.4% CA was about the same but poorer in MM+0.2% CA. The TgDKO cells grew quickly in LB broth and reached log phase at 2-4 hours with a maximum OD600 of 0.8-0.9 after 5 hours of incubation (**Figure2A**). The generation, or doubling time of TgDKO cells in these culture media were different; the best is LB (LB < MM+0.4% CA < MM+0.2% CA) (**Figure2B**). For drug testing, it is essential to allow enough exposure time of the cells to inhibitors. With this concept, MM+CA culture media is

more suitable than LB culture media and that MM+0.4% CA culture media is more favorable than MM+0.2% CA culture media due to higher OD600 signal at 4-5 hours incubation.

5.2.2 Bacterial growth inhibition assay

For assay optimization, the cell density following incubation with various concentrations of antifolate inhibitors was measured at the mid-late log phase after incubation for 4–6 hours. The IC_{50} values of antifolates obtained from each condition were as shown in **Table 4**. The IC_{50} values of six selected compounds determined at 4, 5 and 6 hours after incubation were in the same ranges. According to **Figure 2**, the OD_{600} of the control untreated cells after 4– and 5–hours incubation is within the linear ranges (OD_{600} = 0.30–0.55) while the cells may have reached its plateau at 6 hours incubation (OD_{600} = 0.55–0.70). Therefore, 5–hours incubation is selected as suitable incubation time for determining IC_{50} values.

With this initial exploration, we found that WR99210 and P65 are more active against TgDKO cells than Pyr while Cyc and Tmp are inactive (**Table 4**) Although the majority of the results are in good agreement with enzyme inhibition assay, the IC_{50} –values from TgDKO assay are unacceptably high. Furthermore, P218, which is the most active by enzyme inhibition assay become inactive by TgDKO assay. An alternative better surrogate host is needed. Nonetheless, the results suggested that flexible antifolates is promising for anti-Tg development and it is worth to further explore flexible antifolates in the library.

(A)

(B)

Time interval	Generation time or doubling time* (min)		
	LB	MM + 0.2% CA	MM + 0.4% CA
1-2 hr	32.3	42.2	36.6
2-3 hr	29.4	38.6	36.8
3-4 hr	75.7	43.5	43.5
4-5 hr	433.9	82.8	74.9

*G = t / n = t / 3.3 log b/B

G = generation time

t = time interval in minutes or hours

n = number of generations

B = number of bacteria at the beginning of a time interval

b = number of bacteria at the end of a time interval

OD600 = 1 =1x10^9 cell/ml
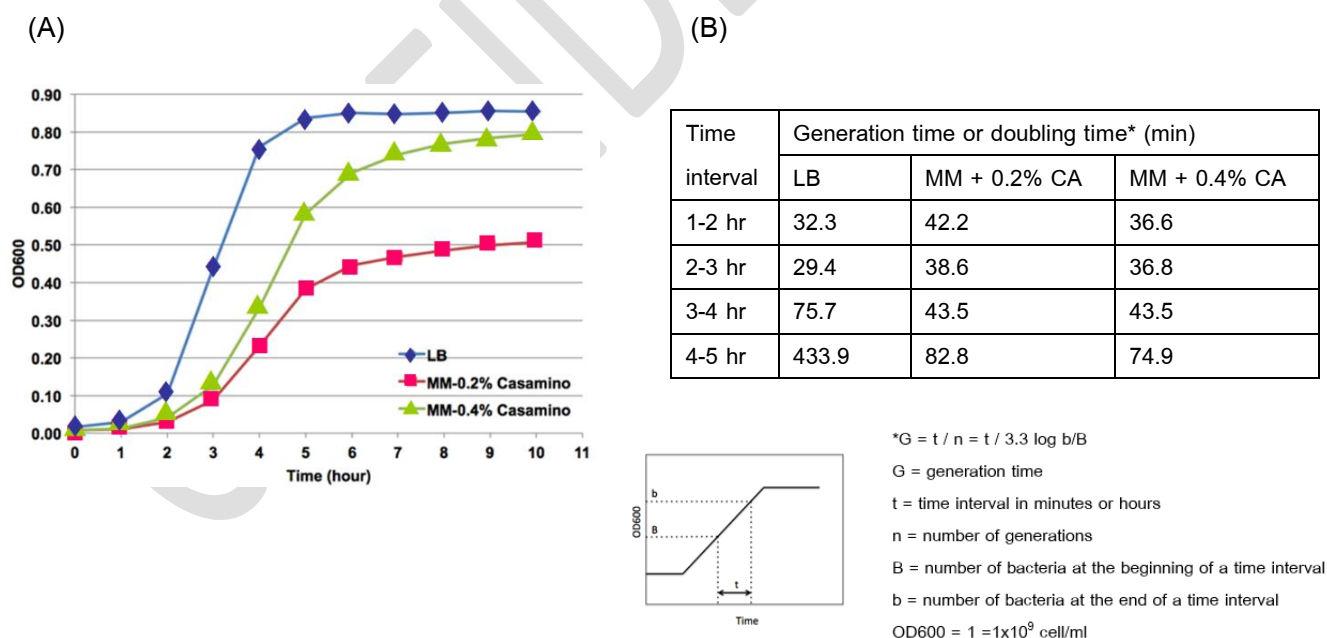
Figure 2. Growth curve and generation time of TgDKO in different culture media.

Table 4. IC$_{50}$ values calculated from TgDKO cell growth at three different time points and from enzyme inhibition assay

Incubation time	IC$_{50}$ values (µM) against TgDKO cell at each incubation time		
	4 hrs	5 hrs	6 hrs
Rigid type			
Pyr	153.6 ± 15.6	179.1 ± 26.5	238.0 ± 21.7
Cyc	>500	>500	>500
Semi-rigid type			
Tmp	>500	>500	>500
Flexible type			
P65	36.9 ± 1.02	38.7 ± 1.99	40.6 ± 4.38
P218	>500	>500	>500
WR99210	6.0 ± 1.37	6.72 ± 0.76	7.77 ± 1.38
OD$_{600}$ range of control	0.30–0.45	0.45–0.55	0.55–0.70

*TgDKO was grown in MM plus 0.4% CA. OD$_{600}$ was measured at 4, 5, and 6 hours after incubation with antifolates. The IC$_{50}$ value is displayed mean ± standard deviation from at least 3 experiments.

5.2 Effect of efflux pump inhibitor on growth inhibitory assay

To address the high IC$_{50}$ values and the discrepancy results of P218 obtained from the bacterial TgDKO model and that from the enzyme inhibition assay, we hypothesized that P218 may possibly be pumped out from the cells into the medium by the action of an efflux pump. To test this hypothesis, effect of efflux pump inhibitors (EPIs) on P218 susceptibility of TgDKO were explored. Five EPIs, phenylalanine arginine ß-naphthylamide (PAßN), carbonyl cyanide *m*-chlorophenyl hydrazine (CCCP), ethylenediaminetetraacetic acid (EDTA), probenecid and verapamil, were employed for the test. These inhibitors have been reported to enhance the drugs action of antibiotics (Gupta *et al.*, 2014, Lamers *et al.*, 2013). PAßN is peptidomimetric acting as a competitive inhibitor specifically for resistance-nodulation-division (RND) family efflux pumps. CCCP is an energy uncoupler and acts to dissipate the proton motive force. EDTA, a chelator of divalent cations such as Mg^{2+}, is important to membrane transport. Probenecid is an inhibitor of anion membrane transport. Verapamil is an inhibitor of drug efflux pump proteins such as P-glycoprotein and its mechanism is to block voltage-dependent calcium channels. We found that EPIs inhibit TgDKO growth to varying degrees (**Table 5**). CCCP and EDTA are moderately inhibitory with IC$_{50}$ 25–57 µM, while PAßN and probenecid and verapamil are inactive at 10 mM.

To explore the effect of EPIs on antifolate susceptibility of TgDKO, the final concentration of the EPIs used were 5 µM of CCCP, 25 µM of EDTA, 1 mM of PAßN, probenecid and verapamil. As shown in **Table 6**, only two EPIs, CCCP and PAßN were found to enhance the inhibitory effect of antifolates by 2 – 15 folds and by 6 – 144 folds, respectively. EDTA, probenecid and verapamil did not cause much change (less than 2–fold). These suggest that RND family efflux pumps and energy from the proton motive force may be involved in increased susceptibility of TgDKO to antifolates while divalent cation, anion membrane transport or voltage-dependent calcium channels may not. Indeed, the TolC RND efflux system has been reported to modulate methotrexate-resistance in *E. coli* (Kopytek et al., 2000). Therefore, a *tolC*-knockout cell was used to confirm the finding.

To confirm the involvement of *TolC* RND efflux system in antifolate susceptibility of TgDKO, PEcΔTFTolC, a PEcΔTF strain whose *TolC* gene was knocked out, was transformed with TgDHFR-TS expressing plasmid and the transformance (TgDKO-TolC) was tested for its antifolate susceptibility. As expected, we found that TgDKO-TolC showed 6–50 fold more susceptible to antifolates (Pyr, P218, P195, P65 and P39) than TgDKO (**Table 6**). Therefore, these data confirm that TolC-dependent efflux pump plays an important role in antifolate susceptibility of PEcΔTF-TgDT and in line with previous report on TolC-dependent methotrexate resistant *E. coli* and in concordance with the PAßN results above.

It should be note that while TgDKO-TolC could reduce the IC_{50}, However, the IC_{50} value of P218 by TgDKO-TolC is higher than the other compounds while it is the most potent one by enzyme inhibition assay. This could be complicated by the presence of carboxylic group in P218. Its ester analogs, P195 are still high, as compared to P218. These information prompt us to employ TgDKO-TolC as a new surrogate model for evaluation of antifolate potency against TgDHFR-TS.

While P218 is the most effective inhibitor among the new antifolates by enzyme inhibition assay, it ranks almost the last by TgDKO-TolC assay. P195, an ester derivative of P218, is 10-fold more active than P218 by TgDKO-TolC assay but 5-fold less active by enzyme inhibition assay (**Table** 6). P65 (flexible) and P39 (rigid) with hydrophobic side chain also showed similar discrepancy as compared to P218. The term cLogP, a computed octanal-water partition coefficient, was taken into consideration. In general, drug-like compounds with high permeability are characterized by an optimum cLogP between 3 and 5 (Leeson & Springthorpe, 2007). As listed in **Table 6**, the cLogP value of P218 calculated using ChemDraw software is markedly low as compared with that of P195 (2.333 vs 4.234, respectively). Analysis of relationship between lipophilic or LogP value and the IC_{50} value of compounds by bacterial assay from a larger data set should give a clear and conclusive results (see section 5.3).

Table 5. Inhibitory effect of efflux pump inhibitors on the growth of TgDKO.

Efflux pump inhibitor	IC$_{50}$ (μM)
CCCP	24.7 ± 1.5
EDTA	56.5 ± 6.0
PAßN	>10000
Probenecid	>10000
Verapamil	>10000

Table 6. Effect of EPIs on efficiency of antifolate against TgDHFR-TS via bacterial model.

Cpd	cLogP*	IC$_{50}$ (μM)							Enzyme assay
		TgDKO with:							
		No EPI	5 μM CCCP[a]	25 μM EDTA[a]	1 mM PAßN[a]	1 mM Probenecid[a]	1 mM Verapamil[a]	ΔtolC, no EPI	
PYR	3.003	225.5 ± 48.6	49.8 ± 7.6	306.7 ± 62.8	35.0 ± 6.1	184.0 ± 6.3	176.4 ± 18.6	38.1 ± 3.3	0.23 ± 0.03
P218	2.333	>1000	>1000	>1000	628.9 ± 55.3	>1000	>1000	32.5 ± 4.9	0.059 ± 0.001
P195	4.234	175.5 ± 40.5	11.3 ± 2.3	157.2 ± 44.9	1.22 ± 0.27	171.7 ± 32.2	143.7 ± 13.5	3.93 ± 0.57	0.305 ± 0.033
P65	5.034	146.6 ± 39.7	43.6 ± 5.8	217.0 ± 55.7	3.68 ± 0.77	67.7 ± 6.7	83.5 ± 10.0	5.75 ± 1.5	0.125 ± 0.022
P39	4.396	72.0 ± 8.4	29.3 ± 7.7	79.4 ± 14.2	0.60 ± 0.06	48.2 ± 7.4	56.2 ± 16.4	1.45 ± 0.22	0.076 ± 0.002

[a]The concentration of CCCP, PAßN and probenecid inhibits 5% growth, while EDTA and verapamil inhibit 25% growth, approximately. *Calculated by ChemDraw software

5.3 IC$_{50}$ determination using bacterial and enzyme assay

To get more conclusive results for the interference of TolC function on antifolate screening assay, we tested 71 compounds against both TgDKO and TgDKO-TolC cells. We found that Among 71 compounds tested, 46 compounds testing by TgDKO-TolC assay tend to have lower IC$_{50}$ value and more active than testing by TgDKO assay. For example, 25 compounds showed no growth inhibition at 500 µM by TgDKO assay while 17 of these are active (IC$_{50}$ = 0.1 – 300 µM) by TgDKO-TolC assay. However, only five compounds, P25, P36, P37, P38 and P62 with similarity in chemical structure, are more active to TgDKO than TgDKO-TolC cells (**Figure 3**). Interestingly, cLogP values of these 5 compounds are less than 3 (cLogP = 1.05–2.47). This would indicate that not only TolC function interferes bacterial assay, properties of compounds such as lipophilicity and permeability should be concerned.

In testing our compounds libraries, we selected a total of 218 compounds including pyrimidine, dihydrotriazine and quinazoline core structure with rigid, semi-rigid and flexible type substituent at S2-S3 subsite (Table 1) and structure details in **Appendix 8.1**). All compounds were tested by TgDKO-TolC cell assay and 149 compounds of which were also tested by TgDHFR-TS enzyme inhibition assay. Distribution of compounds according to their structures and potencies against both assays were summarized in **Table 7**. Compounds in Pyr-Rigid, Pyr-Flexible and Cyc-Flexible groups show wide range of IC$_{50}$-values against TgDKO-TolC cells from less than 1 µM to more than 100 µM while Quinazolines within 1-100 µM ranges. Among these, 24 compounds (3 in Pyr-Rigid, 17 in Pyr-Flexible and 4 in Cyc-Flexible) are the most active with IC$_{50}$-values lower than 1 µM against both TgDKO-TolC cells and TgDHFR-TS enzyme. However, there is no correlation between the two measurement assays for IC$_{50}$ determination of 133 compounds (**Figure 4**). Sixteen of 149 compounds were not in the study because some compounds were inactive in bacterial assay (IC$_{50}$>10µM) and some compounds were inactive in enzyme assay (IC$_{50}$>10µM), possibly from their solubility.

To establish clearer understanding on the effect of cLogP on TgDKO-TolC assay, CLogP values of 133 compounds were estimated for analysis. The calculation was based on two software tools, ChemDraw Ultra8.0 (CambridgeSoft) and MarvinSketch version 16.7.4.0 (ChemAxon Ltd.). The results obtained from both software are slightly difference. Nonetheless, the scatter plotted between cLogP calculated from 2 software tools versus IC$_{50}$ values from TgDKO-TolC assay show the same tendency where compounds with low cLogP tend to have high IC$_{50}$ or inactive and vice versa for

those with high cLogP (**Figure 5A**). We, therefore, selected cLogP values calculated from ChemDraw for further analysis for all 218 compounds. Of which, 21 compounds (3 of dihydrotriazines and 18 of pyrimidines) are inactive (IC_{50} > 100 µM) and 15 of them have cLogP lower than 2. For the rest of the compounds (197 compounds), IC_{50} determined by TgDKO-TolC were plotted versus cLogP (**Figure 5B**). According to core structure scaffolds, we found that most of compounds in dihydrotriazine series, except WR99210, exhibit cLogP lower than 1.2 and all of them are inactive (IC_{50} > 100 µM). From **Figure 5B**, we can set criteria of clogP by less than 2 for inactive compounds (IC_{50} > 50 µM for enzyme assay and IC_{50} > 100 µM TgDKO-TolC assay).

To validate correctness of data determined by TgDKO-TolC, we set criteria of cLogP and IC_{50} determined by both enzyme and bacterial inhibition assays for active, moderate, and inactive compounds as shown in **Table 8**. We classify the data as either active or moderate or inactive compounds, which IC_{50} values determined by both enzyme and bacterial assays are in the same zone. We found that correctness is approximately 14% for compounds with cLogP<2, while it increases to 62% for compounds having cLogP>3 (**Figure 6A**). The correlation of IC_{50} determined by both assays of corrected data is more than 0.6 for each group of cLogP (**Figure 6B–6D**). We can conclude that the TgDKO-TolC surrogate system is suitable for testing compounds having cLogP greater than 3.

