

Interdomain Twists of Human Thymidine Phosphorylase and its Active-Inactive Conformations: Binding of 5-FU and its Analogues to hTP vs. DPD

Tiffany Tozer¹, Kali Heale¹, Caroline Manto Chagas¹, Andre Luis Branco de Barros² and Laleh Alisaraie^{1,3,*}

1. School of Pharmacy, Memorial University of Newfoundland, 300 Prince Philip Dr., A1B 3V6, St. John's, Newfoundland, Canada

2. Department of Clinical and Toxicological Analysis, Universidade Federal de Minas Gerais (UFMG), 6627 Presidente Antônio Carlos Avenue, Belo Horizonte, Minas Gerais, Brazil

3. Department of Chemistry, Memorial University of Newfoundland, A1B 3X7, St. John's, Newfoundland, Canada

**Corresponding author*

Laleh Alisaraie

School of Pharmacy, Memorial University of Newfoundland, 300 Prince Philip Dr., A1B 3V6, St. John's, Newfoundland, Canada

Phone: +1 (709) 864 2734

Fax: +1 (709) 864 4819

Email: laleh.alisaraie@mun.ca

Abstract

5-Fluorouracil (5-FU) is an anticancer drug, which inhibits human thymidine phosphorylase (hTP) and plays a key role in maintaining the process of DNA replication and repair. It is involved in regulating pyrimidine nucleotide production, by which it inhibits the mechanism of cell proliferation and cancerous tumor growth. However, up to 80% of the administered drug is metabolized by dihydropyrimidine dehydrogenase (DPD). This work compares binding of 5-FU and its analogues to hTP and DPD, and suggests strategies to reduce drug binding to DPD to decrease the required dose of 5-FU. An important feature between the proteins studied here was the difference of charge distribution in their binding sites, which can be exploited for designing drugs to selectively bind to the

hTP. The 5-FU presence was thought to be required for a closed conformation. Comparison of the calculation results pertaining to unliganded and liganded protein showed that hTP could still undergo open-closed conformations in the absence of the ligand, however, the presence of a positively charged ligand better stabilizes the closed conformation and rigidifies the core region of the protein more than unliganded or neutral liganded system. The study has also shown that one of the three hinge segments linking the two major α and α/β domains of the hTP is an important contributing factor to the enzyme's open-close conformational twist during its inactivation-activation process. In addition, the angle between the α/β -domain and the α -domain has shown to undergo wide rotations over the course of MD simulation in the absence of a phosphate, suggesting that it contributes to the stabilization of the closed conformation of the hTP.

1. Introduction

Human Thymidine Phosphorylase (hTP) or platelet-derived endothelial cell growth factor (PD-ECGF) is involved in pyrimidine nucleoside metabolism, affecting tumour growth and angiogenesis (Norman et al., 2004). 5-FU is widely used for treatment of colon, breast, pancreatic, stomach and cervical cancer; however, similar to most anticancer drugs and their metabolites (Chagas, Moss, & Alisaraie, 2018; Flynn, Heale, & Alisaraie, 2017) 5-FU causes adverse effects. In addition, up to 80% of the administered drug is metabolized by dihydropyrimidine dehydrogenase (DPD) (Longley, Harkin, & Johnston, 2003).

In addition to the key role of thymidine phosphorylase in cancer development and platelet activation *in vitro*, hTP affects thrombosis *in vivo* through several signaling pathways and therefore it is considered an important target in cardiovascular drug discovery research (W. Li & Yue, 2018). hTP plays a major part in maintaining the process of DNA replication and repair by its involvement in regulating pyrimidine nucleotide production (Heldin, Usuki, & Miyazono, 1991; Mitsiki et al., 2009), and thus its inhibition affects the mechanism of cell proliferation and cancerous tumor growth.

The structure of hTP has been found to contain four molecules in the form of two asymmetric dimers. One monomer of hTP consists of an α -domain as well as an α/β -domain each connected by three loops that facilitate conformational changing of the enzyme from its closed to open conformations.

The α -domain is comprised of six α -helices ($\alpha 1$ - $\alpha 4$, $\alpha 8$ - $\alpha 9$) making up one side of the substrate binding site as well as the dimer interface (El Omari et al., 2006). The α/β -domain is comprised of a mixed β -

sheet (β 1- β 5, β 13) surrounded by 10 α -helices (α 5- α 7, α 10- α 18), and two antiparallel β -sheets (β 6, β 9, β 9, β 11 and β 8, β 10, β 12) (Norman et al., 2004). The active site of hTP is at the interface of the α - and α/β -domain.

The crystal structure can be classified into the space group of $P2_1$ (El Omari et al., 2006). In this paper, helices and β -strands are numbered according to the numbering by Norman *et al.* (Norman et al., 2004) who determined the structure of hTP in its closed conformation bound with the molecular inhibitor 5-chloro-6-[1-(2-iminopyrrolidinyl) methyl] uracil hydrochloride (CMU).

hTP catalyzes the reaction converting thymidine to thymine. It has been found that phosphate binds first, followed by thymidine. Thymine is, however, the first to leave the site followed by deoxyribose-1 phosphate (Krenitsky, 1968; Schwartz, 1971). (**Figure 1**)

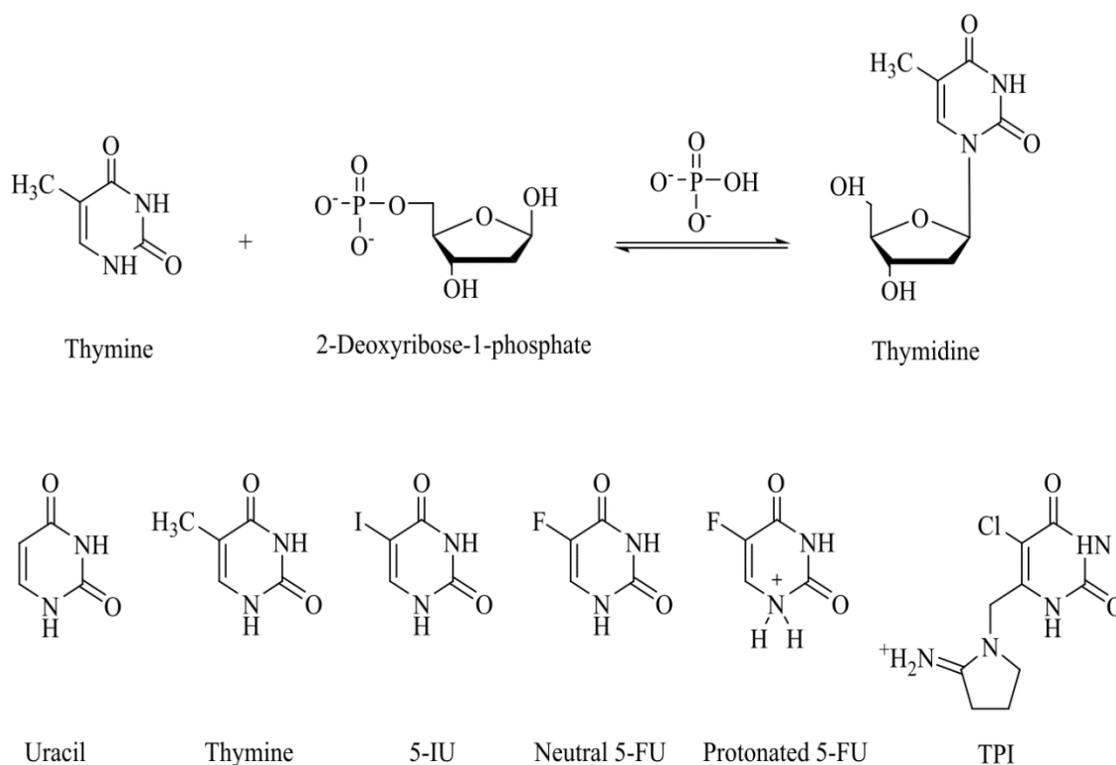


Figure 1: Chemical reaction of thymine and 2-deoxyribose-1-phosphate with phosphate producing thymidine catalyzed by hTP. Chemical structures of uracil and thymine analogues studied here include neutral and protonated 5-FU. TPI is thymidine phosphorylase inhibitor (5-chloro-6-[1-(2-iminopyrrolidinyl) methyl] uracil hydrochloride)

Binding of a pyrimidine base is thought to result in 20° rotation to form the fully closed active site of hTP (Pugmire & Ealick, 1998). The closed (active) conformation in the crystal structure of hTP bound to CMU has been solved in the absence of phosphate and without a His150 and Ala239 hydrogen bond (Norman et al., 2004).

All active site residues are conserved between hTP, *Escherichia coli* thymidine phosphorylase (EcTP) and pyrimidine nucleoside phosphorylase found in *Bacillus stearothermophilus* (BsPYNP). However, Phe210 in EcTP and Phe207 in BsPYNP, are exceptions and both correspond to Val241 in hTP. A conserved glycine rich loop in the enzymes (residues 144-154 of hTP) forms key hydrogen bonds and is thought to play a role in stabilization of the closed (active) conformation (El Omari et al., 2006).

The pKa of 5-FU is estimated to be ~ 8.10 at 10 °C (Parker & Stivers, 2011), and 7.68 at 25 °C (Kremer, Mikita, & Beardsley, 1987). Thus, it is likely that the ligand binds to the hTP in its protonated form and as studied here, since the technical difficulties in protein crystallography, used for solving the structure of 5-IU and hTP, are limiting factors in detecting hydrogen atoms.

The crystal structure of 5-FU and some of its analogues with the hTP and DPD are not available; hence, we have examined their binding interactions with the hTP and DPD by means of molecular docking, and 5-FU in both neutral and protonated forms by molecular dynamics (MD) simulations. The molecular docking experiments were carried out to obtain the initial binding mode of the ligand. The open-closed and conformational twists of hTP have been thought to correspond to its inactive-active modes, which were studied by MD simulations and Normal Mode Analysis (NMA). This allows evaluation at the atomistic level as well as large-scale conformational movements. This was to answer the open question concerning the requirement of ligand for a closed conformation. This is seen in the crystal structure of 5-IU, which could be specific to its favored crystal packing and may be different in solution. To screen the conformational changes of the protein, its unliganded and liganded monomer and dimer were immersed in a box of water molecules in separate experiments, and evaluated under physiological conditions during MD simulations to investigate conformational changes of the system.

2. Materials and Method

2.1. Preparation of the Starting Structures

hTP homodimer complexed with 5-IU was retrieved from Protein Data Bank (PDB) with accession code of 2WK6 (Mitsiki et al., 2009). The unsolved 3D structure of the missing residues in A and B

chain of the dimer (residues 1-34, 482) were constructed according to the sequence data (Uniprot accession code of hTP *Homo sapiens*, P19971) and merged with its corresponding incomplete X-ray structure (2WK6) (Mitsiki et al., 2009). The structure of the constructed segment was confirmed through Phyre 2 homology modeling search engine (Kelley & Sternberg, 2009). The three dimensional structure of the complete sequence (P19971) resulting from the merged X-ray structure and the modeled N-tail was subjected to energy minimization and molecular dynamics simulation for fold recognition and inspection of any potential change of the protein secondary structure.

