BIOMIRROR

An Open Access Journal

The Design and Optimisation of Novel Human Dihydrofolate Reductase Inhibitors for the Management of Proliferative Disease

[a] Graziella Portelli*,
[a] Claire M Shoemake

[a] Department of Pharmacy, Faculty of Medicine and Surgery, University of Malta, MALTA.

ARTICLE INFO

ABSTRACT

Tetrahydrofolate (THF) mediates DNA and RNA synthesis through production of purine and thymidylate precursors. During this process THF is reduced to the inactive dihydrofolate (DHF) and recycled back to the active DHF via a redox reaction, catalysed by dihydrofolate reductase (DHFR). DHFR inhibition prevents cellular growth, hence drug design at this locus is considered valuable with DHFR antagonists having clinical relevance in proliferative disease management. This study utilised methotrexate (MTX) as lead molecule in the design and optimisation of novel DHFR antagonists. PDB crystallographic deposition 1U72 (Cody et al., 2005) [3] describing the holo MTX: human DHFR complex was modelled in SYBYL-X® v1.2 (Tripos) and affinity of MTX for the cognate receptor measured in X-SCORE v1.2 (Wang et al.,1998)to establish baseline affinity. Structure activity data and 2D-topology maps generated in PoseView v1.1 (Stierand and Gastreich, 2011)[6] guided the creation of 7 seeds in which moieties considered non-critical for binding and clinical effect were computationally modified using the GROW module of LigBuilder v1.2 (Wang et al., 2000)[7].Each of the 7 seeds yielded 200 novel structures which were classified according to pharmacophore structure, physiochemical parameter and binding affinity. This molecular cohort was assessed for Lipinski Rule compliance which reduced the total number of viable molecules to 177. These were rendered in UCSF Chimera v1.8 (Pettersen et al., 2004)[5] and Accelerys Draw® v4.1(Accelrys Software Inc., 2013)[1] for visualisation and pharmacophoric growth deduction. The optimal structures combining affinity and Lipinski Rule compliance from each pharmacophoric group were identified, which could be further optimised for in vitro validation on the premise that they hold promise as clinically use anti-proliferative drugs.

Introduction:

Tetrahydrofolate a derivative of Vitamin B9, acts a one-carbon donor during reactions necessary for the synthesis of essential nucleotide precursors, thymidylate and methionine amongst other metabolites. At the end of these reactions, tetrahydrofolate becomes reduced to the insufficient folate form, dihydrofoalte.

Many of the enzymes which make use of tetrahydroflate recycle the dihydrofolate back to tetrahydrofolate. However the enzyme Thymidylate Synthase is the only enzyme utilising one-carbon-transfers that ultimately fails to reproduce tetrahydrofolate. It is the Dihydrofolate Reductase enzyme (DHFR) that recycles the dihydrofolate to tetrahydrofolate via a redox reaction in the presence of NADPH co-factor.

Since the DHFR enzyme is solely responsible for maintaining the *in-vitro* necessary pools of tetrahydrofolate, this enzyme has a crucial role in supporting the production of DNA and RNA in all living organisms. Moreover, inhibition of this ubiquitous enzyme leads to cessation of cellular growth, hence making the DHFR enzyme an ideal target in humans for inhibiting proliferation in rapidly dividing cells.

The drug methotrexate targets the human DHFR, resulting in cell growth inhibition when it binds to this enzyme instead of Dihydrofolate. Apart from its indications in inflammatory diseases such as rheumatoid and psoriatic arthritis, MTX is to date the main antimetabolite used in chemotherapy since its FDA approval in 1953.

It is against the background of this information that this study was carried out. The aim was to discover and optimise high *in-silico* binding affinity antifolate ligands for the hDHFR enzyme, with predicted oral bioavailability and the potential to be developed into clinically useful agents.

Methodology:

Molecular modelling was carried out in SYBYL-X® v1.2 (Tripos). The *holo* hDHFR: MTX complex illustrated in Figure 1, was edited such that any water molecules considered as non-critical to binding were eliminated. The co-crystallized NADPH heteroatom's and water molecules in the vicinity of the Ligand Binding Pocket (LBP) were retained. MTX as illustrated in Figure 2, was extracted from its LBP and exported, together with the *apo* DHFR receptor as illustrated in Figure 3, into X-SCORE v1.2 (Wang et al.,1998) [7] for the baseline ligand binding affinity (*pKd*) for all novel structures designed in this study.