5.4 *In vitro T. gondii* testing

We have successfully developed cell culture assay by monitoring green fluorescence in *T. gondii*. IC_{50} value of Pyr against *T. gondii*-GFP cell assay from our development is comparable to the value from standard method (0.58 ± 0.04 µM) using uracil uptake assay (Reynolds & Roos, 1998). Some of potent compounds identified from enzyme inhibition and/or TgDKO-TolC growth assay with IC_{50} lower than Pyr were tested against *T. gondii* cell assay (**Table 9**). Compounds B001 and B004, containing quinoline and carboxylic moiety at S3 subsite, are the most active compounds against *T. gondii* cell assay with IC_{50} < 5 nM, while their ester analogs (B002 and B003) are the most active compound against TgDHFR-TS enzyme. Similarly, many of pyrimidine-flexible compounds having carboxylic group at S3 subsite including P218, B001, and B003 are active against TgDHFR-TS enzyme (IC_{50} < 0.15 µM) and more active than their ester analogs (P195, B002, and B004 respectively). In contrast, the ester compounds tested by TgDKO-TolC assay are more active than the carboxylic compounds by 8–880 folds IC_{50} reduction. Three pair of compounds (P195/P218,

B002/B001, and B004/B003) in the ester/carboxylic series were tested against *T. gondii* cell assay. Results show that three ester compounds (P195, B004, and B250) are more active than the carboxylic compounds (P218, B003, and 251), while the IC_{50} values of two sets (B002/B001 and B252/B253) are comparable.
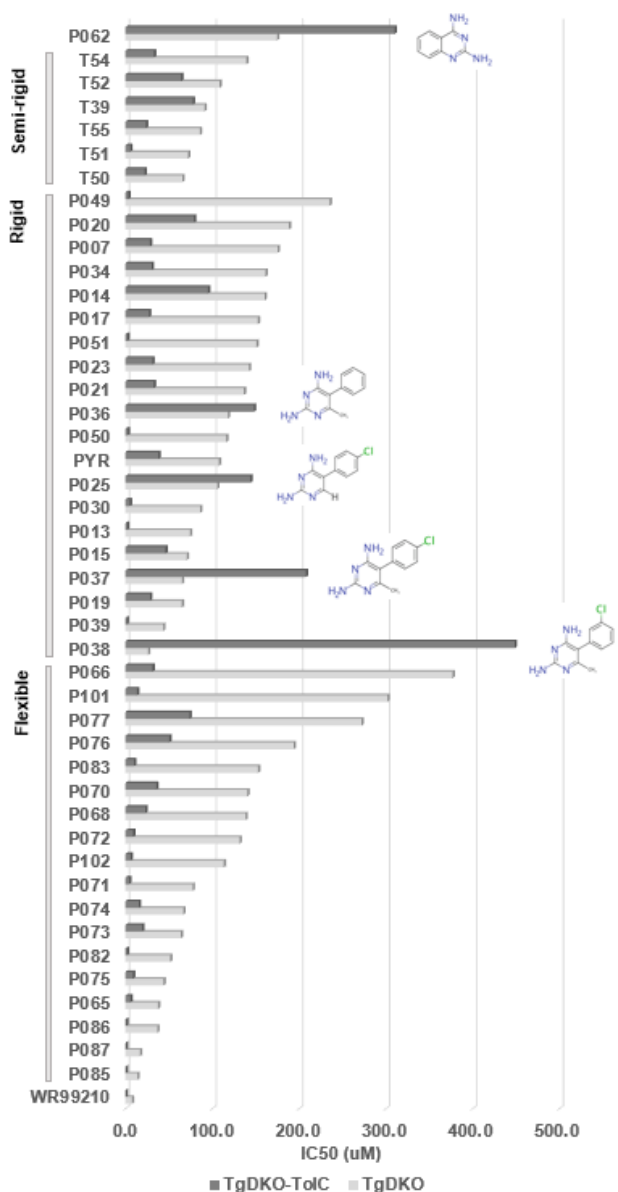
Figure 3. Comparison IC_{50} values of 46 compounds evaluated by TgDKO and TgDKO-TolC bacterial assay.

Five of 46 compounds, P25, P36, P37, P38 and P62, testing by TgDKO-TolC assay show IC50 value higher than testing by TgDKO.

Table 7. Distribution of antifolate compounds according to their structure and potency.

Type of compound	No. of compound tested	No. of compound in each of IC$_{50}$ range against TgDKO-TolC			
		0.01–1 µM	1–10 µM	10–100 µM	>100 µM
Pyr-Rigid	42 (32)	3 (21)	10 (7)	15 (1)	14 (3)
Pyr-Flexible	85 (54)	17 (52)	45 (2)	17 (–)	6 (–)
TMP-Semi-rigid	55 (27)	– (3)	3 (15)	25 (7)	27 (2)
Cyc-Rigid	3 (3)	– (3)	– (–)	– (–)	3 (–)
Cyc-Flexible	15 (15)	4 (12)	2 (3)	– (–)	9 (–)
Quinazoline	18 (18)	– (11)	8 (2)	9 (–)	1 (5)
Total	218 (149)	24 (102)	68	68 (–)	60 (10)

*Parentheses are number of compounds testing using both bacteria TgDKO-TolC and TgDHFR-TS enzyme assay.
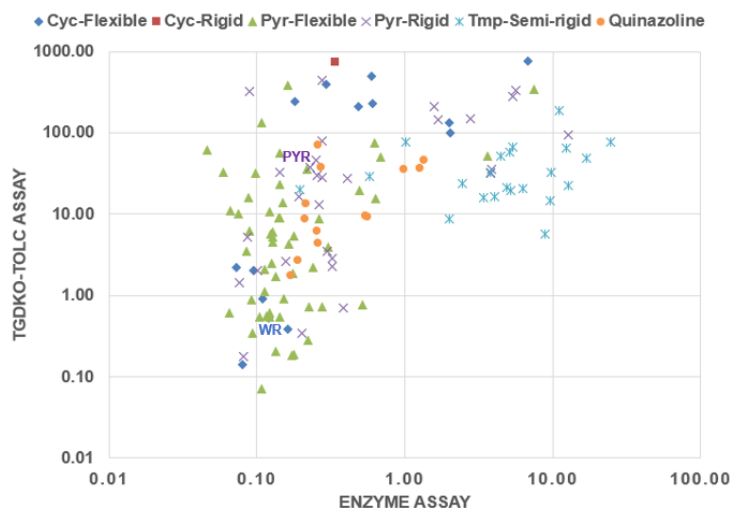
Figure 4. Scatter plots of IC$_{50}$-values of 133 compounds tested against TgDHFR-TS enzyme and TgDKO-TolC assay.

(A)
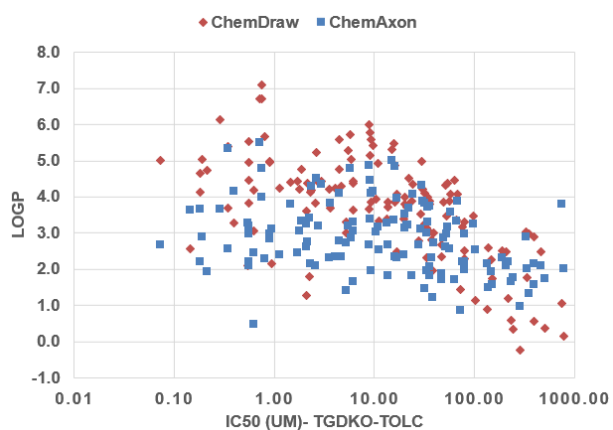
(B)
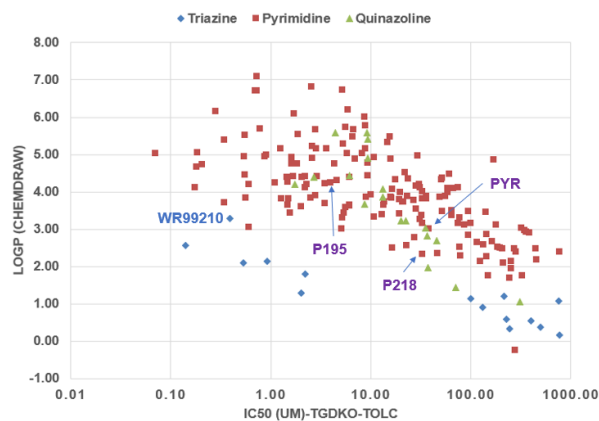
Figure 5. Correlation of IC$_{50}$ values and LogP values.

LogP values of 133 compounds were calculated by two software ChemDraw Ultra8.0 and MarvinSketch version 16.7.4.0 (ChemAxon Ltd.), which are plotted versus IC$_{50}$ values TgDKO-TolC assay (A). (B) Correlation of cLogP (from ChemDraw) and IC$_{50}$ values from TgDKO-TolC assay.

Table 8. Criteria for identification compounds testing by enzyme and bacterial assay.

Criteria	Active	Moderate	Inactive
cLogP (ChemDraw)	>3	2–3	<2
IC$_{50}$ -Enzyme assay	<1 μM	1–50 μM	>50 μM
IC$_{50}$ -TgDKO-TolC assay	<10 μM	10–100 μM	>100 μM
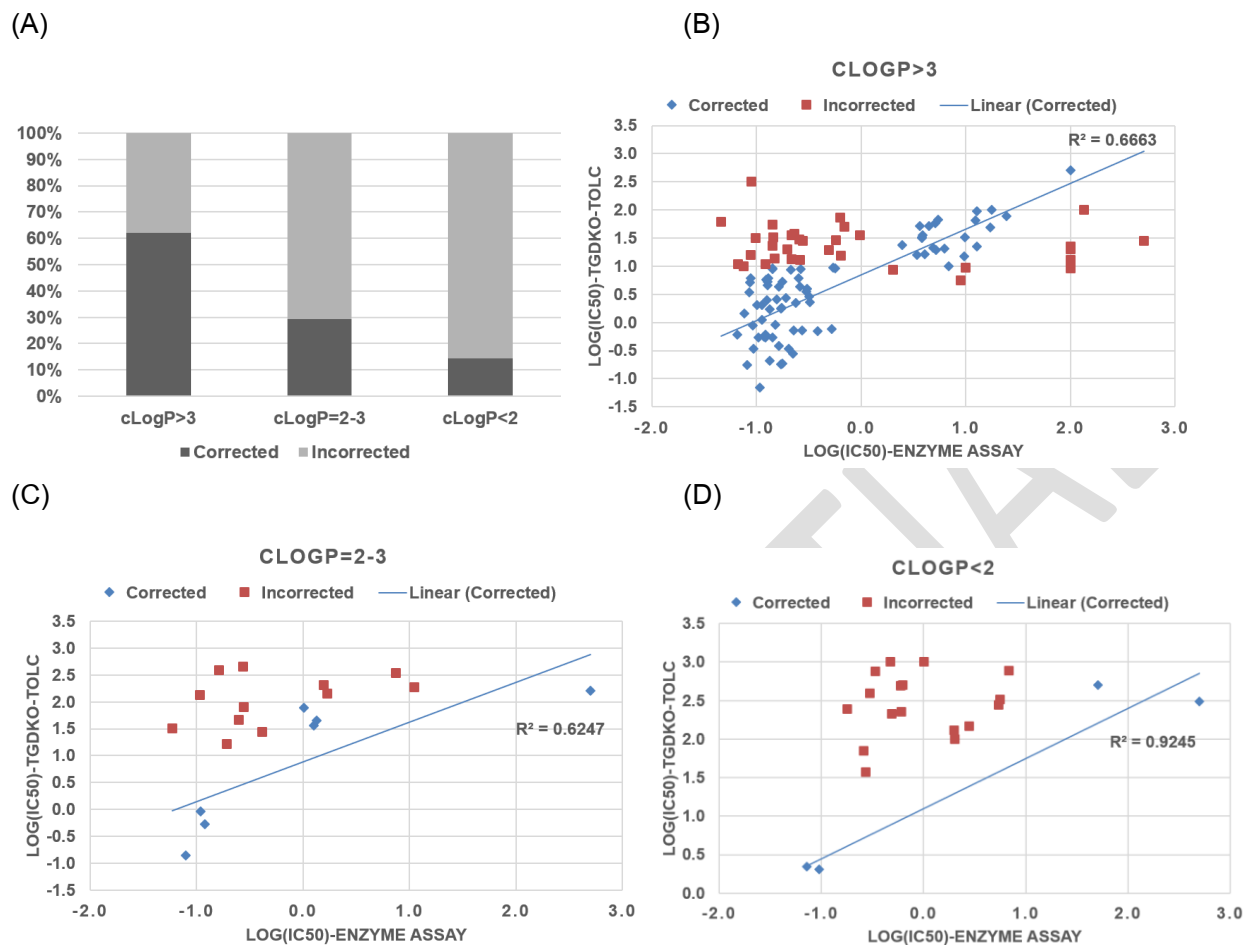
Figure 6. Effect of cLogP values of compounds on correctness of IC$_{50}$ data from TgDKO-TolC assay.

(A) Percentage of corrected and incorrected data according to cLogP values. Correlation of IC50 values between two different assays of compounds with cLogP>3 (B), cLogP=2–3 (C), and cLogP<2 (D).

Table 9. Inhibitory data of good potency compounds against TgDHFR-TS enzyme assay, TgDKO-TolC bacterial cell assay and *T. gondii* cell assay.

| Scaffolds/series	Code	IC$_{50}$ (μM)		
		Enzyme assay	TgDKO-TolC	*T. gondii*
Pyrimidine-Rigid	Pyr	0.23 ± 0.03	38.1 ± 3.3	0.41 ± 0.005
	P08	0.39 ± 0.018	0.705 ± 0.043	ND
	P30	0.087 ± 0.006	5.14 ± 0.55	ND
	P39	0.076 ± 0.002	1.45 ± 0.22	0.03 ± 0.004
	P40	0.081 ± 0.003	0.176 ± 0.018	0.03 ± 0.001
	B15	0.203 ± 0.085	0.34 ± 0.05	ND
Pyrimidine-Flexible	P218 (carboxylic)	0.059 ± 0.001	32.5 ± 4.9	2.19 ± 0.13
	P195 (ester)	0.305 ± 0.033	3.93 ± 0.57	0.04 ± 0.002
	P65	0.125 ± 0.022	5.75 ± 1.5	0.37 ± 0.01
	B001 (carboxylic)	0.148 ± 0.023	13.6 ± 3.9	0.002 ± 0.0002
	B002 (ester)	0.172 ± 0.013	0.181 ± 0.05	0.004 ± 0.0002
	B003 (carboxylic)	0.046 ± 0.004	61.2 ± 3.8	0.53 ± 0.11
	B004 (ester)	0.108 ± 0.039	0.070 ± 0.007	0.0006 ± 0.0001
Pyrimidine-Semi-rigid	TMP	50.4 ± 13.7	>500	27.9 ± 1.2
	T44	>10	9.39 ± 2.57	ND
	T51	8.95 ± 0.44	5.63 ± 0.26	ND
	T57	0.196 ± 0.02	19.7 ± 2.1	ND
	T58	0.578 ± 0.043	29.2 ± 3.9	ND
	T69	2.01 ± 0.53	8.60 ± 1.96	ND
Dihydrotriazine	Cyc	0.623 ± 0.084	>500	ND
	WR99210	0.163 ± 0.014	0.39 ± 0.11	0.03 ± 0.001
	C420	0.110 ± 0.035	0.91 ± 0.17	ND
	C448	0.080 ± 0.011	0.14 ± 0.01	ND
	C452	0.120 ± 0.007	0.53 ± 0.12	ND
Quinazoline	Q06	0.259 ± 0.030	4.37 ± 0.75	ND
	Q11	0.170 ± 0.007	1.74 ± 0.2	ND
	Q14	0.191 ± 0.019	2.70 ± 0.77	ND

ND: Not determined. The IC$_{50}$-values are shown as mean ± standard deviation from at least 3 independent experiments.