A library consisting of 5-FU protonated, 5-FU neutral, uracil, thymine, TPI and 5-IU was built up using SYBYL®-X 2.1.1, a commercial molecular modelling and simulation software package developed by Certera™. The ligands were minimized step-wise using Pullman atomic charges and Tripos force field starting from 1.0 kJ/mol to 0.001 kJ/mol energy gradient each through 1000 iterations. The crystalized structure of DPD was retrieved from Protein Data Bank with 1H7X accession code (Doreen Dobritzsch, Gunter Schneider, Klaus D. Schnackerz, & Ylva Lindqvist, 2001).

The side chains and atom charges were calculated, water molecules were removed, and protonation states were corrected for further energy minimization with AMBER7 FF99 force field. The DPD structure (1H7X) was solved with a resolution of 2.0 Å from recombinant pig liver, which shares a high amino acid sequence identity with human DPD with nearly 93% sequence identity (D. Dobritzsch, G. Schneider, K. D. Schnackerz, & Y. Lindqvist, 2001). FlexX commercial docking software (Gohlke, Hendlich, & Klebe, 2000; Rarey, Kramer, & Lengauer, 1999; Rarey, Kramer, Lengauer, & Klebe, 1996) embedded in LeadIT software package (version 2.1.8) was used to dock 5-FU (protonated and unprotonated) and its analogues into the binding site of hTP and DPD in separate experiments, to search for the most energetically favourable binding conformations of each ligand with hTP (Mitsiki et al., 2009).

2.2. Docking

Both hTP and DPD X-ray structure contain a ligand with an unprotonated uracil, and thus the ligand was used as a reference for docking of the ligand library. A spherical region with a radius of 6.5 Å surrounding the reference was selected as the target site for docking into the binding site of hTP or DPD. The output docking solutions were ranked according to the total binding energies ($\Delta G_{\text{binding}}$), calculated based on the Böhm scoring function (Bohm, 1998).

2.3. Standard MD Simulations

Five separate systems were set up for MD simulations of hTP: i. liganded dimer, ii. unliganded dimer, iii. unliganded monomer, iv. monomer complexed with neutral 5-FU and v. monomer complexed with the protonate 5-FU. Complex systems were generated according to the docking of 5-FU into the hTP binding site (Mitsiki et al., 2009). For Molecular dynamics (MD) simulations the topology of neutral and protonated 5-FU were generated using Automated Topology Builder (ATB, version 1.0) (Malde et al., 2011) based on 53A6 force field (Oostenbrink, Villa, Mark, & van Gunsteren, 2004). All MD simulations were performed with GROMACS package (B. Hess, Kutzner, van der Spoel, & Lindahl, 2008; Van Der Spoel et al., 2005), Gromos 96 and 53A6 force field; since it was tested for similar systems and its performance generated accurate structural data in several other simulations of *holo* or *apo* proteins presented in other publications on various research projects, and were comparable with the experimental data (Guvench & D MacKerell, 2008; Z. Li & Alisaraie, 2015; Power, Smith, Downer, & Alisaraie, 2017). Each of the five structures was immersed in a pre-equilibrated solvent box with simple point charges of water molecules. To neutralize the negative charge of the hTP monomer, sodium cations (Na^+) were added to the solvent box for each system. The non-bonded interactions including van der Waals (vdW) and electrostatic interactions were modelled by Lennard-Jones and electrostatic potential, respectively, with a cut-off distance of 1.4 nm for each (Oostenbrink et al., 2004). The Particle Mesh Ewald (PME) algorithm (Darden, York, & Pedersen, 1993) was used to calculate electrostatic interactions contributing to energies and forces of the systems. LINear Constraint Solver (LINCS) (Berk Hess, Bekker, Berendsen, & Fraaije, 1997) was used to constrain all bond lengths including hydrogen bonds. Each simulation was first energy minimized using the steepest descent method followed by 500.0 ps all-bonds position restraint. The time constant for pressure coupling was set to 1.0 ps and compressibility was set to $4.5 \times 10^{-5} \text{ bar}^{-1}$. The temperature was set to 300 K and temperature coupling to 0.1 ps. The constraints were gradually released during the MD simulations. The root mean square deviation (RMSD) of the protein backbone atoms of hTP in each system was monitored at 300 K for MD trajectories of 100.0 ns. The MD experiments were repeated or extended up to 500 ns (e.g. the unliganded monomer, liganded and unliaganded dimer extended) for better conformational samplings and assessing reoccurrence of the events seen in the first round of the MD simulation in each experiment. All simulations converged before 40.0 ns, and thus frames after 40.0 ns of the trajectories were analyzed throughout the study.

(Figure S1)

2.4. Normal Mode Analysis

GROMACS 5.1.4 double precision was used for Normal Mode Analysis (NMA) (Bahar, Lezon, Bakan, & Shrivastava, 2010) to observe the large-scale conformational changes of the protein. The eigenvectors values of a Hessian matrix were calculated following a thorough minimization of the monomer structure using leap-frog algorithm (Guvench & D MacKerell, 2008), coulomb and van der Waals modifiers with the Verlet list scheme. An ensemble of conformations was generated and scaled to Cartesian coordinates for data analysis with a mass-weighted process. MD simulations were performed on the Compute Canada high performance computer clusters (ACENET, WestGrid and Graham).

3. Results and Discussion

The amino acid sequence of hTP (2WK6) (Mitsiki et al., 2009) was compared to BsPyNP (1BRW) (Pugmire & Ealick, 1998) and EcTP (1AZY) (Pugmire, Cook, Jasanoff, Walter, & Ealick, 1998) using Clustal Omega tool (Kelley & Sternberg, 2009) which showed EcTP and BsPYNP share approximately 40-50% sequence-similarities to hTP. The residues 1-32 of the N-terminus of hTP were not present in BsPyNP or EcTP and also their secondary structure was not solved in 2WK6.pdb (Mitsiki et al., 2009). **(Figure S2)**

hTP residues from Val113 to Gly237 were highly conserved, and this region encompasses a glycine loop (residues 144-154) as well as the binding site of uracil and catalytic phosphate (Walter et al., 1990). All key residues including those directly involving in binding were conserved except for Val241, which was phenylalanine in BsPyNP and EcTP.

The glycine loop (residues 144-154) is 100% conserved among the three proteins. His116, Arg202, Ser217, and Lys221 of the loop are involved in hydrogen bonding with 5-IU according to the crystal structure (Mitsiki et al., 2009). Similar interactions were reported in the crystal structure of the hTP complex with uracil derivative, 4-chloro-6-(1-(2-iminopyrrolidinyl) methyl uracil (CMU) (Norman et al., 2004), as well as in hTP and thymine (2J0F) (El Omari et al., 2006), or in BsPyNP and uracil complex (1BRW) (Pugmire & Ealick, 1998).

Up to five hydrogen bonds were recognized in the region throughout the MD trajectory of hTP-5-FU complex. However, only one hydrogen bond was reported to form between 5-chloro-6-[1-(2-iminopyrrolidinyl) methyl] uracil hydrochloride (CMU) and a water molecule with the imino group of the CMU (Norman et al., 2004). **(Figure S3)**

Both neutral and protonated 5-FU were docked into the closed conformation of hTP. The protonated 5-FU was buried and covered by majorly negative electrostatic potential area of the binding site. (**Figure 2**)

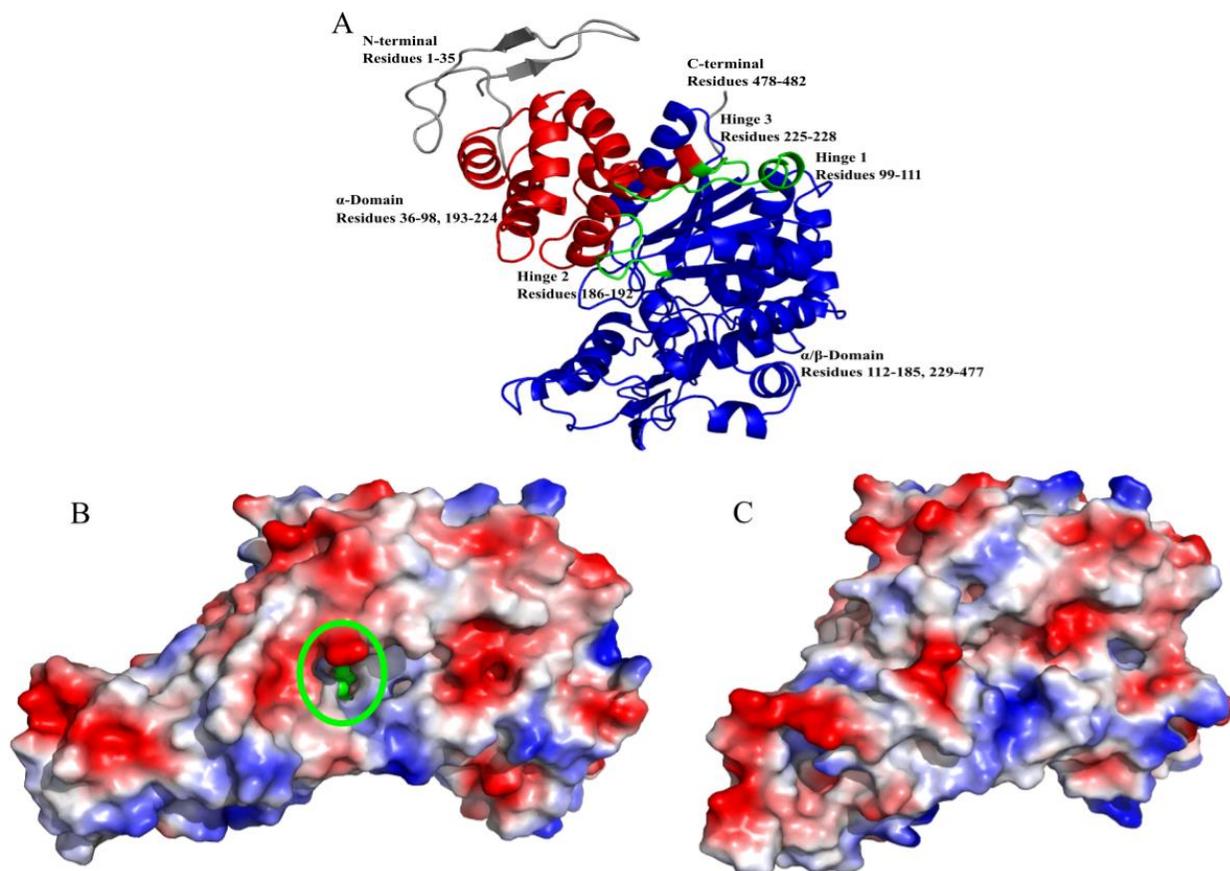


Figure 2: **A)** Monomeric hTP with α -domain and α/β -domain labeled (red and blue respectively) with three hinges (green). Residues 95-112 (region A), residues 159-182 (region B), residues 199-218 (region C), residues 382-388 (region D) and residues 405-415 (region E), **B)** In the monomer complexed with neutral 5-FU in an open conformation, the ligand (green sphere, also circled with green line) can be visualized in the binding site compared to **C)** the monomer complexed with protonated 5-FU in a closed conformation, in which the ligand is not visible, covered by mainly negative electrostatic potential in the interior region of the binding site. Positive and negative electrostatic potential (blue and red respectively), neutral (white).