Figure 1:hDHFR depicted in *rainbow* surface with MTX and NADPH. Image generated with UCSF Chimera v1.8(Pettersen *et al*., 2004)
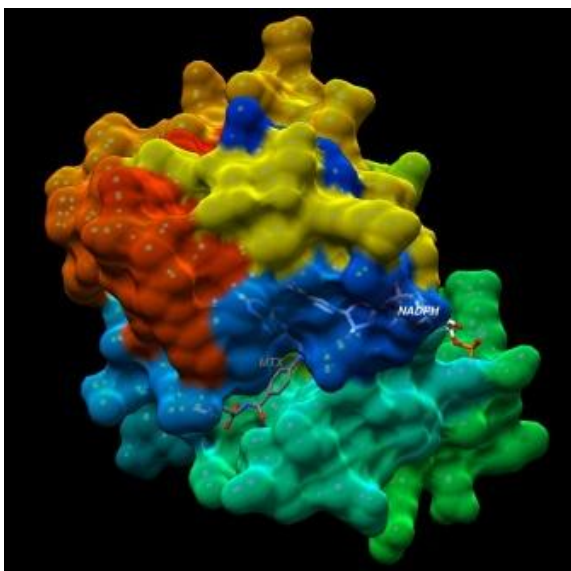
Structure activity data from the literature (Cody *et al*., 2005)[3] (Oefner *et al*., 1988)[4] guided the creation of seed structures which are molelular fragments capable of sustaining structure directed growth. 7 seed structures were designed differing both in magintude and growing site loci, designated as*Hspc* atoms as depicted in Figure 4 and Table 1.The designed structures ensured that the *de novo* deisgn exercise would explore maximal pharmacophoric space and novel structural breadth.

De novo molecular growth was carried out in LigBuilder v1.2 (Wang *et al*., 2000)[7]. The POCKET algorithm of this programme was used to delineate the 3D

LBP of the hDHFR enzyme depicted in Figures 5-7 based on the bioactive conformation of MTX to propose a general pharmacophoric structure on which all the novel structures generated would be based.

Figure 2: MTX as 3D structure depicted in grey with heteroatoms. Image generated with UCSF Chimera v1.8 (Pettersen *et al.*, 2004)
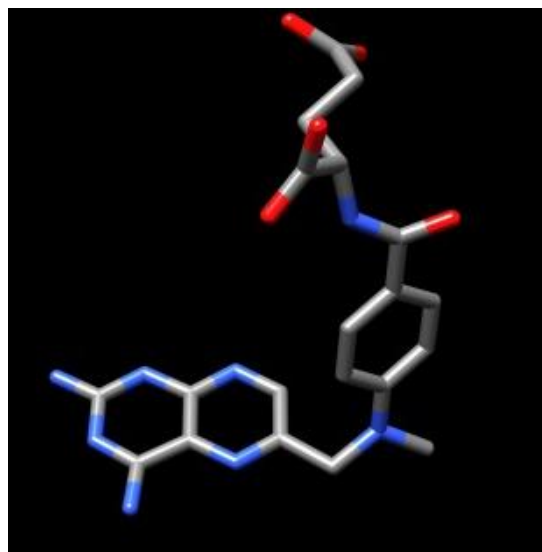

Figure 3: *Apo* hDHFR depictedaccording to strands (purple), helix (red), coils (yellow). Image generated with UCSF Chimera v1.8 (Pettersen *et al*)

Using the GROW algorithm of LigBuilder v1.2 (Wang *et al*., 2000)[7] each seed structure was introduced into the LBP map, generated in the POCKET algorithm of LigBuilder v1.2 (Wang *et al*., 2000)[7]and allowed growth within its confines, initiating from the pre-designed growing site. 200 molecules were generated from each seed and were organized by the PROCESS algorithm of LigBuilder v1.2 (Wang *et al*., 2000)[7] into a molecular database with in which molecules were segregated into families according to LBA.

Physicochemical parameters; including *log*P, molecular weight and synthetic feasibility score were also included. The generated small molecules for each seed were filtered for Lipinski Rule compliance (Lipinski *et al.,* 2001). This reduced molecular cohort was analysed from a pharmacophoric perspective with the pharmacophores specific to each family for each seed being identified.