5.5 Purification and characterization

Crystal structure of full-length bifunctional TgDHFR-TS has been reported at 3.5 Å resolution and the resolution of structure has improved to 2.2 Å by removing two non-essential surface loops in the DHFR domain (Sharma *et al.*, 2013). We have therefore adopted this approach for our study. Truncation of the non-essential loops of TgDHFR-TS (residues 50–72 and 202–218 on loop 1 and 3) was generated and purified by affinity and anion exchanger chromatography. Purification and characterization of both full-length and truncated versions are summarized in **Table 10**. Yield and percent recovery of the purified truncated version was 2.6- and 8- folds lower than that of the purified full-length version. The purified truncated enzyme catalyzes the reduction of dihydrofolate to tetrahydrofolate with a turnover rate (k_{cat}) of 1.38 sec^{-1}, which is about 2.4-fold lower than that of the full-length version (k_{cat} = 3.3 sec^{-1}).

Table 10. Purification and characterization of TgDHFR-TS.

Constructs	Enzyme/ Step	Total volume (mL)	[Protein] (mg.mL^{-1})	Total DHFR activity (units)	Specific activity (U.mg^{-1})	k_{cat} (s^{-1})	Purification (fold)	Recovery (%)	Yield (mg/L culture)
Full-length TgDHFR-TS[a]	Crude	100	20	900	0.45	-	(1)	(100)	-
	Purified	2.3	35.2	235	2.91	3.30	6.5	26.1	13.5
Truncated TgDHFR-TS[b]	Crude	225	12.4	2483	0.89	-	(1)	(100)	-
	Purified	1	62.3	81	1.30	1.38	1.5	3.2	5.2

[a]Prepared from 6 L of culture. [b]Prepared from 12 L of culture.

5.6 Crystallization of TgDHFR-TS

5.6.1 Crystallization of truncated TgDHFR-TS

According to previous publication on first crystal structure of TgDHFR-TS (Sharma *et al.*, 2013), we cannot crystallize truncated version with their conditions. Therefore, we screened to find optimal condition for crystallization. Microbatch technique was performed with a ratio of 1:1 of the mixture of protein complex and crystallizing solution. For initial crystallization screening, approximately 628 crystallizing solutions comprised of various salts, precipitants and buffers with pH ranging from 4.0-9.5 were used. The crystallizing solutions employed are HR2-110, HR2-112, HR2-118, HR2-144, HR2-108, PR1-1 and PR1-2 from Hampton research, and JCSG and PACT from QIAGEN. Following the initial screenings, 23 solutions found to produce needle-like and tiny crystals (**Table 11**) were

further optimized using the grid screening method where series of mixtures were built from three-dimensional matrices with the three different factors. These are types of buffers precipitants and salts. Two types of buffers with pH ranging from 7.5 to 9.0 (HEPES, pH 8.0 & 8.5, and Bis-Tris propane, pH 7.5, 8.0, 8.5, and 9.0), three types of polyethylene glycol (PEG) with different average molecular weight (20% of PEG3350, PEG4000 or PEG6000) and five types of salt (ammonium acetate, ammonium chloride, lithium sulfate, potassium thiocyanate and sodium acetate) were used. The screening was also performed using various percentage of PEG (PEG400 (16-24%v/v), PEG3350 (12-20%w/v), PEG4000 (12-20%w/v), MPD (30-50%v/v)) in either HEPES pH 8.0, 8.5, 9.0, 9.5 or MES pH 7.0 and 8.0. By all attempts to crystallize the truncated protein-inhibitor complex, the crystals were too small to be picked up for X-ray diffraction.

Table 11. Crystallization solution hits for truncated TgDHFR-TS.

Solution	Salt	Buffer	Precipitant
PACT			
48		0.1 M Tris pH 8.0	20% (w/v) PEG6000
JCSG			
3	0.18 M tri-Ammonium citrate		20% (w/v) PEG3350
52	0.2 M Lithium sulfate	0.1 M Tris pH 8.5	1.26 M Ammonium sulfate
57	1.6 M Magnesium sulfate	0.1 M MES pH 6.5	
67	0.8 M Succinic acid pH 7.0		
69	2.4 M Sodium malonate pH 7.0		
70	1.1 M Sodium malonate	0.1 M HEPES pH 7.0	0.5% (v/v) Jeffamine ED-200
81	0.1 M Potassium thiocyanate		30% (w/v) PEG MME 2000
82	0.15 M Potassium bromide		30% (w/v) PEG MME 2000
86	1 M Ammonium sulfate	0.1 M Bis Tris pH 5.5	1% (w/v) PEG3350
93	0.2 M Lithium sulfate	0.1 M Bis Tris pH 5.5	25% (w/v) PEG3350
94	0.2 M Ammonium acetate	0.1 M Bis Tris pH 5.5	25% (w/v) PEG3350
96	0.2 M Ammonium acetate	0.1 M HEPES pH 7.5	45% (v/v) MPD
HR2-108			
A1	1.8 M Sodium acetate trihydrate pH 7.0	0.1 M Bis Tris propane pH 7.0	
A2	2.8 M Sodium acetate trihydrate pH 7.0	0.1 M Bis Tris propane pH 7.0	
A3	1.5 M Ammonium chloride	0.1 M Sodium acetate trihydrate pH 4.6	
A4	1.5 M Ammonium chloride	0.1 M Bis Tris propane pH 7.0	
A7	3.5 M Ammonium chloride	0.1 M Bis Tris propane pH 7.0	
HR2-128			
3		0.1 M Sodium citrate tribasic dihydrate pH 5.6	0.5 M Ammonium phosphate monobasic
4		0.1 M Tris hydrochloride pH 8.5	1.0 M Ammonium sulfate
11		0.1 M Sodium citrate tribasic dehydrate pH 5.6	0.5 M Ammonium phosphate monobasic
34		0.1 M Sodium acetate trihydrate pH 4.6	1.0 M Sodium formate
39		0.1 M HEPES sodium pH 7.5	1.0 M Ammonium sulfate 2% v/v PEG400

5.6.2 Crystallization of full-length TgDHFR-TS

With the failure to obtain high quality crystals from the truncated version of TgDHFR-TS, we decided to focus on the full-length TgDHFR-TS instead. At first, we were able to obtain crystal in microplate and 2.9 Å resolution data were collected. However, reproducibility problem and tiny crystals have still unsolved. The presence of glycerol in protein storage buffer is well known for stability and solubility of protein. At high concentration (>10%), it hinders crystal growth (Vera *et al.*, 2011, Ciccone *et al.*, 2015). Therefore, we removed glycerol in protein by exchanging buffer to 100 mM KP, pH 7.0. Microbatch and hanging drop vapor diffusion techniques were used for crystal screening set up. Crystallization screening was performed using the PEG/ION1 and PEG/ION2 from Hampton research (96 crystallizing solutions) and homemade (Tg335 and Tg6K) solutions (10 types of salts and 6 buffers in 20% w/v either PEG3350 or PEG6000). The plates were incubated at 15, 24 and 30°C. Amorphous precipitation was mostly found in the screening plates incubated at 30°C. The tiny crystals were found in 28 solutions from microbatch at 24°C (**Table 12**).

Further optimization was performed using hanging drop vapor diffusion technique in 500 μL reservoir. A series of crystallization solutions was prepared from six types of salts (LiNO$_3$, LiSO$_4$, NaNO$_3$, NaBr, KBr and KSCN) with concentration varied from 0.05 to 0.2 M and 12–22% (w/v) PEG4000 in 0.1 M buffer (Bis-Tris propane pH 7.5, Tris pH 8.0 or pH 8.5). Various concentrations of TgDHFR-TS were prepared as 5, 10, 15 and 20 mg mL^{-1}. The volume ratio of protein to crystallizing solution was 1:1, 1.5:1 and 2:1. The plates were incubated at 15°C and 24°C. We found that with the 1:1 ratio of 10 mg mL^{-1} protein to crystallizing solution, rod-shaped crystals were observed on the plate incubated at 15°C within 1 month in the solutions of 14–22% PEG4000 and 0.2 M of salt (LiNO$_3$, NaBr or KBr) in 0.1 M Bis-Tris propane pH 7.5. Plate-shaped crystals were also found at 24°C in similar crystallization solutions. Crystals have still tiny size and are difficult to pick up for diffraction screening. To improve the quality and size of the crystals, a series of 96 additives and 96 silver bullets (Hampton Research, USA) were added to the selected crystallizing solution LN1/4K (0.2 M LiNO$_3$, 14% PEG4000, 0.1 M Bis-Tris propane pH 7.5). Different crystal forms were obtained in solution LN1/4K with 38 additives as shown in **Figure 7.** We selected additive no. 43 (100 mM guanidine hydrochloride) and no. 69 (2% w/v benzamidine hydrochloride) for further optimization by varying ratio of protein:crystallizing solution and concentration of additive. We observed that single crystals were found in the solution with additive no. 69, while multiple and tiny crystal was observed in control wells without additives (**Figure 8A**). After varying concentration of benzamidine HCl, size

of crystals was improved to 50×150×450 μm^3 (**Figure 8B**). Similarly, single crystals were found in solution with 100 mM guanidine HCl after 7-day incubation at 24°C.

Table 12. Crystallization solution hits for the full-length TgDHFR-TS.

Solution	Salt	Buffer	Precipitant
PEG/Ion1			
2	0.2 M Potassium fluoride		20% (w/v) PEG3350
4	0.2 M Lithium chloride		20% (w/v) PEG3350
10	0.2 M Sodium iodide		20% (w/v) PEG3350
11	0.2 M Potassium iodide		20% (w/v) PEG3350
15	0.2 M Lithium nitrate		20% (w/v) PEG3350
17	0.2 M Sodium nitrate		20% (w/v) PEG3350
18	0.2 M Potassium nitrate		20% (w/v) PEG3350
19	0.2 M Ammonium nitrate		20% (w/v) PEG3350
20	0.2 M Magnesium formate dihydrate		20% (w/v) PEG3350
21	0.2 M Sodium formate		20% (w/v) PEG3350
22	0.2 M Potassium formate		20% (w/v) PEG3350
23	0.2 M Ammonium formate		20% (w/v) PEG3350
24	0.2 M Lithium acetate dihydrate		20% (w/v) PEG3350
31	0.2 M Lithium sulfate monohydrate		20% (w/v) PEG3350
34	0.2 M Potassium sulfate		20% (w/v) PEG3350
38	0.2 M Ammonium tartrate dibasic		20% (w/v) PEG3350
PEG/Ion2			
14	8% (v/v) Tacsimate pH6.0		20% (w/v) PEG3350
16	8% (v/v) Tacsimate pH7.0		20% (w/v) PEG3350
22	0.2 M Ammonium citrate tribasic pH7.0		20% (w/v) PEG3350
24	0.2 M DL-Malic acid pH7.0		20% (w/v) PEG3350
30	0.2 M Ammonium tartrate dibasic pH7.0		20% (w/v) PEG3350
45	0.15 M Cesium chloride		15% (w/v) PEG3350
46	0.2 M Sodium bromide		20% (w/v) PEG3350
Tg335			
19	0.2 M Lithium sulfate	0.1 M Tris pH 7.0	20% (w/v) PEG3350
48	0.2 M Sodium bromide	0.1 M Bis Tris propane pH 8.0	20% (w/v) PEG3350
54	0.2 M Sodium nitrate	0.1 M Bis Tris propane pH 8.0	20% (w/v) PEG3350
55	0.2 M Sodium thiocyanate	0.1 M Tris pH 7.0	20% (w/v) PEG3350
Tg6K			
58	0.2 M Sodium thiocyanate	0.1 M Tris pH 8.5	20% (w/v) PEG6000

Ad#1: 10 mM Barium chloride dihydrate
Ad#3: 10 mM Calcium chloride dihydrate
Ad#8: 10 mM Strontium chloride hexahydrate
Ad#9: 10 mM Yttrium(III) chloride hexahydrate
Ad#10: 10 mM Zinc chloride
Ad#12: 10 mM Nickel(II) chloride hexahydrate
Ad#14: 10 mM Praseodymium(III) acetate hydrate
Ad#15: 100 mM Ammonium sulfate
Ad#16: 100 mM Potassium chloride
Ad#17: 100 mM Lithium chloride
Ad#18: 200 mM Sodium chloride
Ad#19: 50 mM Sodium fluoride
Ad#20: 100 mM Sodium iodide
Ad#22: 100 mM Potassium sodium tartrate tetrahydrate
Ad#24: 100 mM Cesium chloride
Ad#25: 100 mM Sodium malonate pH 7.0
Ad#26: 10 mM L-Proline
Ad#27: 10 mM Phenol
Ad#28: 3% v/v Dimethyl sulfoxide
Ad#29: 10 mM Sodium bromide
Ad#30: 3% w/v 6-Aminohexanoic acid
Ad#36: 10 mM Taurine
Ad#41: 10 mM Sarcosine
Ad#42: 10 mM Trimethylamine hydrochloride
Ad#43: 100 mM Guanidine hydrochloride
Ad#44: 10 mM Urea
Ad#45: 10 mM b-Nicotinamide adenine dinucleotide hydrate
Ad#48: 1 mM GSH/GSSG
Ad#55: 3% w/v Sucrose
Ad#57: 3% w/v D-Sorbitol
Ad#61: 3% v/v Ethylene glycol
Ad#62: 3% v/v Glycerol
Ad#63: 300 mM NDSB-195
Ad#65: 200 mM NDSB-211
Ad#67: 200 mM NDSB-256
Ad#68: 0.015 mM CYMAL ® -7
Ad#69: 2% w/v Benzamidine hydrochloride
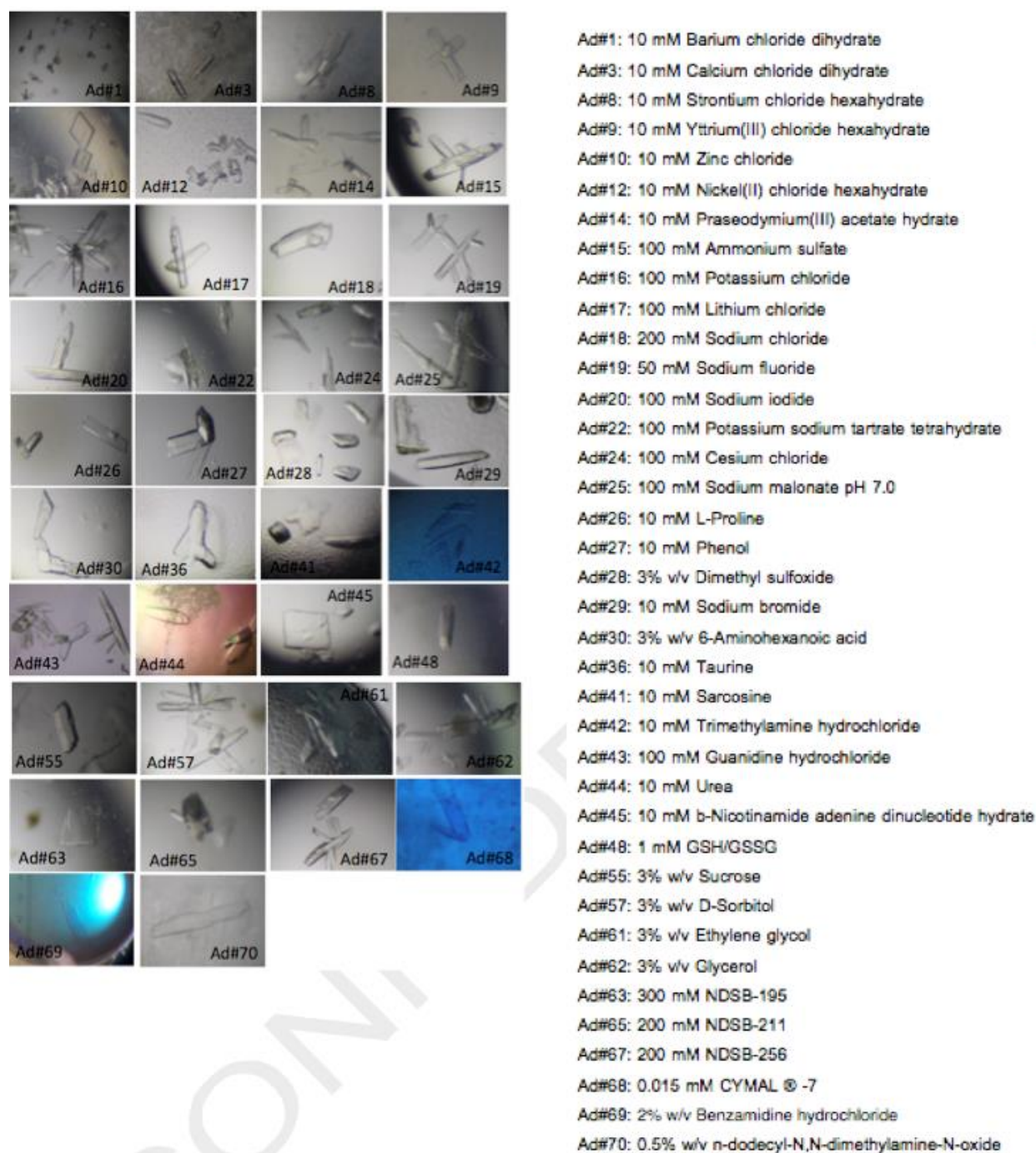Ad#70: 0.5% w/v n-dodecyl-N,N-dimethylamine-N-oxide

Figure 7. Crystals in crystallizing solution LN1/4K with 38 additives.