3.1. *The Open-Closed Conformations*

By investigating five regions with significant differences in the root mean square fluctuation (RMSF) of the residues, the occurrence of open or close conformations of the unliganded, neutral or protonated ligand-enzyme complex was evaluated: **(Table 1 & Table 2)**

Region A consisted of residues 95-112, which encompassed hinge 1 (Norman et al., 2004). **(Figure 2A)**

In the unliganded protein, Gly99 and Glu106 showed a slight increase in fluctuation of the unliganded monomer, and the monomer complexed with protonated 5-FU. During the MD simulation, these residues fluctuated towards each other and away. Broad peak fluctuations ranging from 95-101 (centred at residue Gln97) in the monomer complexed with neutral 5-FU were due to the coil in this region undergoing a pinching motion in which residue Gln97 is apex. This motion contributed to the open-close movement of hTP. **(Figure 3)**

Region B was comprised of residues 159-182 (in α/β domain). Part A in this region showed increased fluctuation in the unliganded dimer compared to the other three systems (i.e. liganded-monomer, protonated-ligand in monomer, and dimer). **(Figure 2A)**

The unliganded dimer and the monomer complexed with protonated 5-FU showed fluctuations with defined peaks at Gly163, Glu171, Val175, Gln179.

This region was not adjacent to the binding site nor was it proximal to the interface between chains in the dimer; hence, no significant difference was expected among the five systems (i.e. monomer liganded (protonated, unprotonated), unliganded, monomer or dimer). **(Figure 3 & Figure 5)**

Region C with residues 199-218 in α -domain, consisted of an area with increased fluctuations in the unliganded monomer, and in the complexed hTP with neutral 5-FU. The unliganded monomer showed a broad peak in RMS fluctuation centred at residue Val208. This residue was the apex of a loop made of Arg202-Ser210 between helix 8 (Pro193-Thr205) and helix 9 (Leu211-Val224) in the α -domain. **(Figure 2A)**

Region D consisted of residues 382-388. All systems showed a peak in RMS fluctuation at residue Pro384 (liganded systems) or Ala385 (unliganded systems) and the unliganded systems showed increased fluctuation over the liganded systems, indicating the latter was more stable as a result of the ligand binding. **(Figure 3A & Figure 4)**

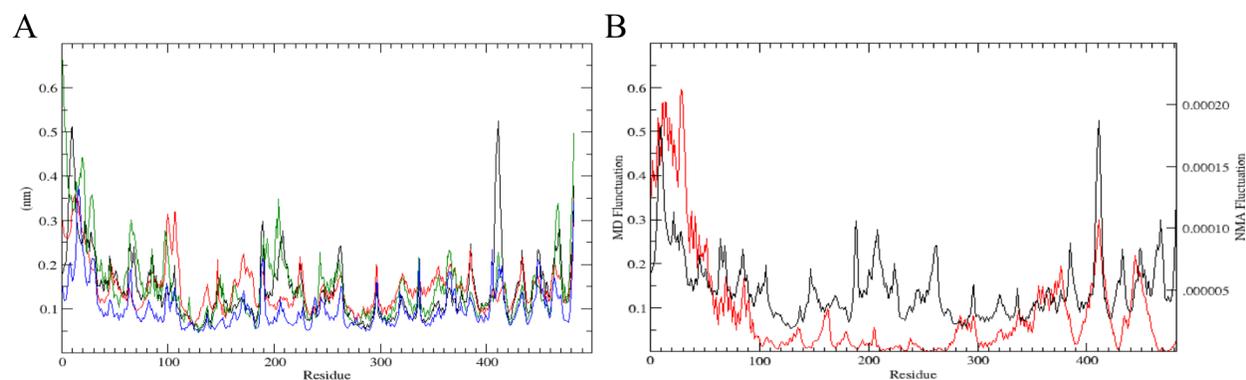


Figure 3: A) The root mean square fluctuation (RMSF) after convergence (at 40 ns) of the backbone atoms of hTP of the unliganded monomer (black), unliganded dimer chain A (red), RMSF of monomer complexed with neutral 5-FU (green), and monomer complexed with protonated 5-FU (blue) obtained from MD simulations. **B)** RMSF plots from the NMA (red) and MD (black) data of the unliganded monomer are overlaid. Region A consisted of residues 95-112, region B residues 159-182, region C residues 199-218, region D residues 382-388 and region E consists of residues (405-415).

Interestingly the RMSF of the B and C regions, including the binding site, showed more stabilization upon binding of the protonated ligand over that of the neutral. The protonated ligand bound protein presented more stabilisation of the binding site region than in both unliganded monomer and the dimer. As it can be expected, the RMSF graph of the neutral ligand bound dimeric system and monomeric systems have very similar fluctuation patterns. **(Figure 3 and Figure S4)**

Region E comprised of residues (405-415). The RMSF plot of the unliganded monomer showed the most visually significant peak spanning residues 405-415 with an apex at residue Ala411. The absence of this large peak in the dimer and liganded systems suggested that dimerization and ligand binding in closed active conformation caused long lasting interactions that stabilized the segment. As well, this segment in PYNP contributed significantly to the stabilization of the closed conformation of the liganded subunit (Pugmire & Ealick, 1998). In hTP, residues 65-70 correspond to the residues of PYNP that participate in the interactions observed (Pugmire & Ealick, 1998). **(Figure 3B & Figure 4)**

Similar to the RMSF graph of the MD simulations, the stabilizing segment showed a significant peak in the RMSF graph of the NMA. This data also confirmed that this region (residues 405-415)

stabilized the closed conformation of hTP as it showed significant fluctuation in the open-open twist movement depicted by NMA of the unliganded system. (Figure 3 to Figure 5)

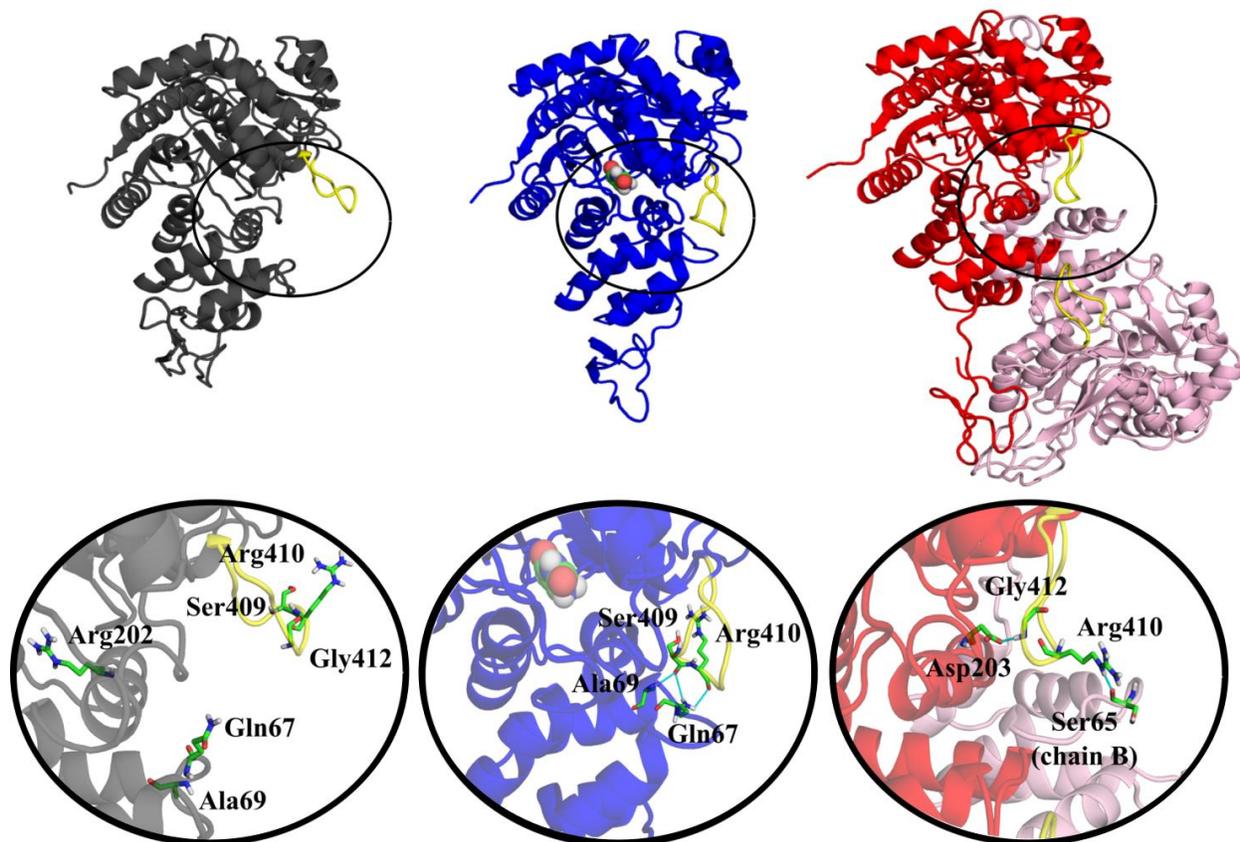


Figure 4: Stabilizing segment (residues 405-415) in yellow of unliganded monomer (grey), monomer liganded with protonated 5-FU (blue), and unliganded dimer (red) and chain B (pink) with intra- and inter-chain interactions (cyan). The ligand 5-FU (sphere representation). Atoms: nitrogen (blue), hydrogen (white) and oxygen (red) carbon (green).

NMA of the unliganded monomer showed larger scale movements during the open-close conformations. There were two major movements that involved sub-domains: i. displacing α -domain, which consisted of the N-terminal to helix 4, and the linking loop between helix 8 (Pro193-Thr205) and helix 9 (Leu211-Val224), and ii. helix 6 (Thr154-Leu158), helix 11 (Ala284-Met294), β 6- β 7 and β 9- β 11 loops of the α / β -domain. (Figure 3 & Figure 5)

Based on the NMA data, the involvement of the three hinges in the open-closed movement was assessed. Hinge 1 appeared to pivot the movement of hTP domains. Hinge 2 and 3 and the

segments in their close proximities on either side showed insignificant movements in the NMA. (Figure 5)

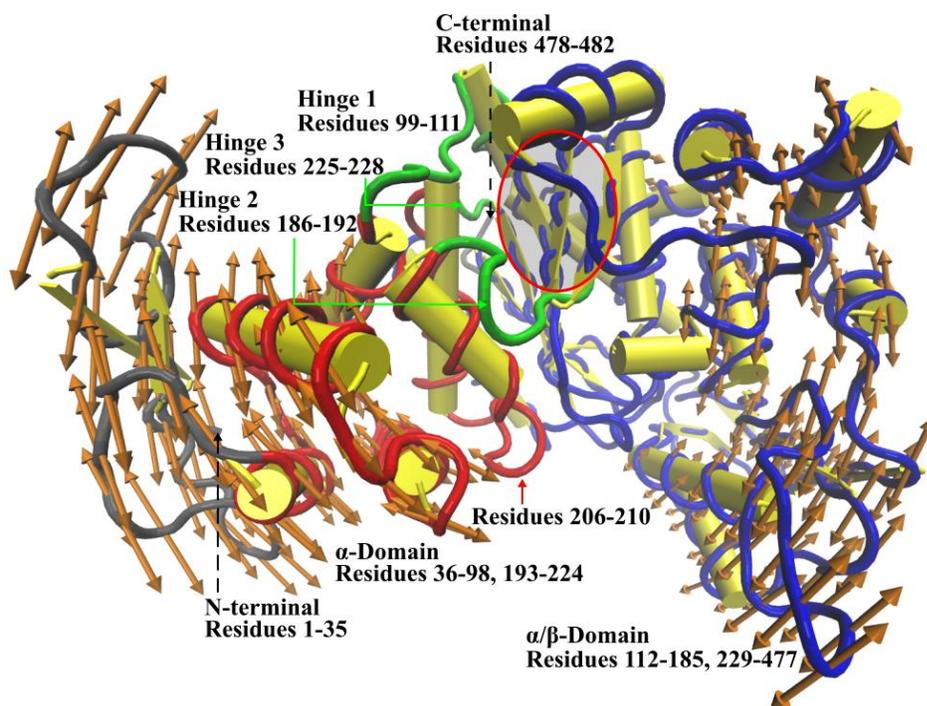


Figure 5: Normal Mode Analysis of hTP (yellow cylinder), α -domain (red), α/β -domain (blue), hinges (green) and N-terminus (residues 1-35, gray). Arrows (orange) demonstrate the open-closed direction of movement of hTP. The binding site was also shown (gray transparent oval). Arrow length is proportional to the magnitude of displacement of the segment of the protein.