Figure 4: The 7 Seeds A-G; aligned in accordance to their magnitude with respect to MTX. Image generated with UCSF Chimera v1.8 (Pettersen *et al.,* 2004)
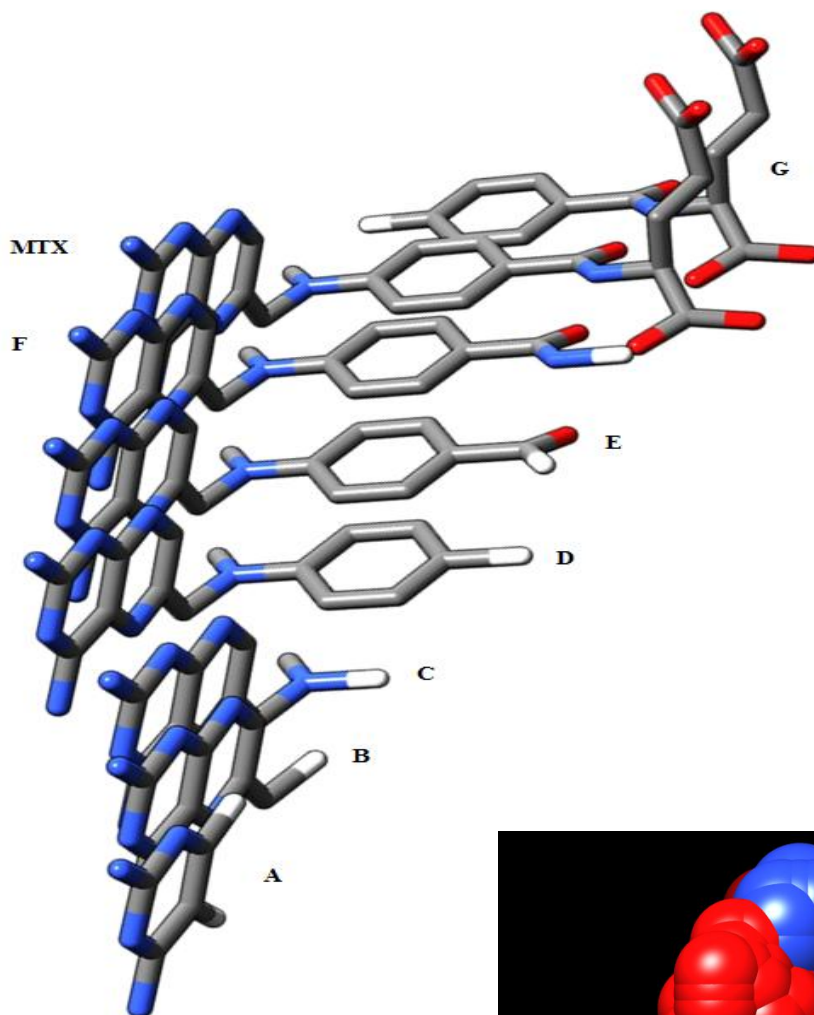