Figure 8. Crystals in crystallizing solution LN1/4K with and without additive no.69.

(A) Crystals in conditions of 1:2 and 2:3 ratio of protein (10 mg/mL) to crystallizing solution with and without 2% additive no.69.　(B) Crystals in solution LN1/4K with 1–3% benzamidine HCl (Bz, homemade solution) compared to 2% additive #69 (2%Ad69, commercial solution from Hampton Research, USA) using 2:3 ratio of protein (12 mg mL^{-1}) to the solution.

5.7 Structure determination of TgDHFR-TS-inhibitors complexes

We successfully crystallized and collected diffraction data of 15 inhibitors complexes with the full-length TgDHFR-TS protein. Co-complex structures have been refined to the highest resolution of each data set varying from 2.05 to 2.85 Å (**Table 2**). The final R-factor or R$_{work}$, the value for measurement of the discrepancy between the crystallographic model and the experimental X-ray diffraction data, of all structures are shown in **Table 2**. After first cycle of refinement, one data set of B001 complex cannot be further processed due to high R-factor (>30%), possibly from multiple lattices or intermolecular disordered in crystal packing, which are hard to solved. The final R-factor of two co-complex structures (TMP and P195) was quite high (25–27% R$_{work}$), while that of another two complexes (P161 and B005) was moderate (22% R$_{work}$). Therefore, these five data sets (B001, TMP, P195, P161, and B005) were not qualify for further analysis in this study.

5.7.1 Overall structure of TgDHFR-TS

All crystals belong to primitive space group P_1 with four or eight molecules per asymmetric unit (ASU), which are designed A to D or A to H and form two biological dimer. We select molecule A of all co-complex structures for further analysis due to the lowest value of temperature factor or B-factor, which is related to the mean square isotropic displacement of the atom. The overall structure of the full-length TgDHFR-TS has topological structural similarity to DHFR-TS enzymes in other apicomplexans, including *Plasmodium falciparum* DHFR-TS (PfDHFR-TS) and *Cryptosporidium hominis* DHFR-TS (ChDHFR-TS) (Yuvaniyama *et al.*, 2003, O'Neil *et al.*, 2003). Sequence similarity

and secondary structural comparison of TgDHFR-TS and PfDHFR-TS are shown in **Figure 9**. The TgDHFR-TS forms a homodimer with TS:TS interdomain contact, and the N-terminal DHFR domains are located on the shoulders of the TS domains as illustrated in **Figure 10A**. The TS:TS interdomain interface and DHFR:TS domain interface display both electrostatic and hydrophobic interactions (**Table 13** and **Figure 10B**). Two ligands in TS active site are located ~34 Å apart, each composed of residues from both subunits. Two DHFR domains are not in contact and distance of two inhibitors binding in the DHFR pocket is ~66 Å. The flexible loop 1 (residues 51–69) and a part of junction region (JR, residues 257–280) cannot be mapped to electron density signal, and thus can be considered as unstructured in the crystal. The equivalent regions are also unstructured in PfDHFR-TS (a part of Insert 2, residues 86–95 and JR linker, residues 232–280) (Yuvaniyama *et al.*, 2003).

Active sites of DHFR and TS are located on different faces of the TgDHFR-TS structure like other apicomplexan DHFR-TS enzymes (Pf- and Ch- DHFR-TS). The JR linker between DHFR and TS moieties in *T. gondii* contains 69 amino acids (residues 253–321). The long linker (55–91 amino acids) is found in other apicomplexans, whereas the shorter linker presents in other kinetoplastids such as *Trypanosoma brucei*, *Trypanosoma cruzi* and *Leishmania major* (Anderson, 2017). It is reported that the JR is essential for catalytic activity of Pf-, Ch- and Tg- DHFR-TS enzymes (Chaianantakul *et al.*, 2013, Sharma *et al.*, 2013, Vargo *et al.*, 2009). A crossover helix in JR were found in all apicomplexan enzymes. The crossover or kinked helix of TgDHFR-TS contains an approximately 20–residue α-helix (residues 284–303, α7–α8), which is longer than that of the equivalent PfDHFR-TS (10–residue α-helix) and ChDHFR-TS (11–residue α-helix) features. The unique kinked helix shows mainly hydrophobic interactions with the adjacent monomer on the catalytic helix (α1), loop 1 (α2), β7, β9, β11 of DHFR and JR (**Figure 10A**), all of which are important in domain stabilization. The non-conserved hydrophobic interactions between the crossover helix and the adjacent monomer were also found in ChDHFR-TS (O'Neil *et al.*, 2003) while electrostatic interactions were mostly found in PfDHFR-TS (Yuvaniyama *et al.*, 2003). The unique kinked helix of TgDHFR-TS, which is absent in monofunctional human DHFR, would be a target for non-active site inhibitor of TgDHFR-TS as same as ChDHFR-TS (Martucci *et al.*, 2013, Martucci *et al.*, 2009).

5.7.2 Structure of TS domain

The TS domain of *T. gondii* contains 289 amino acid (residues 322–610). As TS in other species, the substrate, dUMP, binds in the TS active site with residues from both molecule A and B (**Figure**

11A). The binding residues around TS active site are conserved including Cys489, which acts as nucleophile attacking the dUMP, and four Arg residues (Arg344 and Arg510, plus Arg469 and Arg470 from the other molecule) that bind phosphate group of dUMP. The hydrogen bonding and hydrophobic interactions of dUMP binding to TS active site are drawn using LigPlot (Laskowski & Swindells, 2011) (**Figure 12**) and listed in **Table 13**. Methotrexate or MTX is found to bind in the active site cavity of TgTS. The binding interactions between TgTS and dUMP and MTX are shown in **Figure 11B**. These interactions are in line with MTX binding to *P. falciparum* TS (unpublished data).

5.7.3 Structure of DHFR domain

The DHFR domain consists of 252 amino acids at N-terminus. It contains three flexible loops; loop1 (residues 41–74), loop2 (residues 103–124) and loop3 (residues 184–227) as shown in **Figure 13**. A part of loop 1 (residues 51–69) was not modeled in the reported structures due to no electron density map which it is equivalent to the unstructured part of Insert2 of PfDHFR (residues 86–95). Overall folding of DHFR domain is similar to the DHFR from other species. The r.m.s. deviations of superposition of ten co-complex structures of TgDHFR were from 0.13 to 0.40 Å. The disordered structures on the flexible loop1 and loop3 regions are clearly observed as shown with high B-factor (**Figure 13A**). Inhibitors and NADPH adopt in the DHFR active site as shown in **Figure 13B-C**. A 2D view of NADPH–protein interactions are shown in **Figure 12**. Around the DHFR active site, loop PRKF (residues 88-91) located next to the DHFR active site and loop 2 (residues 103-124) are a high degree of movement (**Figure 13A**). The 2,4 diaminopyrimidine core of all inhibitors makes hydrogen bonding interactions with the side chain of Asp31 and the main chain carbonyl of Val8 and Val151 in the active site of TgDHFR (**Figure 14** and **Figure 15**). It also makes interactions through one water molecule with the side chain of His34, Thr172 and the main chain carbonyl of Val9, Tyr170. Stacking interactions were found between the pyrimidine ring and the conserved Phe35, together with a phenyl group at the C^5 position of pyrimidine in all four rigid-type inhibitors, Pyr, P39, P40 and B126 and the nicotinamide ring of NADPH. The 2,4-diaminopteridine ring of MTX binding in DHFR active site is in an inverse orientation (180°) in comparison with 2-amino,4-oxo pteridine of DHF (**Figure 16**).

Steric clash between the *p*-Cl phenyl group of Pyr and the side chain of Thr83 predicts low affinity binding of PYR to TgDHFR, as observed (**Figure 14A**). The movement of chlorine atom from

the p- position (Pyr) to the m- position (P40 and B126) mitigates the clash with Thr83. All compounds bind in a different mode to Pyr with much reduced steric clash to Thr83, which explains the lower K_i of all compounds compared with Pyr. In Pyr complex, we observe benzamidine (BDN), additive for crystallization, in the binding pocket with π-π interactions to Phe32, Phe91 and phenyl ring of Pyr (**Figure 14A**). Interestingly, this hydrophobic space is occupied by aliphatic or aromatic hydrocarbon substituted at the C^6 (P39 and P40) or C^5 (P65 and B03) position of the pyrimidine ring (**Figure 17A and 17B**). The occupancy in the hydrophobic pocket was previously reported with the C^5 substituted analogs to improve both potency and selectivity (Welsch $et\ al.$, 2016). This hydrophobic pocket could serve as a promising feature for rational drug design to develop more effective drug against TgDHFR.

We observed hydrogen bonding interactions between Arg97 and the carboxylate group of P218, B160, B253 and B126 (**Figure 17C and 17D**). The specific interactions were found in the binding of P218 to PfDHFR, but not to the hDHFR. Key residues in the equivalent position of TgDHFR, PfDHFR and HsDHFR are listed in **Table 14** for description of required substituents that increase binding affinity/selectivity for TgDHFR.
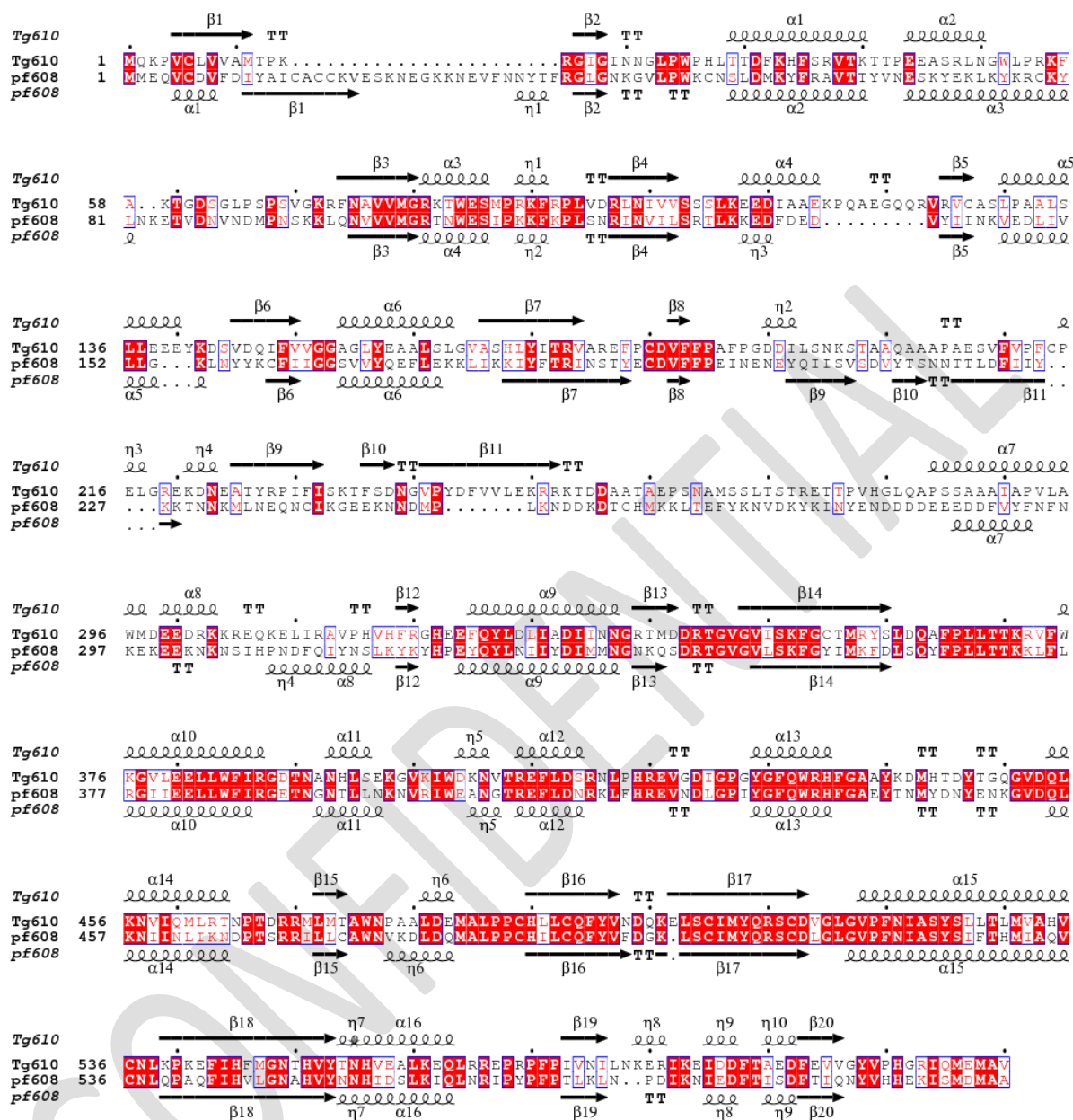
Figure 9. Sequence alignment using clustalX and structures render using EsPript3.0.

Tg610 is TgDHFR-TS complexed with P39 and Pf608 is PfDHFR-TS wild-type complexed with DHF, PDB id: 4DPD.
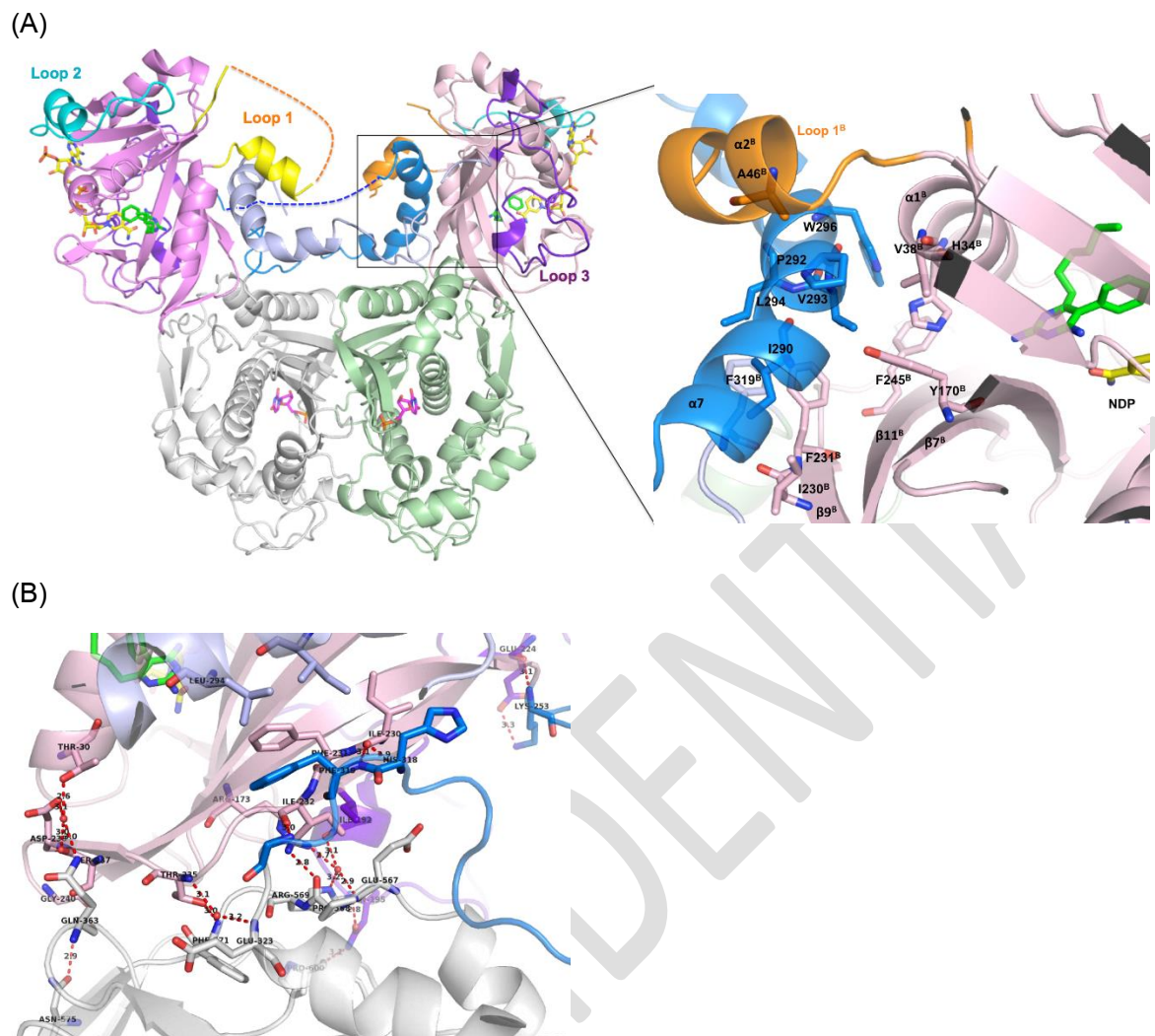
(A)

(B)

Figure 10. Overall structure of TgDHFR-TS and domain interactions.