3.2. α - and α/β - Domains Twist

The hTP is defined as two major subunits of the monomer, an α -domain consisting of six helices ($\alpha 1$ - $\alpha 4$, $\alpha 8$ - $\alpha 9$) and an α/β -domain ($\beta 1$ - $\beta 5$, $\beta 13$) surrounded by helices ($\alpha 5$ - $\alpha 7$, $\alpha 10$ - $\alpha 18$) and two small antiparallel β -sheets ($\beta 6$, $\beta 9$, $\beta 9$, $\beta 11$ and $\beta 8$, $\beta 10$, $\beta 12$) (Norman et al., 2004). The residues between these two domains consisted of three loops (residues 99-111, 186-192, and 225-229) (Norman et al., 2004). (Table 1)

To assess the structural changes or the twists between the two domains during the MD trajectories, distance between the centres of mass (dCOM) of different parts of the α -domain and α/β -domain as well as the entire domains were measured. (Table 1 & Table 2)

Defined as	Residues
------------	----------

α-domain	35-97, 192-225
α/β-domain	112-187, 229-482
α-domain part A	36-96
α/β-domain part A	112-186
Hinge 1	97-111
α-domain part B	34-45, 66-81
α/β-domain part B	393-405, 447-459

Table 1: Important regions of hTP analyzed

In addition, a plane of each α -domain part A and α/β -domain part A were defined based on the coordination of the centres of mass of the respective sub-domains. Three atoms of each sub-domain were chosen to make a plane, as follows: The plane of α -domain part A was composed of α -carbon of Asp45, Gly64, and Ala94, and the plane of α/β -domain part A was composed of α -carbon of Gly137, His150, and Pro170. The angle between the two planes was measured over the simulation trajectory. In order to account for any expansion of folding at the end of the α - and α/β -domains defined in the X-ray structure, we expanded the residue range by one or two residues for some regions of hTP. **(Table 1 & Figure 6)**

Unliganded monomer					
<i>Group</i>	<i>Maximum Time</i> <i>(ns)</i>	<i>Maximum</i> <i>Distance (nm)</i>	<i>Minimum Time</i> <i>(ns)</i>	<i>Minimum</i> <i>Distance (nm)</i>	<i>Difference (nm)</i>
<i>AD-ABD</i>	63.0	2.86	53.0	2.64	0.22
<i>ADA-ABDA</i>	41.4	2.89	72.2	2.76	0.13
<i>ADB-ABDB</i>	80.0	3.80	53.0	3.17	0.63
<i>Group</i>	<i>Maximum Time</i> <i>(ns)</i>	<i>Maximum</i> <i>Angle (°)</i>	<i>Minimum Time</i> <i>(ns)</i>	<i>Minimum</i> <i>Angle (°)</i>	<i>Difference (°)</i>
<i>ADA-ABDA</i>	81.6	75.3	60.7	46.6	28.7

Unliganded dimer

<i>Group</i>	<i>Maximum Time</i> (<i>ns</i>)	<i>Maximum</i> <i>Distance (nm)</i>	<i>Minimum Time</i> (<i>ns</i>)	<i>Minimum</i> <i>Distance (nm)</i>	<i>Difference (nm)</i>
<i>AD-ABD</i>	80.0	2.79	64.0	2.69	0.10
<i>ADA-ABDA</i>	48.0	3.01	99.0	2.87	0.14
<i>ADB-ABDB</i>	80.0	3.35	83.0	3.23	0.12

<i>Group</i>	<i>Maximum Time</i> (<i>ns</i>)	<i>Maximum</i> <i>Angle (°)</i>	<i>Minimum Time</i> (<i>ns</i>)	<i>Minimum</i> <i>Angle (°)</i>	<i>Difference (°)</i>
<i>ADA-ABDA</i>	54.0	44.0	49.0	30.0	14.0

Monomer complexed with neutral 5-FU

<i>Group</i>	<i>Maximum Time</i> (<i>ns</i>)	<i>Maximum</i> <i>Distance (nm)</i>	<i>Minimum Time</i> (<i>ns</i>)	<i>Minimum</i> <i>Distance (nm)</i>	<i>Difference (nm)</i>
<i>AD-ABD</i>	59.0	2.97	82.0	2.78	0.19
<i>ADA-ABDA</i>	73.0	3.14	82.0	2.90	0.24
<i>ADB-ABDB</i>	59.0	3.85	82.0	3.34	0.51

<i>Group</i>	<i>Maximum Time</i> (<i>ns</i>)	<i>Maximum</i> <i>Angle (°)</i>	<i>Minimum Time</i> (<i>ns</i>)	<i>Minimum</i> <i>Angle (°)</i>	<i>Difference (°)</i>
<i>ADA-ABDA</i>	81.0	59.7	53.0	40.2	19.5

Monomer complexed with protonated 5-FU

<i>Group</i>	<i>Maximum Time</i> (<i>ns</i>)	<i>Maximum</i> <i>Distance (nm)</i>	<i>Minimum Time</i> (<i>ns</i>)	<i>Minimum</i> <i>Distance (nm)</i>	<i>Difference (nm)</i>
<i>AD-ABD</i>	41.0	2.67	63.0	2.55	0.12
<i>ADA-ABDA</i>	43.0	2.94	62.0	2.84	0.10
<i>ADB-ABDB</i>	41.0	2.88	62.0	2.53	0.35

<i>Group</i>	<i>Maximum Time</i> (<i>ns</i>)	<i>Maximum</i> <i>Angle (°)</i>	<i>Minimum Time</i> (<i>ns</i>)	<i>Minimum</i> <i>Angle (°)</i>	<i>Difference (°)</i>
<i>ADA-ABDA</i>	98.0	64.6	43.0	51.1	13.5

Table 2: Comparison of the maximum and minimum dCOM and angle between two planes of α -domain part A (i.e. ADA) and α/β -domain part A (i.e. ABDA) and the respective times during the simulation of unliganded monomer, unliganded dimer part A and part B (ADB, ABDB), monomer complexed with neutral 5-FU, and monomer complexed with protonated 5-FU.

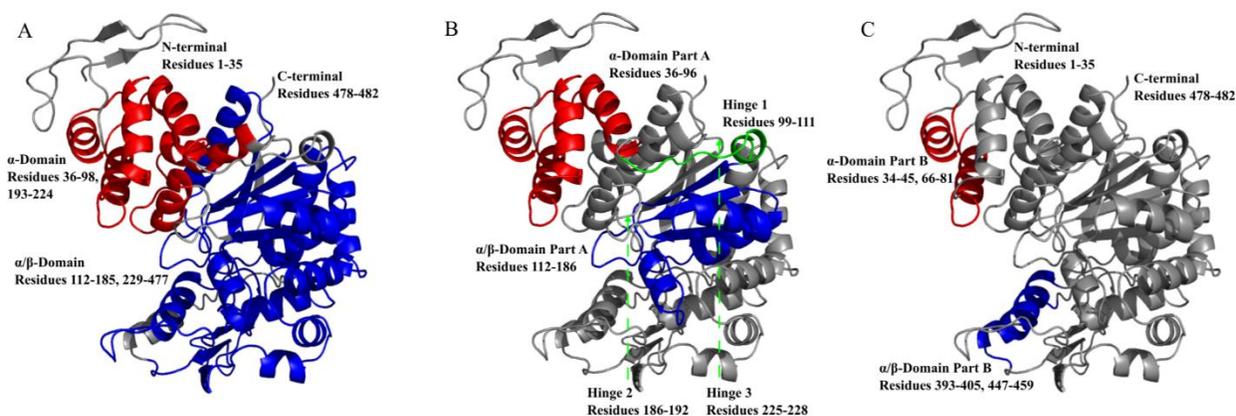


Figure 6: Segment pairs of unliganded monomeric hTP analyzed for the distance between centres of mass and angle between two planes. **A)** α -domain defined by the residues 36-98 and 193-224 (red) and α/β -domain defined by residues 112-185 and 229-477 (blue). **B)** Sub-domains named α -domain Part A (red) defined by residues 36-96 and α/β -domain Part A (blue) defined by residues 112-186. **C)** Sub-domains named α -domain Part B (red) defined by residues 34-45 and 66-81 and α/β -domain Part A (blue) defined by residues 393-405 and 447-459.

3.3. Unliganded hTP

The maximum dCOM of α -domain part A and α/β -domain part A of the unliganded monomer was at ~ 41 ns and the minimum dCOM was at ~ 72 ns. The maximum dCOM and minimum dCOM represent structures of the open (inactive) and closed (active) conformations respectively. **(Figure S5)**

A new turn in hinge part A was expanded in open conformation at the minimum dCOM by three residues (at 41.4 ns helix of residues 106-111; at 72.2 ns helix of residues 109-111) than it was in the closed conformation at maximum dCOM.

Helicity of the new turn (Glu106-Arg109) in hinge 1 region (Gln97-Gln111) and helix 5 (Val125-Ala134) was evaluated. These helices, which are in the proximity of the binding site, participate in a helical conformation. In over 65% of the simulation, the residues 124-126 and 128-133 appeared as a helix. Leu127 contributed to the formation of the helix during only 18% of the simulation, meaning during 72% of this time period, this region consisted of two α -helices ($\alpha 5$ A, $\alpha 5$ B) instead of one helix ($\alpha 5$) of residues 125-134 as referenced in the crystal structures of hTP (Mitsiki et al., 2009; Norman et al., 2004). **(Figure S6)**

The angle between planes (composed of: i. Asp45, Gly64, and Ala94 (α -domain) and ii. Gly137, His150, and Pro170 (α/β -domain)) was also measured, which varied from 35° to 85° . The

minimum angle between these two planes (35°) corresponded to an open or active conformation and the maximum angle (85°) corresponded to a closed or inactive conformation. **(Figure S7)**

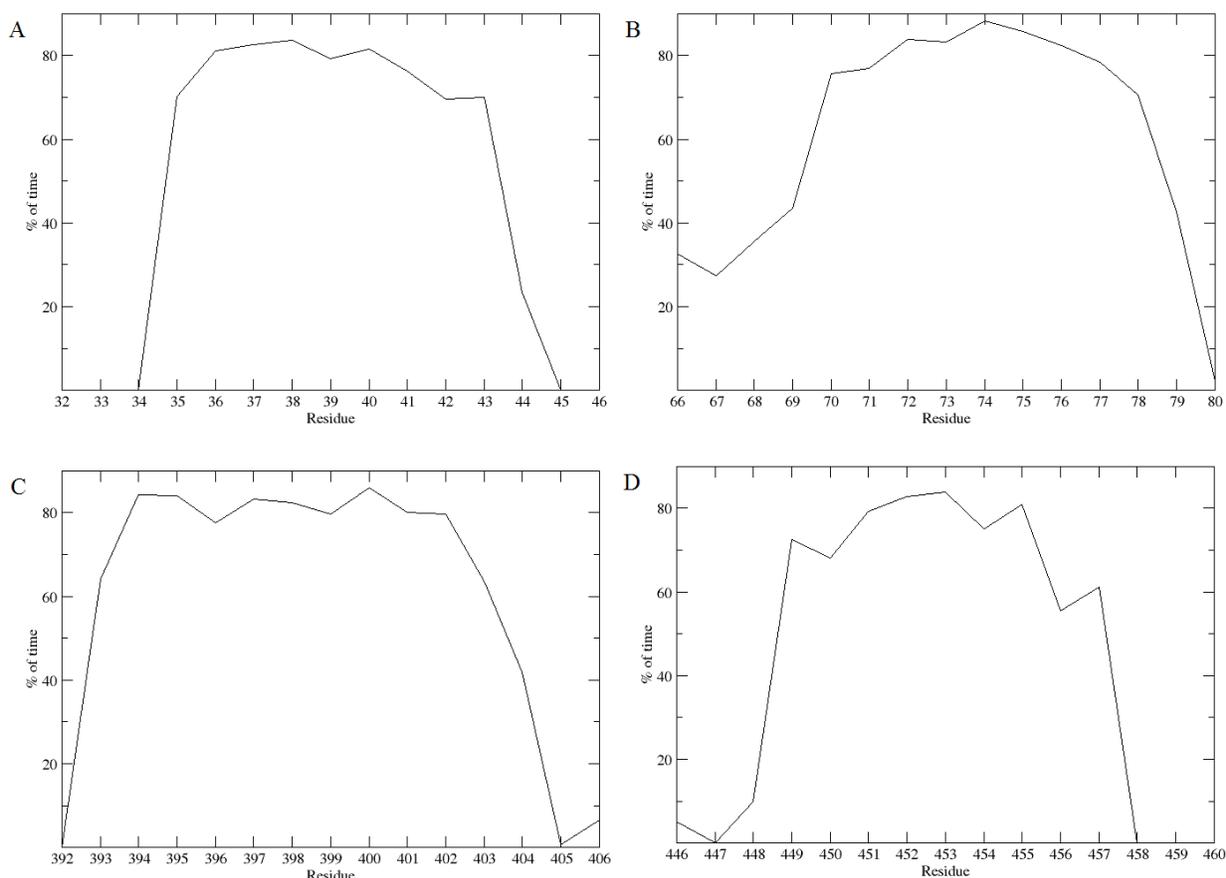


Figure 7: Helicity of α -domain and α/β -domain Part B, **A)** helix 1, **B)** helix 3, **C)** helix 17, and **D)** helix 18 in the unliganded monomer after convergence to 100 ns.

The plots of the time-dependent dCOM and angle between two planes of α -domain part A and α/β -domain part A were overlaid and analyzed. The timeframes of the absolute minimum and maximum dCOM did not correspond to the absolute maximum and minimum angle between two planes. The minimum dCOM was observed at ~ 72 ns (2.76 nm) and the maximum dCOM was at ~ 41 ns (2.89 nm). The maximum angle between two planes was observed at ~ 81 ns (75.3°), and the minimum angle between two planes was at ~ 60 ns (46.6°).

Thus, it is not possible to identify two particular timeframes that best represent the open and closed conformations; however, these still confirm that the open-closed conformations exist in the absence of the ligand, as well. **(Figure S8)**

The RMSF plot of the unliganded monomer was examined for peaks in fluctuation in the α -domain part A and α/β -domain part A regions. **(Figures 2 & Figure 3, Table 1 & Table 2)**

For the α -domain part A region, a peak in fluctuation was defined as a fluctuation greater than 2 Å. Four areas showed such peaks: Asp45-Gly46, Asn63-Gly64, Gln67-Ala69, and Asp84-Glu86. For the α/β -domain part A, the apex of the peaks appeared at residues Gly147, Thr151, and Ser169, respectively. Gly147 is in close proximity to the binding site of 5-FU and Leu148 is part of the hydrophobic pocket. Increased fluctuation in this area could be indicative of movement consistent with the unliganded protein opening and closing. **(Figure S9)**

The dCOM of the α -domain part B and α/β -domain part B was evaluated. The maximum dCOM was at ~80.0 ns, which corresponds to an open (inactive) conformation, and at ~53.0 ns, which corresponds to a closed (active) conformation. **(Figure S10)**

The helicity of all four helices of α -domain part B ($\alpha 1$ and $\alpha 3$) and α/β -domain part B ($\alpha 17$ and $\alpha 18$) was examined. Helix 1 of the α -domain part B was relatively stable with all residues in helix formation greater than 65% of the time. Conversely, the helix 3 appeared to vary in length throughout the simulation. Helix 17 of α/β -domain part B was relatively stable over 60% of the time and helix 18 of α/β -domain part B showed more variability ranging from 55%-65% over the course of simulation. **(Figure 7)**

Further details on the dCOM between α -domain and α/β -domain of the unliganded monomer is available in the supplementary information. **(Figure S11)**

3.4. *Liganded hTP*

The dCOM of α -domain and α/β -domain of the monomer complexed with neutral 5-FU was evaluated over the trajectory. The maximum dCOM (2.97 nm) was observed at 59.0 ns and the minimum dCOM (2.78 nm) was observed at 82.0 ns (versus 41 ns and 72 ns as the respective timeframes in the unliganded system). The structure at maximum dCOM corresponds to the open conformation as the domains are further apart which draws the active site open, while the structure at minimum dCOM corresponds to the closed conformation as the domains are closer together and the active site is closed.

In the dCOM plot of both i) the α -domain and α/β -domain and ii) the α -domain part B and α/β -domain part B the maxima and minimum were observed in the same timeframes, 59.0 ns and 82.0 ns respectively. Four α -helices of part B could represent the movements of the α - and α/β -domains in the monomer complexed with neutral 5-FU.

In the closed conformation helix 8 adjacent to hinge 2 and proximal to the active site involved residues 195-198. This allowed Tyr199 to form a 2.5 Å hydrogen bond with 5-FU. In the open conformation at 82.0 ns, helix 8 is longer (residues 195-203), which causes Tyr199 to be farther from the active site, which weakens the interaction.

With protonated 5-FU, dCOM of α -domain and α/β -domain was 2.67 nm at 41.0 ns corresponding to an open conformation, and 2.55 nm at 63.0 ns corresponding to a closed conformation. A key difference in hinge 1 was observed between the secondary structures of hTP at these conformational states. In the open conformation, a helix was formed at residues 106-111 at 41.0 ns appeared to regulate opening the hinge and contribute to the formation of the inactive status (open). **(Figure S12)**

The unliganded monomer had the greatest difference in angle between two planes (i.e. the plane of α -domain part A, that of α/β -domain). This was to be expected as the difference is representative of the opening and closing of hTP. However, dimerization stabilizes hTP, restricting its open-closed movement; hence, the difference was smaller for the unliganded dimer.

Ligand binding, however, further stabilizes hTP and reduces open-closed movement. The monomer complexed with neutral 5-FU has a greater difference in angle between two planes compared to the monomer complexed with protonated 5-FU which suggests a protonated 5-FU molecule is more tightly bound than a neutral 5-FU molecule. **(Table 2)**

The monomer complexed with protonated 5-FU had the lowest average dCOM (2.55 nm), hence, the protonated 5-FU ligand had the largest effect on stabilizing the proximity of the two domains and dimerization had the second largest effect. Both the monomer complexed with protonated 5-FU and dimer systems showed minimal fluctuation in the dCOM plots, meaning the domains were held relatively stable throughout the simulations. The unliganded monomer showed large fluctuations in both plots, indicating that it was likely switching between the open and closed conformation as it lacked stabilization that could be controlled by dimerization or by a ligand. The monomer complexed with neutral 5-FU also showed a high dCOM of the two α - and α/β -domains and high fluctuation in the dCOM plot. **(Figure S12)**

The MD results indicate that when hTP was liganded with neutral 5-FU, the ligand was loosely held in place and the protein continued to undergo an open-closed conformation while the ligand was bound. In comparison, when hTP was bound to protonated 5-FU, the positive charge held the protein more tightly with a distance of 14.9 Å vs. 31.2 Å at the binding site entrance and restricted the open-close movement. **(Figure 8)**

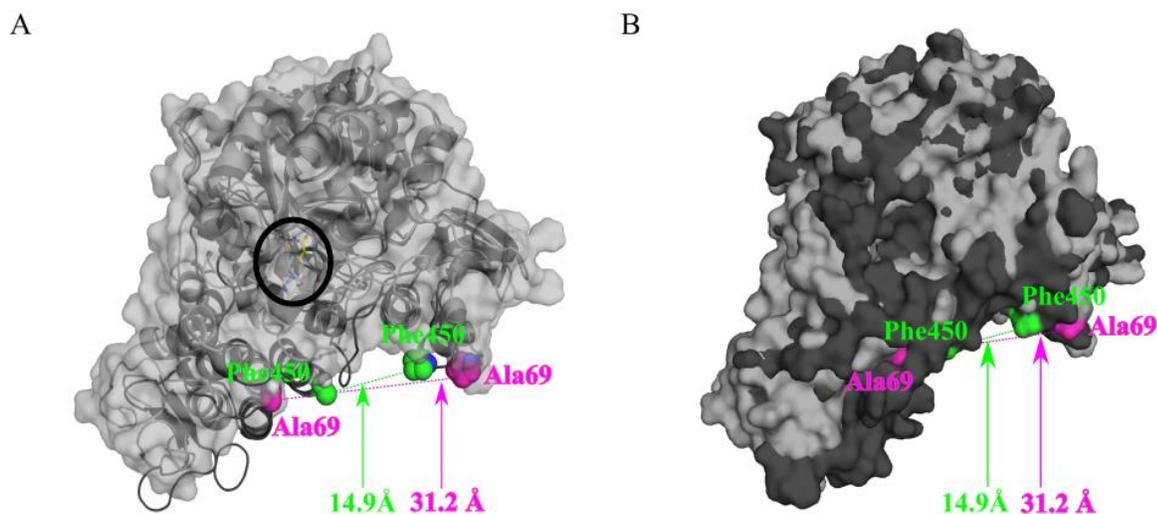


Figure 8: The hTP complexed with protonated 5-FU (black) conforms more tightly to its ligand compared to the neutral 5-FU (grey). This is demonstrated by comparing the hTP structure at the minimum dCOM of α -domain and α/β -domain complexed with neutral 5-FU and complexed with protonated 5-FU (63.0 ns). The distance between Ala69 and Phe450 at the entrance of the binding site (green with protonated 5-FU and magenta with neutral 5-FU) on either side of the protein's hinge demonstrates the difference between the two conformations (14.9 Å vs. 31.2 Å).