(A) The DHFR domains of the homodimer (residues 1-252) are in pink and light pink. TS domains (residues 322–610) are in grey and pale green, and the junction regions separating DHFR and TS domains (residue 253–321) are in blue and light blue. Three flexible loops are highlighted in different colors; loop1 (yellow and orange, residues 41–74), loop2 (cyan, residues 103–124) and loop3 (purple, residues 184–227). Ligands are shown in the stick representation. Co-crystallized ligands are highlighted in different colors: P39, green; NADPH, yellow and dUMP in magenta. The unstructured missing regions at loop 1 (residue 51–69, yellow) and the JR (residue 257–280, blue) are denoted by dashed lines (shown only for molecule A). Kinked helix (residue 284–303, $\alpha 7$–$\alpha 8$) of molecule A (blue) and its interactions to the adjacent DHFR-TS molecule are zoomed and key hydrophobic interactions are shown in the stick representation. (B) Interactions between DHFR and TS domains in molecule A. Residues in the interface of DHFR (pink) and TS (grey) domains including loop3 (purple) and JR linker (blue) are shown in stick drawing. P39 is shown in green stick and NADPH in yellow stick. Dashed lines represent hydrogen bonding interactions.

Table 13. List of ligand-protein and protein-protein interactions

	Interactions	Residues
Ligand—Protein		
Pyrimidine core—DHFR	H-bond	Val8, Asp31, Val151
S3 subsite—DHFR	H-bond	Arg97
	Hydrophobic	Val9, Ala10, Phe32, Phe35, Thr83, Phe91
S4 subsite—DHFR	H-bond	Arg97
	Hydrophobic	Val9, Ala10, Phe35, Thr83, Met87, Pro88
NADPH—DHFR	H-bond	Ala10, Ile17, Arg81, Lys82, Thr83, Ser103, Ser104, Ser105, Gly153, Ala154, Gly155
	Hydrophobic	Val9, Gly18, Asn21, Gly22, Leu23, Trp25, Gly80, Val102, Ala128, Val151, Gly152, Leu156, Tyr157, Ala159, Val182
dUMP—TS	H-bond	Arg344, Arg469[B], Arg470[B], Arg510, Ser511, Asn521, Asp513, His551, Tyr553
	Hydrophobic	Cys489, Gln509, Cys512, Gly517, His490
Protein—Protein		
DHFR:TS (A:A)	H-bond	Arg173—Pro568, Asn195—Arg569/Pro600, Ile230—His318/Phe319, Ile232—Phe319/Gly321/His322, Thr235—Phe571, Ser237—Gln363
	Hydrophobic	Arg176—Val596/Gly597, Ile192—Glu567, Pro229—Gly567/His316, Phe231—Phe319, Pro242—Ile573, Asp244—Pro570
Donated helix (α7):DHFR, TS (A:B)	H-bond	Ser286—Glu249, Glu299—Arg37
	Hydrophobic	Ile290—Phe231, Ala291—Ala46, Val293—Tyr170/Phe231, Leu294—Phe231/Phe319, Trp296—Thr30/His34/Phe245, Met297—Phe231/Phe236/Phe245
TS:TS (A:B)	H-bond	Arg339—Asn498/Asp499, Ser351—Tyr496, Phe353/Gly354—Arg358, Asp343—Arg469, Arg358—Phe353/Gly354, Arg469—Asp343/Arg510/Tyr553, Gln494—Arg510, Tyr496—Ser351, Asn498/Arg499—Arg339, Arg510—Gln494/Arg469, Tyr553—Arg469
	Hydrophobic	Phe353—Tyr496/Ile506/Ile544, Thr356—Thr356, Phe436—Pro478, Leu472—Leu491, Thr474—Trp476/Pro478, Trp476—Thr474, Pro478—Phe436/Thr474, Leu491—Leu472/Leu492, Tyr496—Phr353, Ile506—Phr353, Tyr508—Phe546, Ile544—Phe353, Phe546—Thr356/Tyr508/Phe546

The interactions identified by LigPlot+ of interatomic distances within 2.7 and 3.35 Å for hydrogen-bonding interactions (H-bond acceptor and donor, respectively) and 3.9 Å for hydrophobic interactions.
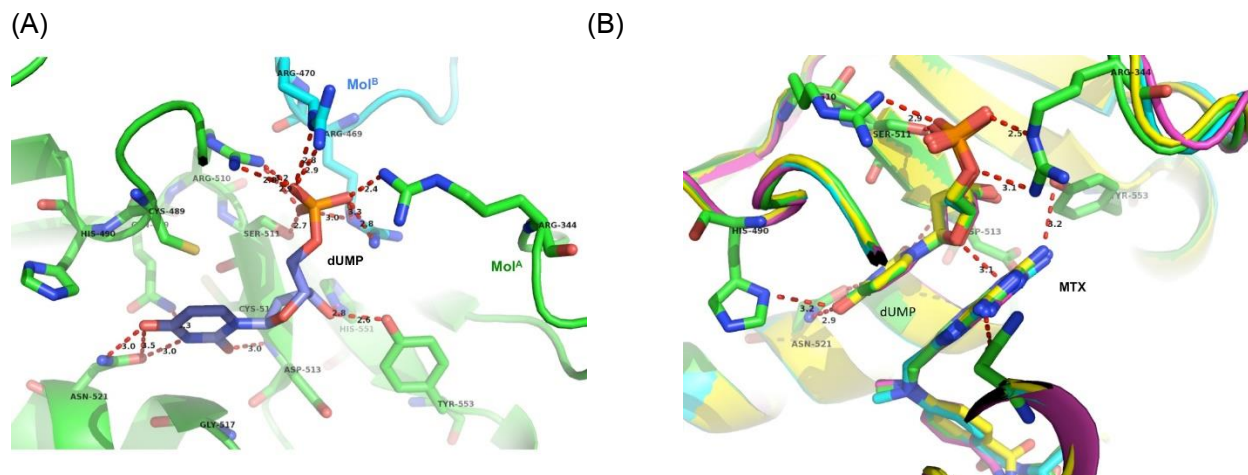
(A)

(B)

Figure 11. TS active site.

(A) Interactions of dUMP binding between mol A and mol B. (B) dUMP and MTX bind in the TS active site of four molecules. The r.m.s. deviations of superposition of mol B, C, and D on mol A are 0.192, 0.152 and 0.313 Å, respectively, using PyMOL.
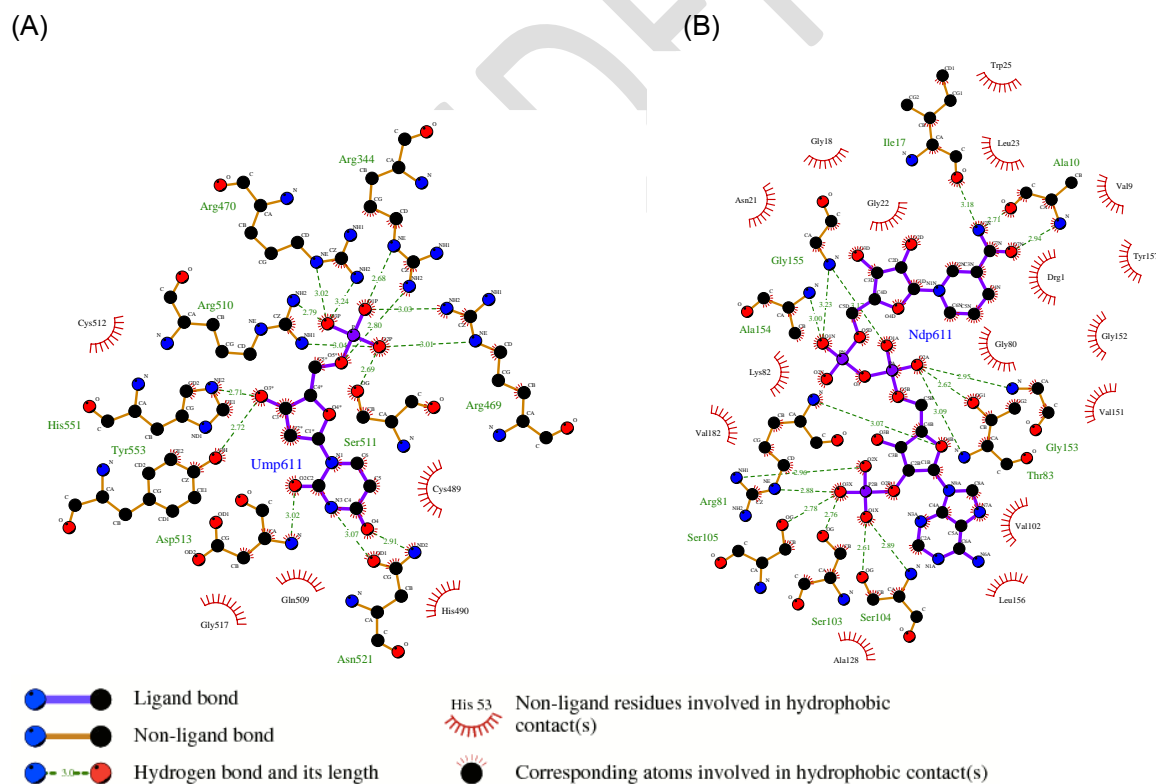
(A)

(B)

Figure 12. Ligand-protein interactions in DHFR and TS active sites.

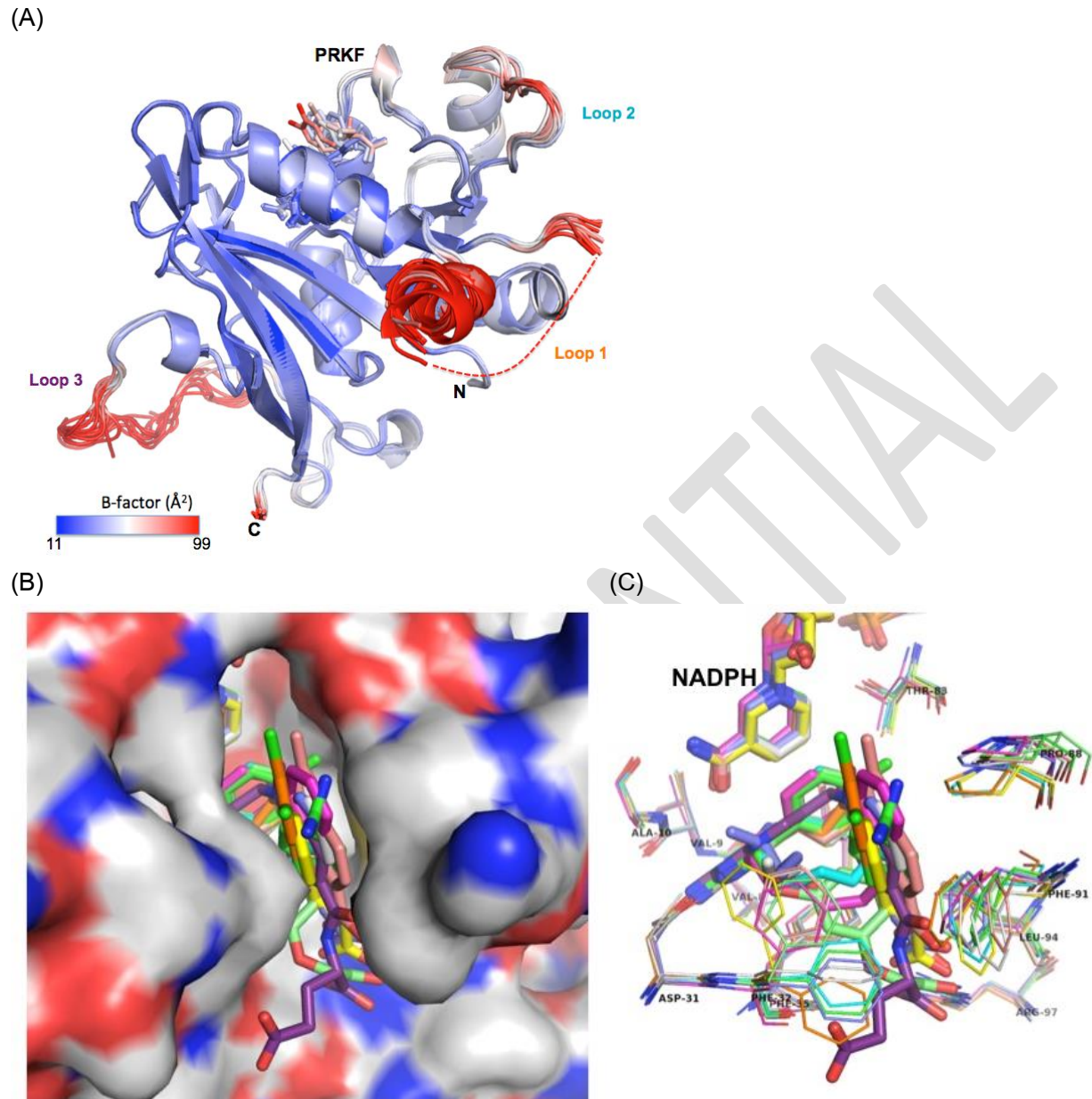Ligand plot shows the interactions in dUMP (A) and NADPH (B) in the binding pockets.

(A)

(B)

(C)

Figure 13. Overall DHFR domain and the active site.

(A) Overall co-complexes structure with all inhibitors together with NADPH is shown with spectrum coloring based on B-factor values (minimum at 11 Å^2, blue - maximum at 90Å^2, red). (B) Protein shows as surface surrounding ligands, inhibitors and NADPH. (C) Key residues in active site are labelled and drawn as ribbon and the ligands are in stick.
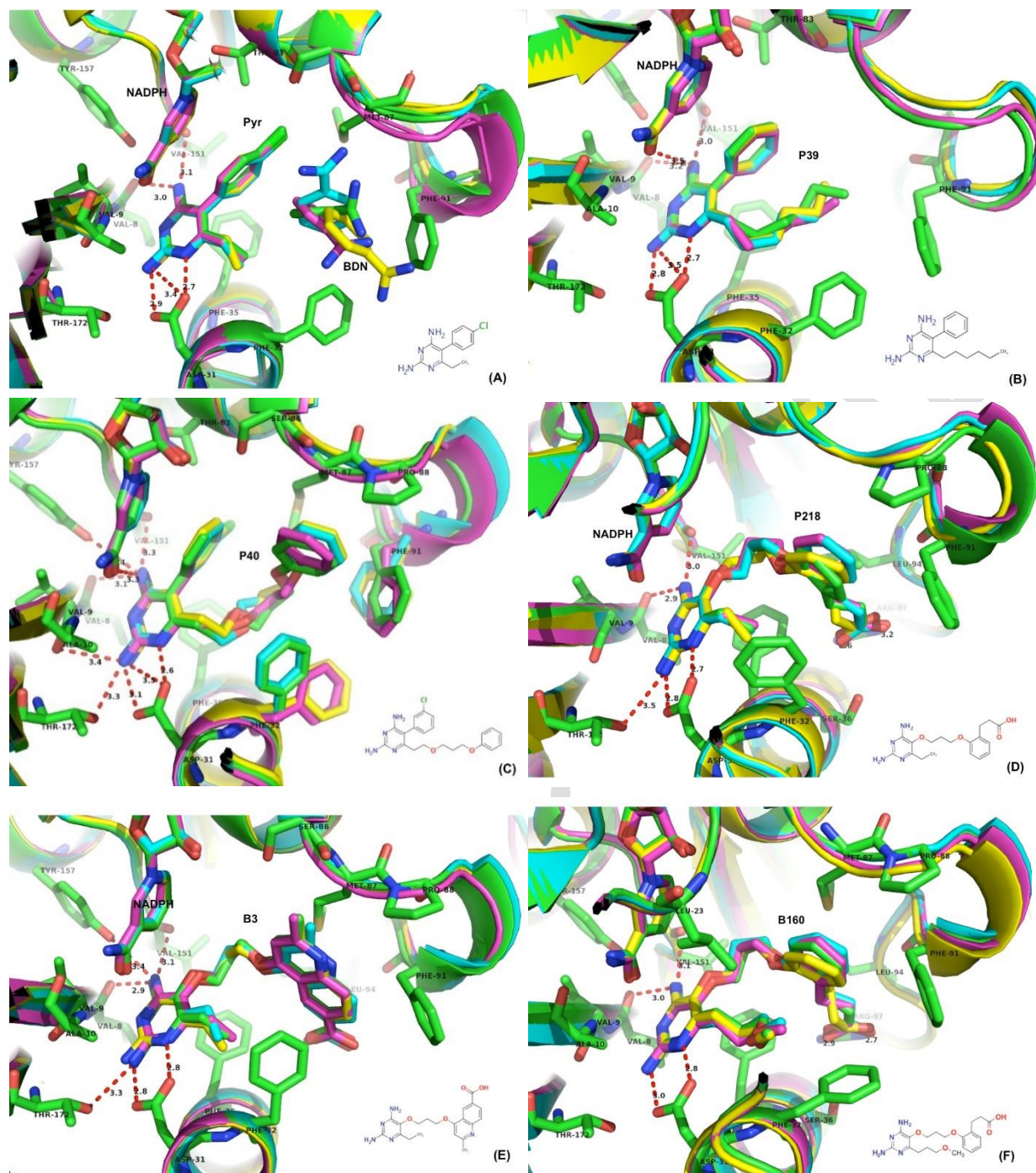
Figure 14. Superposition of 4 molecules of co-complex structures.

Molecule A, B, C and D are drawn in green, cyan, pink and yellow. Residues within 3.8 Å around DHFR active site are shown as stick.
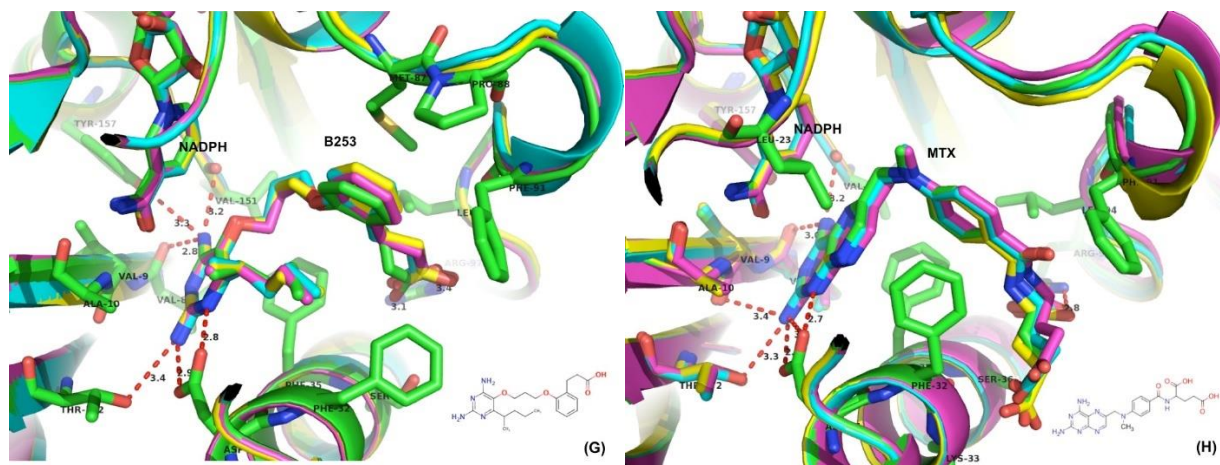
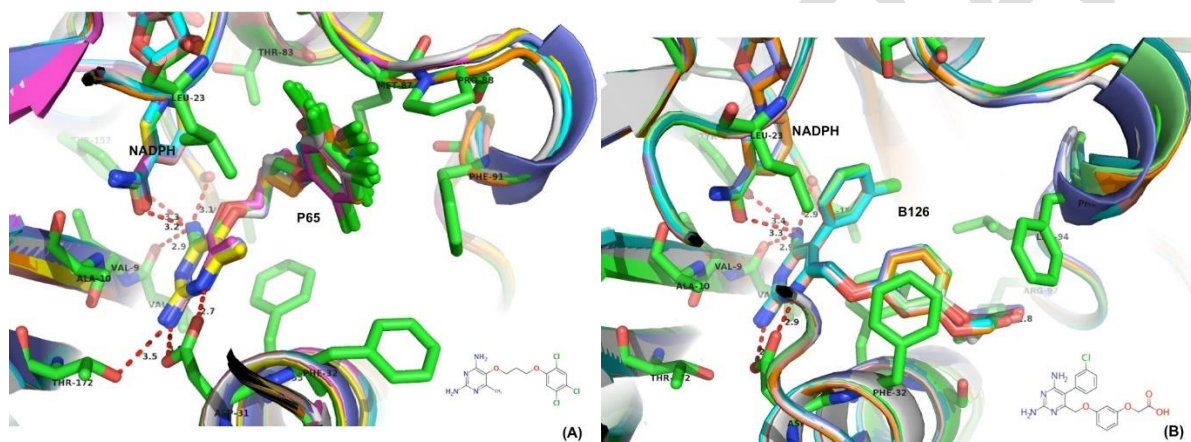**Figure 14. Superposition of 4 molecules of co-complex structures.**

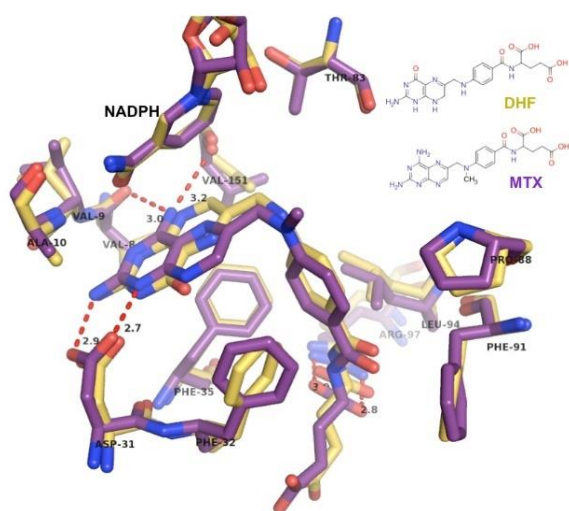**Figure 15. Superposition of 8 molecules of co-complex structures.**

Figure 16. Superposition of MTX and DHF (PDB id: 4EIL) complexed TgDHFR structure.
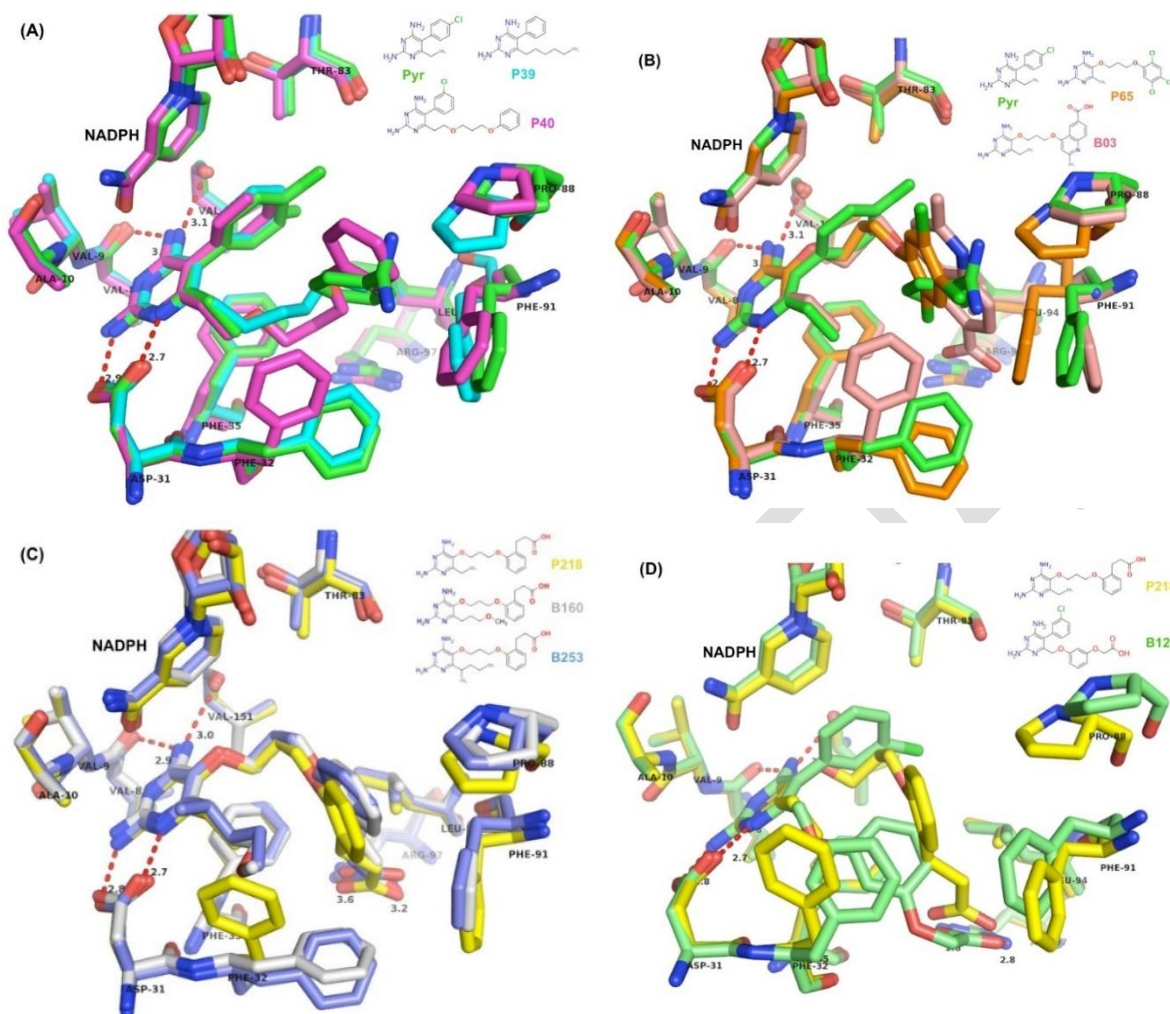
Figure 17. Superposition of co-complex structures.

Table 14. List of key residues in equivalent position of Tg-, Pf- and human DHFR that interact to DHFR (within 4 Å).

TgDHFR	PfDHFR	HsDHFR	DHFR-DHF interaction	Description of substituents for TgDHFR
Major interactions				
Val8	Ile14	Ile7	H-bond (carbonyl)	4-amino of pyrimidine ring
Val9	Cys15	Val8	H-bond (carbonyl)	2-amino of pyrimidine ring
Leu23	Leu46	Leu22	Hydrophobic	Hydrophobic substituent at C^5 or C^6 of pyrimidine ring
Asp31	Asp54	Glu30	H-bond (sidechain)	N^1, 2-amino of pyrimidine ring
Phe35	Phe58	Phe34	π-π interaction	Core structure with aromatic hydrocarbon e.g. pyrimidine, dihydrotriazine, quinazoline, etc.
Pro88	Pro113	Pro61	Hydrophobic	Hydrophobic substituent at C^5 or C^6 of pyrimidine ring
Arg97	Arg122	Arg70	H-bond (sidechain)	Electrostatics group at C^5 or C^6 of pyrimidine ring
Leu94	Leu119	Leu67	Hydrophobic	Hydrophobic substituent at C^5 or C^6 of pyrimidine ring
Val151	Ile164Leu	Val115	Hydrophobic	H-bonding (carbonyl to 4-amino of pyrimidine ring)
Thr172	Thr185	Thr136	H-bond (sidechain)	2-amino of pyrimidine ring
Species specific interactions				
Gly22	Val45	Asp21	van der Waals	Small substituent
Phe32	Met55	Phe31	π-π interaction and hydrophobic	Hydrophobic substituent at C^5 or C^6 of pyrimidine ring
Ser36	Cys59Arg	Gln35	H-bond (Cys59Arg in PfDHFR)	Electrostatics group
Thr83	Ser108Asn	Thr56	No interaction with DHF, but steric conflict with Pyr	Des- or m- substituent phenyl or flexible chain at C^5 of pyrimidine ring
Met87	Ile112	Ile60	Hydrophobic	Hydrophobic substituent at C^5 or C^6 of pyrimidine ring
Phe91	Phe116	Asn64	Hydrophobic and partial π-π interaction in TgDHFR and PfDHFR, but H-bonding to Asn64 in HsDHFR	Hydrophobic substituent at C^5 or C^6 of pyrimidine ring

6. Conclusion and Discussion

The toxicity of high dose regimen of pyrimethamine for toxoplasmosis treatment has pressed an urgent need to develop new safer therapies for toxoplasmosis treatment. To identify compounds effective against TgDHFR-TS, we established bacterial surrogate system and *T. gondii* cell system. We screened 218 compounds, including rigid, semi-rigid and flexible types of three scaffolds (pyrimidines, dihydrotriazines and quinazolines) for the ability to inhibit TgDKO-TolC cell growth. A total of 134 compounds are more active than the standard compound, Pyr (IC$_{50}$<38.1 µM), in which 24 compounds inhibit the cell growth with IC$_{50}$ <1 µM. One advantage of our bacterial system is that it is simple and not complicated for screening new inhibitors against TgDHFR-TS. However, it has limitation of compounds testing with TgDKO-TolC due in part to the hydrophobic/hydrophilic properties and thus cell permeability of the compounds. We found that compounds with cLogP>3 are suitable for testing with our bacterial system, which is in line with an optimum cLogP between 3 and 5 of drug-like compounds with high permeability (Leeson & Springthorpe, 2007).

We have identified a set of effective compounds by direct enzyme inhibition assay, TgDKO-TolC surrogate system and validated by *T. gondii* inhibition assay. Interestingly, compounds with carboxylic group substituent at S3 subsite of pyrimidines are more active than compounds with ester group as determined by enzyme assay while their ester counterparts are more active than the carboxylic compounds as determined by TgDKO-TolC and *T. gondii* cell-based assays, suggesting that hydrophobic compounds or prodrug form of active carboxylic containing antifolates can be used. The most effective compounds (B001, B002, and B004) are flexible antifolates with quinoline moiety at S3 subsite (IC$_{50}$ < 5 nM by *T. gondii* cell assay). Although they are not cytotoxic to HFFs host cell used in *T. gondii* cell assay, they are cytotoxic to Vero and KB cells. Other flexible antifolates are also highly potent against *T. gondii* parasites. These could serve as good templates for development of effective anti-Tg antifolates.

We have successfully solved high resolution structures of the full-length TgDHFR-TS complexed with Pyr and identified effective inhibitors. The co-complex structures with Pyr clarify the conflict interaction between *p*-Cl phenyl of Pyr and Thr83 of TgDHFR, which could explain the low binding affinity of the drug with TgDHFR. Our structural data also indicate that the pyrimidine antifolates that were designed to avoid such conflict and carrying either additional electrostatic group for hydrogen bonding with Arg97 (P218, B160, B253 and B126) or with suitable hydrophobic substituent on C^6 for interacting with the hydrophobic pocket (PRKF loop and Phe32) (P39 and P40) are with high potency against TgDHFR. Moreover, the unique kinked crossover helix in the linker of

TgDHFR-TS provides hydrophobic interactions to the adjacent monomer, which is similar to that found in ChDHFR-TS but different from PfDHFR-TS where most interaction are electrostatic (O'Neil *et al.*, 2003, Yuvaniyama *et al.*, 2003). This information may be targeted for designing non-active site inhibitors. It can be concluded that the tools and knowledge generated from this project should form basis for facilitating discovery of effective and safe anti-Tg antifolates in the near future.
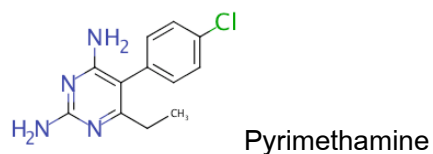
7. References

Anderson, K. S. (2017). Understanding the molecular mechanism of substrate channeling and domain communication in protozoal bifunctional TS-DHFR. *Protein Eng Des Sel* **30**, 253-261.

Chaianantakul, N., Sirawaraporn, R. & Sirawaraporn, W. (2013). Insights into the role of the junctional region of *Plasmodium falciparum* dihydrofolate reductase-thymidylate synthase. *Malar J* **12**, 91.