3.5. Glycine-Rich Loop

The glycine loop (residues 144-154 of hTP) is 100% conserved among the three sequences of TP (i.e. Human TP (2W6K), *Geobacillus sterothemophilus* pyrimidine nucleoside phosphorylase (BsPyNP) (1BRW), and *Escherichia coli* thymidine phosphorylase (EcTP) (1AZY)). The loop was proposed to involve in the catalytic phosphate binding (Walter et al., 1990) in BsPYNP, which resulted in a hydrogen bond between residues His116 and Gly205 of BsPYNP corresponding to His150 and Ala239 of hTP. **(Figure S1 and Figure 9)**

Phosphate binding was thought to be the cause for an 8° rotation of the α/β -domain relative to the α -domain in the closed conformation (Pugmire & Ealick, 1998). In our simulations, water molecules occupied the phosphate-binding site establishing hydrogen bonds that compensated the binding strength for the absence of phosphate ion. The angle between the α/β -domain and the α -domain however underwent rotational twists over the course of simulation all larger than the reported 8° from the X-ray structure (e.g. the unliganded protein from 46.6° to 75.3°) indicating that in the

absence and presence of the catalytic phosphate or a ligand there is a twist between the domains as well as the open-closed conformations. (Table 2)

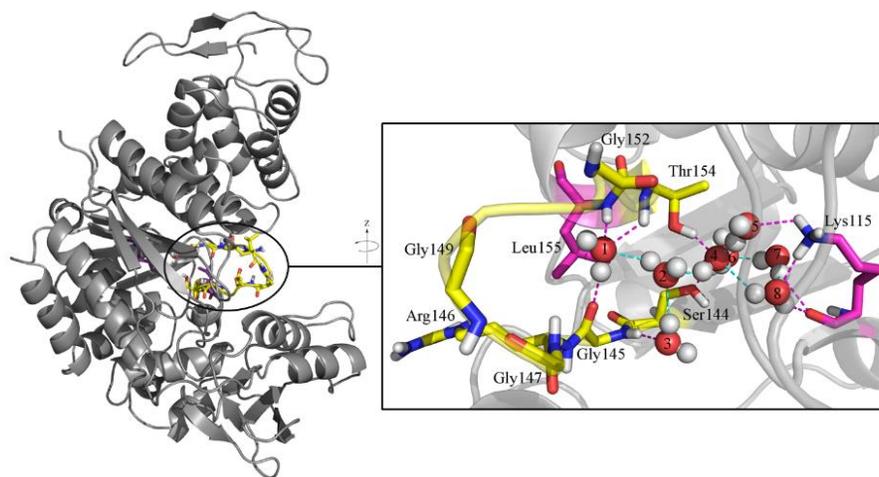


Figure 9: The glycine loop (yellow), non-glycine loop residues (Leu155 and Lys115, magenta) bind to water molecules in the phosphate-binding sites of hTP (grey). Hydrogen bonds (pink and cyan dashed-lines) among the hTP residues and water molecules, or among water molecules the H-bond mediators. Water molecules (spheres) and residues (sticks), hydrogen (white), nitrogen (blue), and oxygen (red).

3.6. *N-terminal Tail*

Residues 1-33 of hTP have not been solved by X-ray crystallography and did not exist in *Geobacillus sterothemophilus* pyrimidine nucleoside phosphorylase (BsPyNP) or *Escherichia coli* thymidine phosphorylase (EcTP). (Figure S1)

This segment was constructed according to the available hTP sequence data and added to the enzyme structures prior to the MD simulations. In the unliganded monomer and the monomer complexed with protonated 5-FU, a short β -sheet composed of two anti-parallel β -strands was observed in this segment throughout most of the MD simulation.

To study the folding of the N-terminal, the simulations of unliganded monomer and dimer (liganded and unliganded) were extended up to 500.0 ns. Throughout the majority of the trajectory, the monomer continued to show a small β -sheet made of β -strands of residues 2-5 and 14-17. This β -sheet lengthened at 350.0 ns involving β -strand residues 1-5 and 14-19.

The RMSF of the protein generated from the MD simulation and NMA were compared. In the N-terminal tail, the NMA-RMSF showed increased fluctuation over the entire segment, while the MD-

RMSF primarily showed a peak in fluctuation in the first 14 residues. This was to be expected, since NMA present collective large-scale movements, whereas MD simulations show individual residues movement on atomistic level. (**Figure 3**)

3.7. *Binding of Uracil Analogues to hTP and DPD*

The position of both neutral and protonated 5-FU varied throughout their MD simulation. The binding modes of the ligand in each status, obtained from docking, was used as the starting structure for the separate MD simulations. The conformation of both docking solutions were very similar to that of 5-IU in its X-ray structure ($\text{RMSD} \leq \sim 0.5 \text{ \AA}$).

The docking solution displays interactions of 5-FU with His116, Leu148, Arg202, Val208, Ile214, Ser217, Lys221, Val241, Lys115, His116, Tyr199, Arg202, Ile214 and Ser217 as very similar to that of 5-IU. The fluorine of 5-FU interacts with oxygen of Thr118, NH₂ of Arg202, and Val 208 and Leu 148. In 5-IU four residues, Leu148, Val208, Ile214, and Val241, stabilize the iodine atom (Mitsiki et al., 2009), and chlorine atom in TPI (Norman et al., 2004) by forming a hydrophobic pocket. Additionally Lys115, His116, Tyr199, Arg202, Ile214 and Ser217 were identified as few key residues by enzymatic activity assays (Mitsiki et al., 2009).

Docking solutions of various analogues of 5-FU including 5-IU, uracil, thymine and 5-chloro-6-[1-(2-iminopyrrolidinyl) methyl] uracil hydrochloride (TPI) exhibit very similar binding modes to the neutral and protonated form of the 5-FU as the overlaid solutions display only minor shift among the ligands binding conformations. (**Figure 10**)

TPI possessed the lowest binding energy -23.87 kJ/mol, because it can additionally interact with Leu148 and Thr118 via its 2-iminopyrrolidinyl fragment, attached to its uracil ring, with Ile 148 and Thr 118. (**Table 3 & Figure 10**)

Due to the possibility of the dynamic movements during MD simulations of 5-FU complex with hTP, the ligand as well as the hTP residues are flexible, thus the ligand adapts to the flexible conformation of the binding site region resulting in different conformations from the initial mode obtained from the rigid docking (i.e. starting structure). However, the ligand revisits the same location as in the crystal structures (El Omari et al., 2006; Mitsiki et al., 2009; Norman et al., 2004) during the MD simulation. 5-FU at the protonated nitrogen creates a hydrogen bond to the polar OH moiety and could interact with aromatic ring of Tyr 199. This explains the result of mutagenesis studies that mutation of Tyr with Phe reduced the activity of the 5-IU, and mutation with Ala resulted in its inactivation (El Omari et al., 2006). Protonated 5-FU forms hydrogen bonds with Ser117 (N), Ser144

(O) Tyr199 (O) and Lys221 (H) and Ser217. At 62.0 ns and 98.0 ns, it was observed that the hydroxyl oxygen of Tyr199 formed two hydrogen bonds with hydrogen atoms of the protonated nitrogen of the ligand. It is worth noting that the rigid docking solutions of the uracil ligands were very close to each other and that 5-IU (both its docking solution and X-ray structure) also validate the docking method and are utilized in this study. The protonated 5-FU shows the least affinity in that conformation compared to the other ligands in the library among which neutral 5-FU presented one of the strongest binding energies to the hTP. Since the MD simulation is more precise and covers a much wider conformational space than the single conformation of the X-ray structure (used for docking), the protonated ligand can be shown to be a stronger stabilizer of the binding site than the neutral form. Thus, the protonated ligand could reduce the fluctuations of the hTP residues in the ligand binding area as discussed earlier on the RMSF graphs.

In the DPD binding site, the docking solution of 5-FU neutral had the lowest binding energy -26.95 kJ/mol, followed by thymine and 5-IU with binding energy of -26.42 kJ/mol and -26.04 kJ/mol, respectively. Except for TPI, the ligands displayed very similar binding conformation, mainly interacting with Asn668, Asn736, Thr737 and Asn609. Protonation of 5-FU affected binding energy of the ligand to the highest value (i.e. low affinity) among the docked ligands (-21.80 kJ/mol).

Additionally, the TPI interacted with Val738, Ser739 and Gln764 via its additional fragment iminopyrrolidinyl. This caused dislocation and rotation of the TPI uracil $\sim 75^\circ$ compared to that of other uracil derivatives. This suggested that the addition of a hydrophobic fragment instead of the imino (polar group) could reduce the affinity of the drug binding to the DPD site. **(Figure 11)**

In addition, the electrostatic surface area of the DPD binding site was mainly positive (Doreen Dobritsch et al., 2001), whereas that of hTP was mainly negative. Thus, stabilization of a positive charge on the uracil analogues (e.g. by addition of a stable quaternary nitrogen) could enhance drug binding in the hTP site over the DPD. **(Figure S13)**

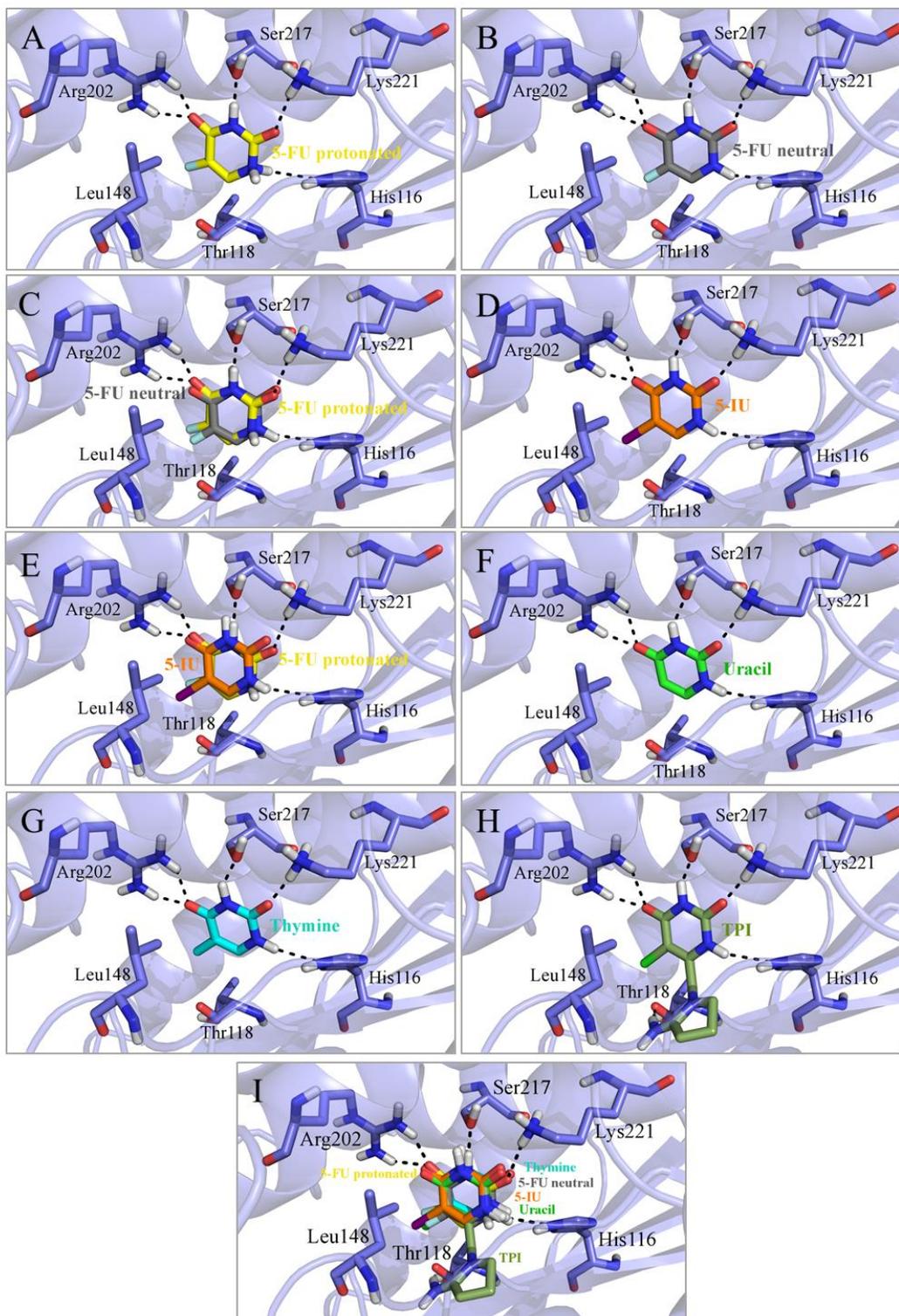


Figure 10: Docking solution and binding interactions of hTP (blue) complexed with **A)** 5-FU protonated (yellow), **B)** 5-FU neutral (grey), **C)** overlay of 5-FU protonated and 5-FU neutral, **D)** 5-IU (orange), **E)** overlay of 5-IU and 5-FU-protonated, **F)** Uracil (lime green), **G)** Thymine (cyan), **H)**

5-chloro-6-[1-(2-iminopyrrolidinyl) methyl] uracil hydrochloride (TPI) (smudge), and **I**) overlay of all ligands. Ligands and residues are shown in sticks; hydrogen bonds are shown in dash lines. Atoms: nitrogen (blue), hydrogen (white) and oxygen (red).

<i>Ligands</i>	$\Delta G_{\text{binding}}$ (kJ/mol)	
	hTP	DPD
<i>5-FU neutral</i>	-22.50	-26.95
<i>5-FU protonated</i>	-19.59	-21.80
<i>5-IU</i>	-19.71	-26.04
<i>TPI</i>	-23.87	-24.56
<i>Thymine</i>	-21.71	-26.42
<i>Uracil</i>	-20.45	-25.57

Table 3: Binding energies of 5-FU neutral, 5-FU protonated, 5-IU, TPI, thymine and uracil docked with hTP and DPD receptors.

In the hTP, the aromatic ring of Tyr199 is in the perpendicular position to the 5-FU ring. There is a possibility of a weak H-bond interaction between the extra hydrogen of the protonated 5-FU and the hydroxyl group of Tyr199 that is 3.2 Å apart. **(Figure S14)**

This weak interaction slightly moves the 5-FU from its position, as compared to that of the neutral ligand. The displacement results in the weakening of the interactions of the protonated 5-FU with other residues (i.e. His116, Thr118, Leu148, Arg202, Ser217 and Lys221). **(Figure 10)**

Since the relatively far-distance H-bond through the extra hydrogen of the protonated form, is not strong enough, it is unable to compensate for the energy lose due to the ligand dislocation. Thus, the overall binding energy of the protonated form is higher (less negative) than the other ligands.

In DPD, there is Asn609 in the proximity of the protonated hydrogen of the 5-FU. While the hydrogen of the neutral 5-FU is also able to form H-bond with the carbonyl oxygen of Asn609, the existence of two hydrogen atoms on the protonated 5-FU, both being capable of hydrogen bonding, causes the ligand to adapt a position where both of its hydrogen atoms can interact with the oxygen of the Asn. **(Figure S14)**

Despite an additional, however, weak H-bond interactions with both of the hydrogen atoms on the protonated 5-FU, the ligand finds an optimum position, which costs it some shifts and lose of

strong interactions with other residues (i.e. Asn668, Asn736 and Thr737) involving the neutral 5-FU or its analogues lacking the proton. This could result in slightly weakening binding energy of the protonated 5-FU as compared to the unprotonated ligands. **(Figure 11)**

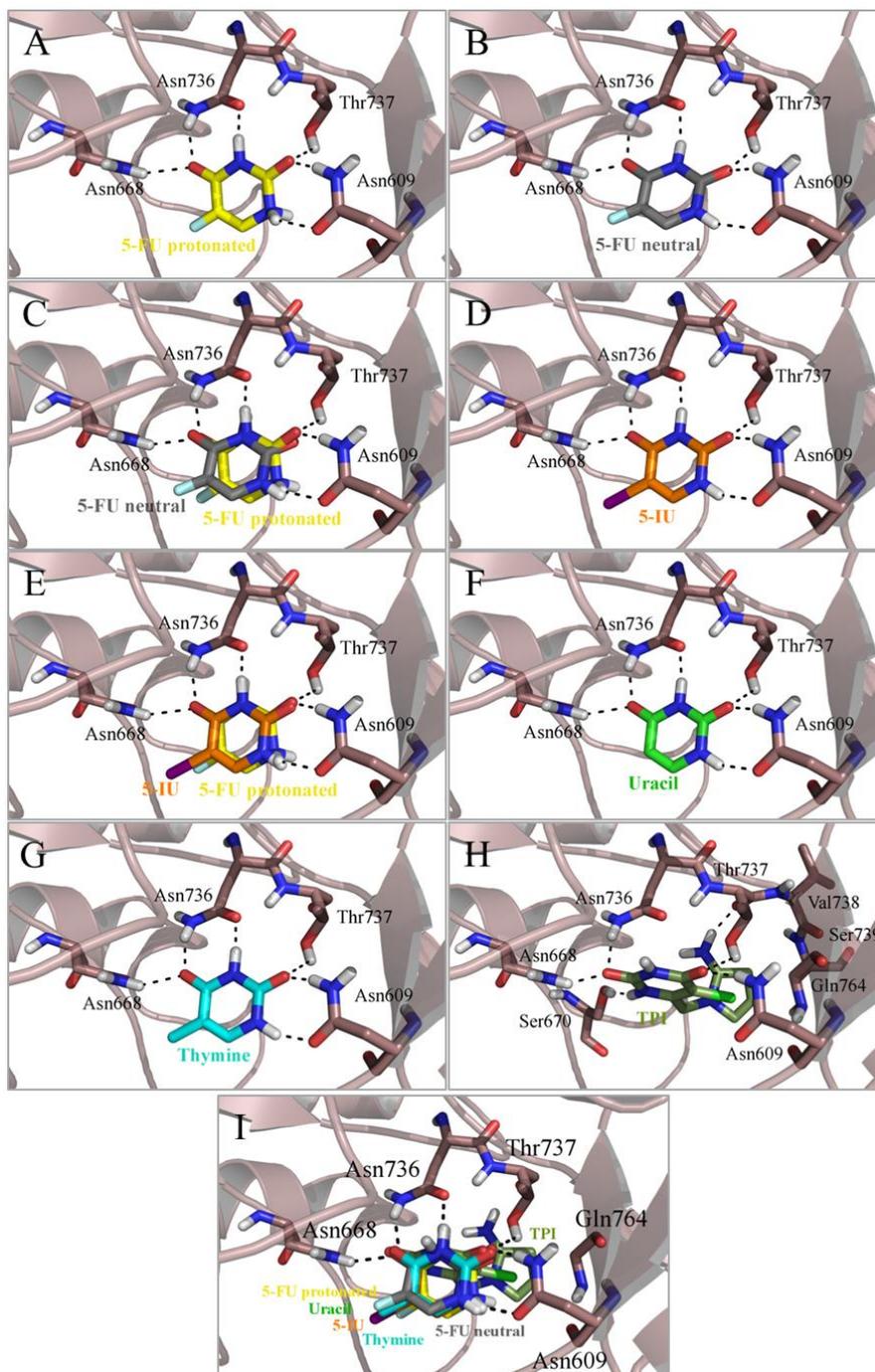


Figure 11: Binding interactions of DPD (salmon) complexed with **A**) 5-FU protonated (yellow), **B**) 5-FU neutral (grey), **C**) overlay of 5-FU protonated and 5-FU neutral, **D**) 5-IU (orange), **E**) overlay of 5-IU and 5-FU-protonated, **F**) Uracil (lime green), **G**) Thymine (cyan), **H**) TPI (smudge), and **I**) overlay of all ligands. Ligands are shown in sticks and hydrogen bonds in dash lines. Atoms: nitrogen (blue), hydrogen (white) and oxygen (red).

4. Conclusions

The MD simulations as well as the NMA carried out in this study, aimed to screen a broad conformational space of the hTP, and to elucidate effects of the interdomain hinges and their twists on the formation of the open-closed (inactive-active) conformations of the enzyme. The study was performed in the presence or absence of the 5-FU, in either of its protonated or unprotonated forms.

The terminal residues are rarely solved in protein crystallography mainly due to the high flexibility of the terminal segments and their extensive exposure to water molecules of the surrounding aqueous environment. This causes poor diffraction data and thus uncertainties when the X-ray diffraction data are associated with the structure of residues in the terminal segments. Through a lengthy MD simulation of up to 500 ns, we monitored the changes of the secondary structure of the N-terminal residues, which has not been solved in the existing TP crystal structures. Two types of folding were observed; a small β -sheet composed of two anti-parallel β -strands seen in the unliganded monomer and the monomer complexed with protonated 5-FU, while a turn with a π -helix shape was seen in the dimer. It is worth noting that the observed secondary structure from MD simulations for the N-terminus may depend on the force field (Cino, Choy, & Karttunen, 2012), hence to further confirm our observation, MD simulations of the protein using other force fields than GROMOS96 are suggested.

In the five systems studied here, a segment of hinge 1 folded into a helix at various timeframes throughout the MD trajectories. The hinge 1 showed that is the major factor involving movements of the domains. The majority of the open-closed movements in the corresponding inactive-active modes of the enzyme appeared to be mediated by only one hinge and the roles of the other hinges in the mechanical flexibility of the domains were found insignificant.

Binding of the catalytic phosphate was thought to be the cause of an 8° rotation of the α/β -domain relative to the α -domain of the hTP prior to pyrimidine binding to the active site. In MD simulations, water molecules occupied the phosphate-binding site establishing hydrogen bonds that compensated the binding strength for the absence of the phosphate ion. The angle between the α/β -

domain and the α -domain also underwent wider rotations over the course of MD simulation of the protein than the reported range from the X-ray structure. This was for $\sim 46.6^\circ$ - 75.3° in the unliganded hTP suggesting that the twist occurs between the domains, in the absence or presence of the catalytic phosphate or a ligand, in open-closed conformations. The pKa value of 5-FU suggests that it binds to the protein in its protonated form. Thus, neutral and protonated 5-FU as well as uracil, thymine, TPI were studied by means of docking.

According to the X-ray structure, it was thought that the substrate is not required for the hTP closed or active conformation and observation of a closed conformation for both liganded and unliganded forms was associated with the enzyme crystal packing. However, analysis of the dCOM variations, measuring the angle between the domains as well as the RMSF data indicated that when hTP was liganded with neutral 5-FU, the ligand was loosely held in place and the protein continued to undergo an open-closed conformation while the ligand was bound. In comparison, when hTP was bound to protonated 5-FU, the positive charge held the protein more tightly with a distance of 14.9 \AA vs. 31.2 \AA at the binding site entrance and restricted the open-close movement. The protonated ligand stabilizes the core and binding site region of the enzyme as it was shown by the RMSF data.