Chitnumsub, P., Yuvaniyama, J., Vanichtanankul, J., Kamchonwongpaisan, S., Walkinshaw, M. D. & Yuthavong, Y. (2004). Characterization, crystallization and preliminary X-ray analysis of bifunctional dihydrofolate reductase-thymidylate synthase from *Plasmodium falciparum. Acta Crystallogr. D Biol. Crystallogr.* **60**, 780-783.

Ciccone, L., Vera, L., Tepshi, L., Rosalia, L., Rossello, A. & Stura, E. A. (2015). Multicomponent mixtures for cryoprotection and ligand solubilization. *Biotechnol Rep (Amst)* **7**, 120-127.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501.

Gupta, S., Cohen, K. A., Winglee, K., Maiga, M., Diarra, B. & Bishai, W. R. (2014). Efflux inhibition with verapamil potentiates bedaquiline in *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother* **58**, 574-576.

Kopytek, S. J., Dyer, J. C., Knapp, G. S. & Hu, J. C. (2000). Resistance to methotrexate due to AcrAB-dependent export from *Escherichia coli*. *Antimicrob Agents Chemother* **44**, 3210-3212.

Lamers, R. P., Cavallari, J. F. & Burrows, L. L. (2013). The efflux inhibitor phenylalanine-arginine beta-naphthylamide (PAbetaN) permeabilizes the outer membrane of gram-negative bacteria. *PLoS One* **8**, e60666.

Laskowski, R. A. & Swindells, M. B. (2011). LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J Chem Inf Model* **51**, 2778-2786.

Leeson, P. D. & Springthorpe, B. (2007). The influence of drug-like concepts on decision-making in medicinal chemistry. *Nat Rev Drug Discov* **6**, 881-890.

Lovell, S. C., Davis, I. W., Arendall, W. B., 3rd, de Bakker, P. I., Word, J. M., Prisant, M. G., Richardson, J. S. & Richardson, D. C. (2003). Structure validation by Calpha geometry: phi,psi and Cbeta deviation. *Proteins* **50**, 437-450.

Martucci, W. E., Rodriguez, J. M., Vargo, M. A., Marr, M., Hamilton, A. D. & Anderson, K. S. (2013). Exploring novel strategies for AIDS protozoal pathogens: alpha-helix mimetics targeting a key allosteric protein-protein interaction in *C. hominis* TS-DHFR. *Medchemcomm* **4**.
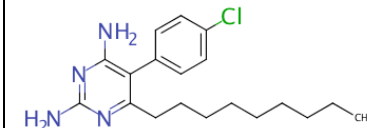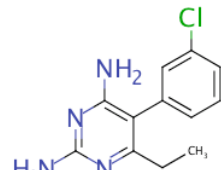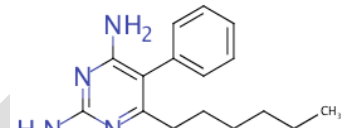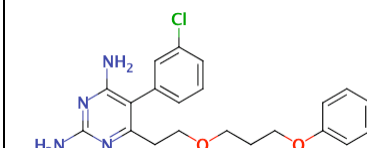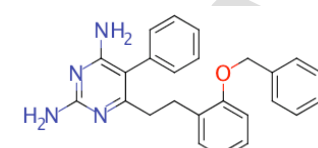
Martucci, W. E., Udier-Blagovic, M., Atreya, C., Babatunde, O., Vargo, M. A., Jorgensen, W. L. & Anderson, K. S. (2009). Novel non-active site inhibitor of *Cryptosporidium hominis* TS-DHFR identified by a virtual screen. *Bioorg Med Chem Lett* **19**, 418-423.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). Phaser crystallographic software. *J Appl Crystallogr* **40**, 658-674.

Murshudov, G. N., Skubak, 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 Crystallogr D Biol Crystallogr* **67**, 355-367.

Nikaido, H. (2001). Preventing drug access to targets: cell surface permeability barriers and active efflux in bacteria. *Semin Cell Dev Biol* **12**, 215-223.

O'Neil, R. H., Lilien, R. H., Donald, B. R., Stroud, R. M. & Anderson, A. C. (2003). Phylogenetic classification of protozoa based on the structure of the linker domain in the bifunctional enzyme, dihydrofolate reductase-thymidylate synthase. *J Biol Chem* **278**, 52980-52987.

Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods in Enzymology* **276**, 307-326.

Reynolds, M. G. & Roos, D. S. (1998). A biochemical and genetic model for parasite resistance to antifolates. *Toxoplasma gondii* provides insights into pyrimethamine and cycloguanil resistance in *Plasmodium falciparum*. *J Biol Chem* **273**, 3461-3469.

The PyMOL Molecular Graphics System, Version 1.3r1Schrödinger, LLC (2010). The PyMOL Molecular Graphics System, Version 1.3r1, Vol.

Sharma, H., Landau, M. J., Vargo, M. A., Spasov, K. A. & Anderson, K. S. (2013). First Three-Dimensional Structure of *Toxoplasma gondii* Thymidylate Synthase-Dihydrofolate Reductase: Insights for Catalysis, Interdomain Interactions, and Substrate Channeling. *Biochemistry* **52**, 7305-7317.

Vagin, A. & Teplyakov, A. (2010). Molecular replacement with MOLREP. *Acta Crystallogr D Biol Crystallogr* **66**, 22-25.

Vargo, M. A., Martucci, W. E. & Anderson, K. S. (2009). Disruption of the crossover helix impairs dihydrofolate reductase activity in the bifunctional enzyme TS-DHFR from *Cryptosporidium hominis*. *Biochem J* **417**, 757-764.

Vera, L., Czarny, B., Georgiadis, D., Dive, V. & Stura, E. A. (2011). Practical Use of Glycerol in Protein Crystallization. *Cryst Growth Des* **11**, 2755-2762.

Welsch, M. E., Zhou, J., Gao, Y., Yan, Y., Porter, G., Agnihotri, G., Li, Y., Lu, H., Chen, Z. & Thomas, S. B. (2016). Discovery of Potent and Selective Leads against *Toxoplasma gondii* Dihydrofolate Reductase via Structure-Based Design. *ACS Med Chem Lett* **7**, 1124-1129.

Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A. & Wilson, K. S. (2011). Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **67**, 235-242.

Yuvaniyama, J., Chitnumsub, P., Kamchonwongpaisan, S., Vanichtanankul, J., Sirawaraporn, W., Taylor, P., Walkinshaw, M. D. & Yuthavong, Y. (2003). Insights into antifolate resistance from malarial DHFR-TS structures. *Nat Struct Biol* **10**, 357-365.
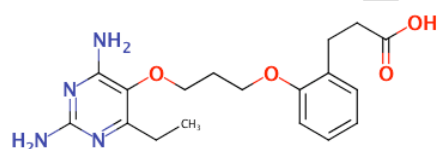
8. Appendix

8.1. Chemical structure of antifolates hits in this study.
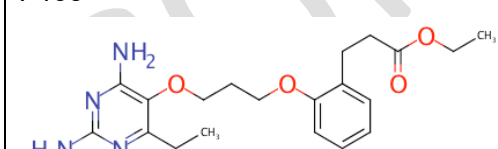
(A) Pyrimidine scaffold: Rigid compounds

Pyrimethamine
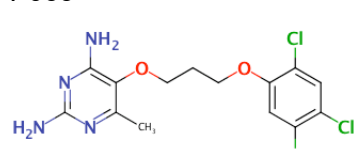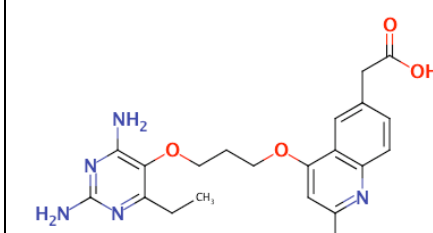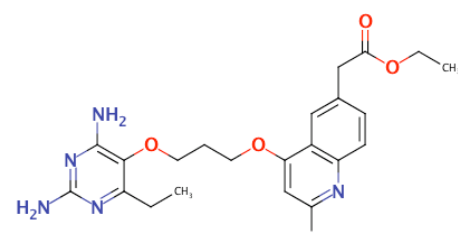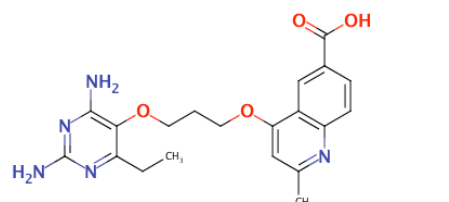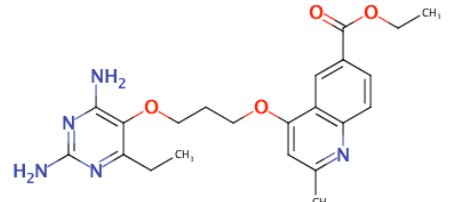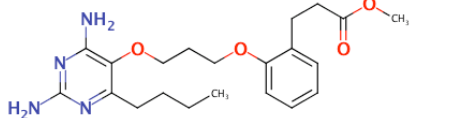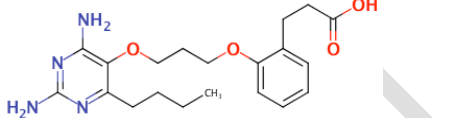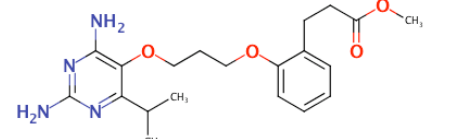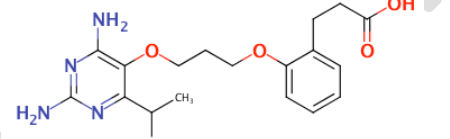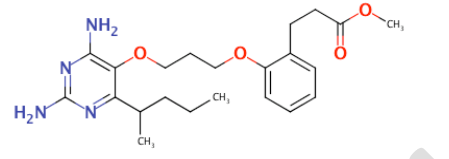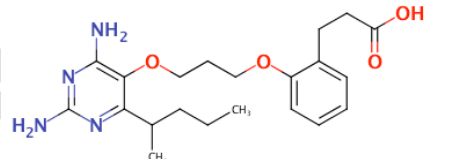
P008	P030	P039
P040	B015	

(B) Pyrimidine scaffold: Flexible compounds
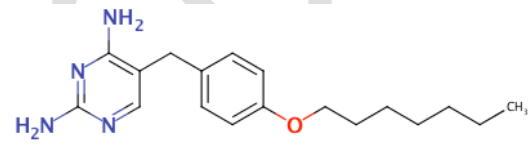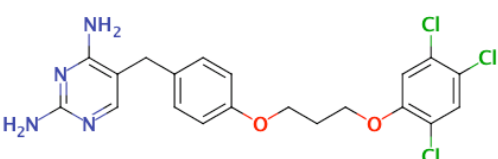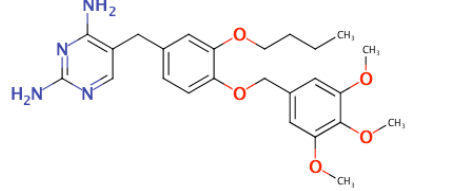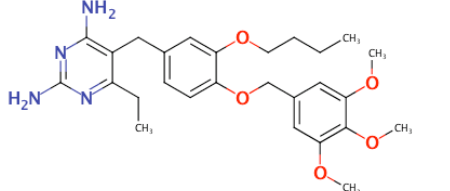
P218

P195	P065
B001	B002

B003	B004

B241	B242

B250	B251

B252	B253

(C) Pyrimidine scaffold: 55 Semi-rigid compounds

Trimethoprim

T44	T51

T57	T58

T59	T69
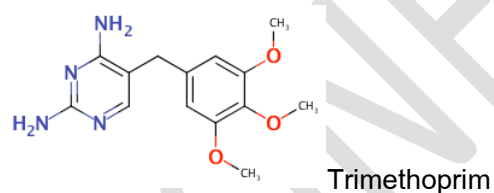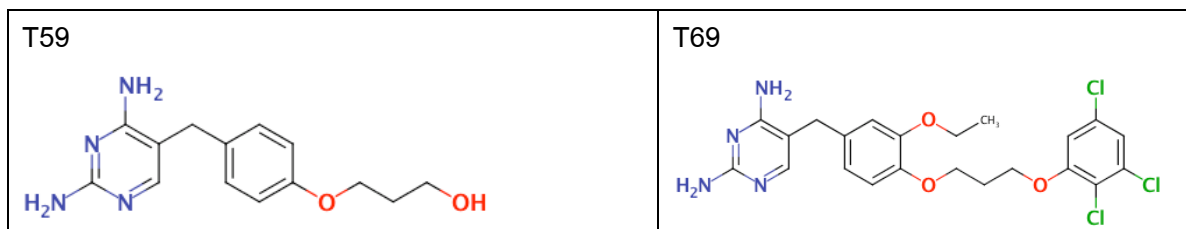
(D) Triazine scaffold: 3 Rigid and 15 Flexible compounds

Cycloguanil

WR99210

C420	C448	C452
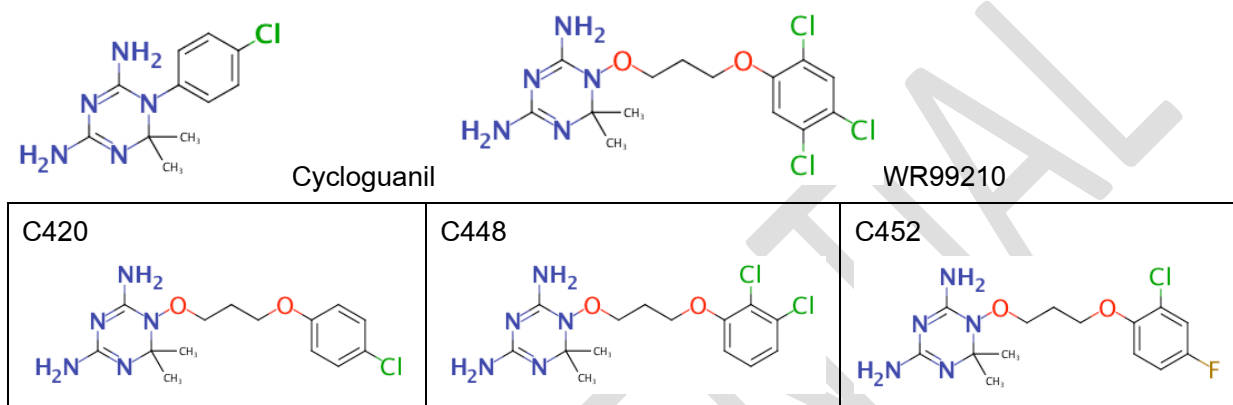
(E) Quinazoline scaffold: 18 compounds

Q06	Q11	Q14
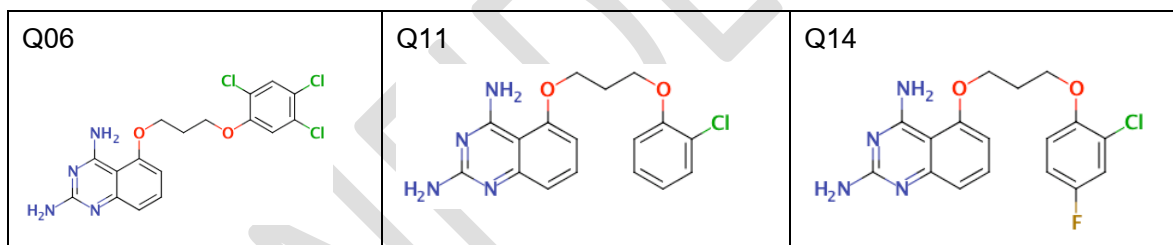
8.2. The IC$_{50}$ values of all compounds tested in this study against TgDHFR-TS using enzyme and bacteria assay

Cpd	IC$_{50}$ (µM)		
	Enzyme assay	PEcΔTF-TgDT	PEcΔTFTolC-TgDT
Pyrimidine scaffold: 42 Rigid compounds			
PYR	0.230 ± 0.030	107.800 ± 23.400	38.100 ± 3.300
P002	0.263 ± 0.035	>500	13.047 ± 2.439
P004	>100	>500	>500
P005	3.884 ± 0.117	>500	35.859 ± 9.861
P007	0.277 ± 0.020	175.525 ± 23.024	28.325 ± 7.441
P008	0.387 ± 0.018	>500	0.705 ± 0.043
P010	>500	>500	162.974 ± 8.821
P013	0.101 ± 0.027	74.467 ± 11.578	2.026 ± 0.548
P014	12.838 ± 2.446	160.371 ± 28.032	95.254 ± 7.116
P015	0.251 ± 0.017	70.694 ± 15.588	46.202 ± 8.098
P016	0.193 ± 0.011	>500	16.462 ± 2.848
P017	0.410 ± 0.053	152.671 ± 22.779	27.267 ± 4.191
P019	>500	65.036 ± 6.780	28.410 ± 1.542
P020	0.276 ± 0.016	188.390 ± 30.689	79.085 ± 11.098
P021	0.144 ± 0.005	136.654 ± 18.237	32.699 ± 4.920
P022	0.157 ± 0.043	>500	2.589 ± 0.530
P023	3.826 ± 0.474	142.278 ± 11.356	31.286 ± 4.013
P025	1.695 ± 0.198	105.822 ± 15.721	144.033 ± 23.489
P028	5.386 ± 0.356	>500	276.330 ± 18.463
P030	0.087 ± 0.006	86.094 ± 3.199	5.145 ± 0.548
P034	0.255 ± 0.042	161.441 ± 4.090	30.292 ± 0.802
P036	2.777 ± 0.093	118.133 ± 18.479	148.120 ± 22.177
P037	1.570 ± 0.112	65.128 ± 8.176	208.074 ± 33.077
P038	0.274 ± 0.049	26.127 ± 8.579	448.493 ± 45.115
P039	0.076 ± 0.002	43.505 ± 8.144	1.447 ± 0.220
P040	0.081 ± 0.003	>500	0.176 ± 0.018
P049	0.301 ± 0.024	235.344 ± 39.651	3.502 ± 0.562
P050	0.321 ± 0.051	116.204 ± 9.461	2.890 ± 0.793
P051	0.325 ± 0.023	151.050 ± 32.473	2.297 ± 0.652
B006	nd	nd	>250

Cpd	IC$_{50}$ (µM)					
	Enzyme assay			PEcΔTF-TgDT		PEcΔTFTolC-TgDT
B007	5.611	±	2.454	nd		328.322 ± 53.542
B008	nd			nd		210.234 ± 40.520
B009	nd			nd		271.330 ± 58.963
B010	nd			nd		93.716 ± 11.831
B011	nd			nd		98.989 ± 22.476
B013	nd			nd		1.696 ± 0.183
B014	nd			nd		2.558 ± 0.599
B015	0.203	±	0.085	nd		0.340 ± 0.048
B016	nd			nd		9.355 ± 2.056
B017	nd			nd		>100
B018	nd			nd		>250
B126	0.090	±	0.017	nd		319.990 ± 16.139
Pyrimidine scaffold: 85 Flexible compounds						
P218	0.059	±	0.001	>500		32.500 ± 4.900
P065	0.125	±	0.022	37.574 ± 2.164		5.750 ± 1.500
P066	0.099	±	0.003	377.024 ± 71.585		31.304 ± 4.452
P067	0.085	±	0.007	<100		3.453 ± 0.415
P068	0.142	±	0.007	nd		23.037 ± 3.053
P070	0.218	±	0.023	140.576 ± 22.609		35.695 ± 3.183
P071	0.127	±	0.008	77.423 ± 6.420		4.571 ± 1.293
P072	0.143	±	0.014	131.551 ± 4.801		9.060 ± 1.075
P073	0.492	±	0.036	63.863 ± 9.257		19.422 ± 3.075
P074	0.638	±	0.084	66.816 ± 1.274		15.401 ± 1.134
P075	0.263	±	0.048	43.914 ± 12.157		8.790 ± 0.755
P076	0.689	±	0.050	193.910 ± 3.843		50.585 ± 10.907
P077	0.629	±	0.094	272.015 ± 38.827		74.032 ± 13.980
P082	0.133	±	0.018	51.608 ± 2.341		1.691 ± 0.491
P083	0.075	±	0.005	152.822 ± 21.939		10.086 ± 1.311
P085	0.143	±	0.013	13.836 ± 3.119		0.540 ± 0.047
P086	0.092	±	0.009	36.821 ± 2.617		0.871 ± 0.155
P087	0.065	±	0.000	16.814 ± 1.345		0.605 ± 0.140
P101	nd			301.933 ± 34.611		13.506 ± 1.056
P102	nd			113.390 ± 6.095		6.081 ± 0.598

Cpd	IC$_{50}$ (μM)		
	Enzyme assay	PEcΔTF-TgDT	PEcΔTFTolC-TgDT
P114	nd	nd	15.952 ± 2.885
P127	nd	nd	17.683 ± 4.827
P128	nd	nd	8.182 ± 0.922
P129	nd	nd	6.949 ± 1.254
P131	nd	nd	22.574 ± 5.644
P136	0.129 ± 0.026	nd	6.089 ± 1.573
P137	7.457 ± 0.131	nd	344.096 ± 17.648
P143	nd	nd	1.477 ± 0.150
P144	nd	nd	1.516 ± 0.193
P148	nd	nd	2.583 ± 0.727
P150	nd	nd	1.551 ± 0.121
P154	0.104 ± 0.060	nd	0.547 ± 0.050
P161	0.129 ± 0.016	nd	5.241 ± 0.262
P165	0.122 ± 0.044	nd	0.606 ± 0.072
P166	nd	nd	1.267 ± 0.307
P172	nd	nd	5.343 ± 0.745
P174	0.162 ± 0.023	nd	383.168 ± 52.354
P175	0.133 ± 0.037	nd	0.206 ± 0.034
P176	nd	nd	>500
P180	nd	nd	2.793 ± 0.303
P191	nd	nd	1.472 ± 0.219
P192	nd	nd	1.621 ± 0.251
P193	nd	nd	2.216 ± 0.218
P194	nd	nd	6.645 ± 1.348
P195	0.305 ± 0.033	nd	3.925 ± 0.572
P198		nd	1.619 ± 0.434
P199	0.151 ± 0.010	nd	0.905 ± 0.208
P200	nd	nd	1.244 ± 0.103
P201	nd	nd	1.865 ± 0.518
P202	nd	nd	2.839 ± 0.407
P204	0.123 ± 0.024	nd	0.547 ± 0.021
P205	nd	nd	3.636 ± 0.058
P206	0.222 ± 0.054	nd	0.282 ± 0.019

Cpd	IC$_{50}$ (μM)		
	Enzyme assay	PEcΔTF-TgDT	PEcΔTFTolC-TgDT
P207	nd	nd	2.819 ± 0.522
P208	nd	nd	6.697 ± 1.564
P209	0.239 ± 0.090	nd	2.224 ± 0.445
P210	0.177 ± 0.017	nd	0.186 ± 0.030
P211	0.142 ± 0.010	nd	8.893 ± 2.272
P212	nd	nd	5.905 ± 0.305
P213	nd	nd	5.187 ± 0.800
P215	0.226 ± 0.023	nd	0.723 ± 0.179
P217	nd	nd	5.566 ± 1.170
P219	0.520 ± 0.038	nd	0.774 ± 0.081
P220	0.275 ± 0.104	nd	0.722 ± 0.060
B001	0.148 ± 0.023	nd	13.655 ± 3.874
B002	0.172 ± 0.013	nd	0.181 ± 0.052
B003	0.046 ± 0.0036	nd	61.230 ± 3.820
B004	0.108 ± 0.039	nd	0.070 ± 0.007
B005	0.117 ± 0.008	nd	0.559 ± 0.051
B012	0.126 ± 0.002	nd	2.497 ± 0.149
B019	nd	nd	>250
B021	nd	nd	>250
B119	0.143 ± 0.021	nd	55.511 ± 4.939
B160	0.107 ± 0.026	nd	133.908 ± 11.140
B169	3.628 ± 1.311	nd	52.123 ± 3.837
B237	0.112 ± 0.017	nd	1.098 ± 0.190
B238	0.088 ± 0.014	nd	15.846 ± 1.200
B239	0.177 ± 0.008	nd	5.293 ± 0.323
B240	0.122 ± 0.032	nd	10.730 ± 2.070
B241	0.165 ± 0.019	nd	4.315 ± 0.630
B242	0.113 ± 0.019	nd	2.068 ± 0.313
B250	0.093 ± 0.009	nd	0.344 ± 0.038
B251	0.066 ± 0.002	nd	10.822 ± 1.778
B252	0.174 ± 0.008	nd	1.831 ± 0.083
B253	0.089 ± 0.012	nd	6.171 ± 0.146
Pyrimidine scaffold: 55 Semi-rigid compounds			

Cpd	IC$_{50}$ (μM)		
	Enzyme assay	PEcΔTF-TgDT	PEcΔTFTolC-TgDT
Tmp	50.398 ± 13.765	>500	>500
T05	nd	nd	454.976 ± 127.071
T06	nd	nd	279.619 ± 73.799
T07	nd	nd	85.987 ± 3.710
T08	nd	nd	143.950 ± 28.355
T09	nd	nd	63.883 ± 3.365
T10	nd	nd	32.469 ± 7.650
T11	nd	nd	27.047 ± 4.747
T13	3.432 ± 0.381	>500	15.889 ± 1.579
T14	nd	nd	178.017 ± 7.945
T15	nd	nd	>1000
T16	nd	nd	>1000
T17	nd	nd	245.857 ± 55.690
T18	nd	nd	>500
T19	nd	nd	>1000
T20	nd	nd	>100
T23	nd	nd	253.114 ± 28.629
T24	5.131 ± 0.845	>500	58.29 ± 58.287
T31	>125	>500	>100
T32	17.761 ± 0.244	>500	>100
T33	9.629 0.469	>500	14.759 ± 4.271
T34	6.830 ± 2.961	>500	>10
T37	5.405 ± 0.239	>500	66.177 ± 11.787
T39	24.661 ± 10.495	91.188 ± 17.949	77.713 ± 5.232
T40	nd	nd	>100
T41	nd	nd	>500
T42	nd	nd	124.537 ± 4.426
T44	>10	>500	9.393 ± 2.568
T45	nd	nd	115.237 ± 15.526
T46	nd	nd	765.217 ± 55.691
T47	nd	nd	359.822 ± 96.459
T48	nd	nd	140.631 ± 14.725
T49	nd	nd	58.935 ± 12.124

Cpd	IC$_{50}$ (μM)		
	Enzyme assay	PEcΔTF-TgDT	PEcΔTFTolC-TgDT
T50	12.675 ± 2.908	65.682 ± 4.579	22.143 ± 3.191
T51	8.956 ± 0.442	72.272 ± 2.485	5.633 ± 0.258
T52	12.394 ± 2.457	108.441 ± 17.899	64.350 ± 5.712
T53	4.050 ± 0.061	>500	16.480 ± 1.325
T54	3.782 ± 0.611	139.373 ± 20.210	32.937 ± 5.111
T55	2.467 ± 0.361	85.840 ± 11.425	23.667 ± 2.366
T57	0.196 ± 0.020	>500	19.740 ± 2.148
T58	0.578 ± 0.043	>500	29.200 ± 3.863
T59	nd	nd	>1000
T61	nd	nd	194.445 ± 25.036
T64	nd	nd	137.466 ± 9.140
T65	11.146 ± 1.593	nd	185.292 ± 38.848
T66	nd	nd	250.694 ± 39.729
T67	5.222 ± 0.353	nd	19.468 ± 4.489
T68	6.275 ± 3.018	nd	20.774 ± 3.077
T69	2.012 ± 0.531	nd	8.602 ± 1.955
T74	17.047 ± 9.090	nd	49.019 ± 7.451
T75	9.840 ± 0.917	nd	32.925 ± 6.039
T77	nd	nd	168.058 ± 22.329
T78	4.883 ± 0.839	nd	21.069 ± 3.374
T79	1.022 ± 0.027	nd	77.389 ± 20.556
T80	4.445 ± 0.648	nd	51.332 ± 7.212
T82	nd	nd	74.387 ± 8.310
Triazine scaffold: 3 Rigid compounds			
CYC	0.623 ± 0.0835	>500	>500
C030	1.011 ± 0.1379	>1000	>1000
C088	0.338 ± 0.0305	>1000	750.563 ± 38.667
Triazine scaffold: 15 Flexible compounds			
WR	0.163 ± 0.014	7.602 ± 1.368	0.387 ± 0.114
C420	0.110 ± 0.035	nd	0.913 ± 0.168
C423	0.095 ± 0.024	nd	2.030 ± 0.148
C425	0.073 ± 0.012	nd	2.210 ± 0.358
C448	0.080 ± 0.011	nd	0.140 ± 0.010

Cpd	IC$_{50}$ (μM)		
	Enzyme assay	PEcΔTF-TgDT	PEcΔTFTolC-TgDT
C452	0.120 ± 0.007	nd	0.530 ± 0.123
C459	0.474 ± 0.036	nd	>1000
C460	2.033 ± 0.438	nd	100.300 ± 16.827
C461	6.831 ± 0.177	nd	772.877 ± 96.399
C462	1.990 ± 0.435	nd	131.043 ± 12.525
C472	0.296 ± 0.036	nd	396.103 ± 64.062
C473	0.486 ± 0.031	nd	212.903 ± 63.525
C474	0.606 ± 0.065	nd	228.613 ± 30.800
C475	0.597 ± 0.047	nd	492.727 ± 46.766
C496	0.180 ± 0.024	nd	244.803 ± 41.831
Quinazoline scaffold: 18 compounds			
Q01	0.259 ± 0.020	nd	70.553 ± 8.269
Q02	0.271 ± 0.009	nd	37.213 ± 6.804
Q03	1.345 ± 0.298	nd	45.477 ± 2.514
Q04	0.979 ± 0.166	nd	35.210 ± 5.039
Q05	>100	nd	20.090 ± 2.537
Q06	0.259 ± 0.030	nd	4.370 ± 0.747
Q07	0.564 ± 0.195	<250	9.139 ± 2.328
Q08	0.215 ± 0.007	<1000	13.284 ± 3.076
Q09	0.542 ± 0.041	nd	9.377 ± 2.296
Q10	>100	<1000	22.233 ± 6.611
Q11	0.170 ± 0.007	nd	1.743 ± 0.195
Q13	0.213 ± 0.032	nd	8.630 ± 1.239
Q14	0.191 ± 0.019	nd	2.703 ± 0.771
Q15	1.267 ± 0.214	nd	36.643 ± 10.286
Q16	>100	nd	9.25 ± 9.253
Q18	0.254 ± 0.036	nd	6.157 ± 1.487
Q19	>100	nd	13.083 ± 0.779
P062	>500	174.777 ± 37.158	310.316 ± 20.302
S1 analog			
T63	nd	nd	>100

nd: not determined

9. Output (Acknowledge the Thailand Research Fund)

9.1 International Journal Publication

Vanichtanankul J, Yoomuang A, Taweechai S, Yuvaniyama J, Tarnchompoo B, Yuthavong Y and Kamchonwongpaisan S. Structural insight into effective inhibitors binding to *Toxoplasma gondii* dihydrofolate reductase thymidylate synthase. Manuscript in preparation

9.2 Application

Bacterial surrogate system and *T. gondii* cell-based assay established in this study will be useful for other researcher to identify effective inhibitors leading to development novel drug.

9.3 Others e.g. national journal publication, proceeding, international conference, book chapter, patent

International conference

1. Vanichtanankul J, Yoomuang A, Taweechai S, Suwanakitti N, Tarnchompoo B, Yuthavong Y and Kamchonwongpaisan S. Evaluation of antifolates against Toxoplasma gondii dihydrofolate reductase. The 7th AOHUPO Congress/9th PST International Symposium. August 6-8, 2014. Abstract PP368-147, page 215.

2. Vanichtanankul J, Suwanakitti N, Yoomuang A, Taweechai S, Tarnchompoo B, Yuthavong Y, and Kamchonwongpaisan S. Enhanced efficiency of antifolates against Toxoplasma gondii dihydrofolate reductase by efflux pump inhibition through bacteria model. The 5th International Biochemistry and Molecular Biology Conference. May 26-27, 2016. APC127, page 157.

Thai petty patent

Vanichtanankul J, Taweechai S, Suwanakitti N, Yoomuang A, Talawanich Y, Tarnchompoo B, Saeyang T and Yuthavong Y. Kamchonwongpaisan S. ระบบทดสอบสารแอนติโฟเลตยับยั้งท็อก โซพลาสม่า กอนดิไอ โดยใช้แบคทีเรีย อีโคไล พร่องยีนไธเอ ยีนโฟลเอ และยีนโทลซี เป็นเซลล์เจ้าบ้าน. Thai petty patent No. 13652. Mar 9, 2018.