All the uracil analogues studied here bind to almost identical location to that of 5-IU in the X-ray structure. By performing docking on various analogues of the 5-FU as well as the MD simulations, both, repetition runs as well as the extra-long MD simulations of up to 500 ns, no other binding site with potential for accommodating 5-FU, better than that of the known active site, was observed. This supports and, is in agreement with, the available experimental data (Mitsiki et al., 2009).

The TPI possesses the lowest binding energy mainly due to its additional side chain compared to the other analogues of uracil. The same ligands were docked into the DPD binding site.

This was studied to investigate how $\sim 80\%$ of the administered 5-FU is directed to the undesired target, DPD. It is worth mentioning that the response and survival rates of patients treated with 5-fluorouracil (5-FU) have been compared to those treated by the co-administration of 5-FU with other chemotherapy agents (Longley et al., 2003). 5-FU, a fluoropyrimidine, used in the treatment of colorectal cancers and many other cancers, when administered with leucovorin (5-FU/LV), irinotecan, oxaliplatin as well as following methotrexate, patient response rates increased. While fluorouracil has shown to be of particular clinical use as a first-line treatment, it clearly results in more side effects as a large amount of the administered drug binds to an undesired target site.

The ligand binding study has shown that the iminopyrrolidinyl of the TPI dislocates and rotates its uracil fragment $\sim 75^\circ$ compared to the uracil conformation in other derivatives in the DPD

binding site. The dislocation was not seen in the conformation of TPI in the hTP site. Thus, the addition of a hydrophobic fragment, instead of the polar group iminopyrrolidinyl attached to uracil, is expected to reduce the binding energy as well as the interaction forces between the ligand and the DPD residues (considering the type of residues in the DPD site). The electrostatic surface area of the DPD binding site that is mainly positive can be exploited for directing positively charged ligand towards hTP, because the electrostatic surface of hTP shows to be mainly negative in its binding site area. Stabilizing a positive charge on a uracil analogue with a quaternary nitrogen is suggested to direct and enhance binding of the ligand to hTP, thus it could increase drug selectivity towards the desired target. The MD results also confirmed increased stability of the hTP binding site through interaction with the positively charged 5-FU relative to the neutral ligand bound hTP and its unliganded form, according to the RMSF data.

The presented results obtained through MD simulations and other accompanying *in silico* methods are complementary to the currently existing information on hTP or DPD from crystallography. This work aimed to fill the knowledge gap caused by the common technical limitations of crystallography such as the inability to detect the protonation state of drugs and protein residues or screening a wide conformational space of the proteins or their bound ligands since the diffraction data generates one static structure of the system.

Acknowledgements

The authors thank Compute Canada (in particular ACENET and WestGrid) for accessing their HPC facilities free of charge and for the technical supports.

L.A. is grateful to Memorial University of Newfoundland and APOTEX CANADA for the funding support.

Data Sharing and Data Accessibility

The data that support the findings of this study are available on request from the corresponding author.

Conflict of Interest

The authors have no conflict of interest.

5. References

- Bahar, I., Lezon, T. R., Bakan, A., & Shrivastava, I. H. (2010). Normal Mode Analysis of Biomolecular Structures: Functional Mechanisms of Membrane Proteins. *Chemical Reviews*, *110*(3), 1463-1497. doi: 10.1021/cr900095e
- Bohm, H. J. (1998). Prediction of binding constants of protein ligands: a fast method for the prioritization of hits obtained from de novo design or 3D database search programs. *J Comput Aided Mol Des*, *12*(4), 309-323.
- Chagas, C. M., Moss, S., & Alisaraie, L. (2018). Drug metabolites and their effects on the development of adverse reactions: Revisiting Lipinski's Rule of Five. *International Journal of Pharmaceutics*, *549*(1), 133-149. doi: <https://doi.org/10.1016/j.ijpharm.2018.07.046>
- Cino, E. A., Choy, W.-Y., & Karttunen, M. (2012). Comparison of Secondary Structure Formation Using 10 Different Force Fields in Microsecond Molecular Dynamics Simulations. *J Chem Theory Comput*, *8*(8), 2725-2740. doi: 10.1021/ct300323g
- Darden, T., York, D., & Pedersen, L. (1993). Particle mesh Ewald: An $N \cdot \log(N)$ method for Ewald sums in large systems. *The Journal of Chemical Physics*, *98*(12), 10089-10092. doi: <http://dx.doi.org/10.1063/1.464397>
- Dobritzsch, D., Schneider, G., Schnackerz, K. D., & Lindqvist, Y. (2001). Crystal structure of dihydropyrimidine dehydrogenase, a major determinant of the pharmacokinetics of the anti-cancer drug 5-fluorouracil. *EMBO J*, *20*(4), 650-660. doi: 10.1093/emboj/20.4.650
- Dobritzsch, D., Schneider, G., Schnackerz, K. D., & Lindqvist, Y. (2001). Crystal structure of dihydropyrimidine dehydrogenase, a major determinant of the pharmacokinetics of the anti-cancer drug 5-fluorouracil. *The EMBO Journal*, *20*(4), 650-660. doi: 10.1093/emboj/20.4.650
- El Omari, K., Bronckaers, A., Liekens, S., Perez-Perez, M. J., Balzarini, J., & Stammers, D. K. (2006). Structural basis for non-competitive product inhibition in human thymidine phosphorylase: implications for drug design. *Biochem J*, *399*(2), 199-204. doi: 10.1042/bj20060513
- Flynn, M., Heale, K. A., & Alisaraie, L. (2017). Mechanism of Off-Target Interactions and Toxicity of Tamoxifen and Its Metabolites. *Chemical Research in Toxicology*, *30*(7), 1492-1507. doi: 10.1021/acs.chemrestox.7b00112
- Gohlke, H., Hendlich, M., & Klebe, G. (2000). Knowledge-based scoring function to predict protein-ligand interactions. *J Mol Biol*, *295*(2), 337-356. doi: 10.1006/jmbi.1999.3371

- Guvench, O., & D MacKerell, A. (2008). *Comparison of Protein Force Fields for Molecular Dynamics Simulations* (Vol. 443).
- Heldin, C.-H., Usuki, K., & Miyazono, K. (1991). Platelet-derived endothelial cell growth factor. *Journal of Cellular Biochemistry*, 47(3), 208-210. doi: doi:10.1002/jcb.240470304
- Hess, B., Bekker, H., Berendsen, H. J. C., & Fraaije, J. G. E. M. (1997). LINCS: A linear constraint solver for molecular simulations. *J Comput Chem*, 18(12), 1463-1472. doi: 10.1002/(SICI)1096-987X(199709)18:12<1463::AID-JCC4>3.0.CO;2-H
- Hess, B., Kutzner, C., van der Spoel, D., & Lindahl, E. (2008). GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J Chem Theory Comput*, 4(3), 435-447. doi: 10.1021/ct700301q
- Kelley, L. A., & Sternberg, M. J. E. (2009). Protein structure prediction on the Web: a case study using the Phyre server. *Nature Protocols*, 4, 363. doi: 10.1038/nprot.2009.2
- Kremer, A. B., Mikita, T., & Beardsley, G. P. (1987). Chemical consequences of the incorporation of 5-fluorouracil into DNA as studied by NMR. *Biochemistry*, 26(2), 391-397. doi: 10.1021/bi00376a009
- Krenitsky, T. A. (1968). Pentosyl Transfer Mechanisms of the Mammalian Nucleoside Phosphorylases. *Journal of Biological Chemistry*, 243(11), 2871-2875.
- Li, W., & Yue, H. (2018). Thymidine phosphorylase: A potential new target for treating cardiovascular disease. *Trends in Cardiovascular Medicine*, 28(3), 157-171. doi: <https://doi.org/10.1016/j.tcm.2017.10.003>
- Li, Z., & Alisaraie, L. (2015). Microtubules dual chemo and thermo-responsive depolymerization. *Proteins: Structure, Function, and Bioinformatics*, 83(5), 970-981. doi: 10.1002/prot.24793
- Longley, D. B., Harkin, D. P., & Johnston, P. G. (2003). 5-Fluorouracil: mechanisms of action and clinical strategies. *Nat Rev Cancer*, 3(5), 330-338.
- Malde, A. K., Zuo, L., Breeze, M., Stroet, M., Poger, D., Nair, P. C., . . . Mark, A. E. (2011). An Automated Force Field Topology Builder (ATB) and Repository: Version 1.0. *J Chem Theory Comput*, 7(12), 4026-4037. doi: 10.1021/ct200196m
- Mitsiki, E., Papageorgiou, A. C., Iyer, S., Thiyagarajan, N., Prior, S. H., Sleep, D., . . . Acharya, K. R. (2009). Structures of native human thymidine phosphorylase and in complex with 5-iodouracil. *Biochemical and Biophysical Research Communications*, 386(4), 666-670. doi: <http://dx.doi.org/10.1016/j.bbrc.2009.06.104>

- Norman, R. A., Barry, S. T., Bate, M., Breed, J., Colls, J. G., Ernill, R. J., . . . Pauptit, R. A. (2004). Crystal structure of human thymidine phosphorylase in complex with a small molecule inhibitor. *Structure*, *12*(1), 75-84.
- Oostenbrink, C., Villa, A., Mark, A. E., & van Gunsteren, W. F. (2004). A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. *J Comput Chem*, *25*(13), 1656-1676. doi: 10.1002/jcc.20090
- Parker, J. B., & Stivers, J. T. (2011). Dynamics of Uracil and 5-Fluorouracil in DNA. *Biochemistry*, *50*(5), 612-617. doi: 10.1021/bi101536k
- Power, B. H., Smith, N., Downer, B., & Alisaraie, L. (2017). Insight into the mechanism of chemical modification of antibacterial agents by antibiotic resistance enzyme O-phosphotransferase-III A. *Chemical Biology & Drug Design*, *89*(1), 84-97. doi: doi:10.1111/cbdd.12835
- Pugmire, M. J., Cook, W. J., Jasanoff, A., Walter, M. R., & Ealick, S. E. (1998). Structural and theoretical studies suggest domain movement produces an active conformation of thymidine phosphorylase. *J Mol Biol*, *281*(2), 285-299. doi: 10.1006/jmbi.1998.1941
- Pugmire, M. J., & Ealick, S. E. (1998). The crystal structure of pyrimidine nucleoside phosphorylase in a closed conformation. *Structure*, *6*(11), 1467-1479.
- Rarey, M., Kramer, B., & Lengauer, T. (1999). Docking of hydrophobic ligands with interaction-based matching algorithms. *Bioinformatics*, *15*(3), 243-250.
- Rarey, M., Kramer, B., Lengauer, T., & Klebe, G. (1996). A fast flexible docking method using an incremental construction algorithm. *J Mol Biol*, *261*(3), 470-489. doi: 10.1006/jmbi.1996.0477
- Schwartz, M. (1971). Thymidine phosphorylase from Escherichia coli. Properties and kinetics. *European Journal Of Biochemistry / FEBS*, *21*(2), 191-198.
- Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E., & Berendsen, H. J. (2005). GROMACS: fast, flexible, and free. *J Comput Chem*, *26*(16), 1701-1718. doi: 10.1002/jcc.20291
- Walter, M. R., Cook, W. J., Cole, L. B., Short, S. A., Koszalka, G. W., Krenitsky, T. A., & Ealick, S. E. (1990). Three-dimensional structure of thymidine phosphorylase from Escherichia coli at 2.8 Å resolution. *J Biol Chem*, *265*(23), 14016-14022.