Figure 5: The delineated LBP produced by POCKET LigBuilder v1.2 (Wang et al., 2000)[7]. Atomic representations are as follows: Nitrogen atoms (Blue) represent hydrogen-bond donor sites; oxygen atoms (Red) represent hydrogen-bond acceptor sites and carbon atoms (Taupe) represent hydrophobic sites.Image generated with UCSF Chimera v1.8 (Pettersen
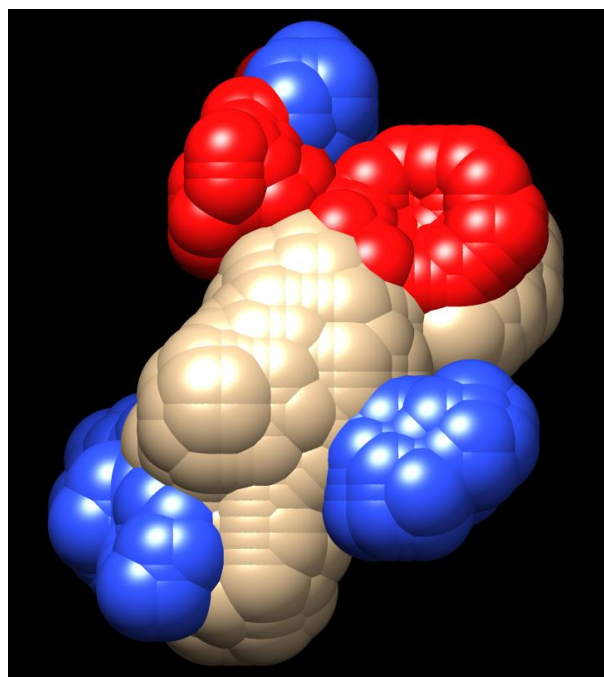
Figure 6: The Pharmacophore model, produced by POCKET LigBuilder v1.2 (Wang et al., 2000)[7]. Atomic representations are as indicated in Figure 5 above
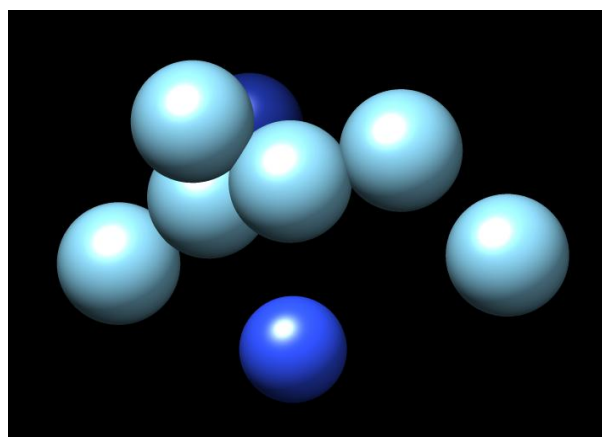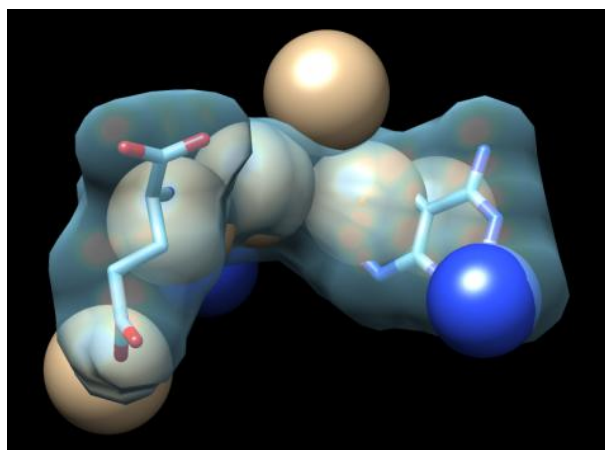

Figure 7: MTX inside the delineated LBPproduced by POCKET LigBuilder v1.2 (Wang et al., 2000). Image generated with UCSF Chimera v1.8 (Pettersen *et al.,* 2004)

The optimal Lipinski Rule compliant (Lipinski *et al.,* 2001) molecule from each family was identified. PoseView v1.1 (Stierand and Gastreich, 2011) was used in order to generate 2D topology maps describing the interactions between the specific molecular moieties of the selected optimal structures and amino acids forming the perimeter of the hDHFR LBP. The LBA and binding energy for the top selected ligands were also calculated using X-SCORE v1.2 (Wang *et al.,*1998), the values of which are illustrated in Table 2.

Results and Discussion

From total of 1400 novel structures 177 were compliant with Lipinski's Rule of 5 (Lipinski *et al.,* 2001), the majority of which derived from seed A followed by seed G. These results are illustrated in Table 1.

Seed A, the smallest seed, comprising exclusively the pyrimidine ring of MTX as illustrated in Figure 4, gave the maximum number of Lipinski rule compliant molecules with a total of 89 molecules from 10 different families. All 10 families had at least one structure that was compliant with Lipinski's rule of 5. The top selected molecule derived from Family 2 which generated the maximum number of Lipinski Rule of 5 (Lipinski *et al.,* 2001).The top selected molecule Result_052 illustrated in Figure 8 had molecular weight of 383 *log*P value of 3.2, *pKd* value of 9.85 and a total of 5 and 7 H-Bond donors and acceptors respectively.

The selected ligand for seed A formed hydrogen bonds represented by dashed lines emanating from single pyrimidine ring, with Glu^{30}, Ile^{115} and Ile^{7}. The pyrimidine ring of the top selected molecule also formed π interactions with the hydrophobic amino acid Phe^{34}. These hydrogen bonds and π interactions are the same as those forged by MTX via its pteridine ring. The ligands' end terminal comes in contact with a hydrophobic pocket formed byIle^{60} and Phe^{3}as depicted by the green spline segments. In MTX these amino acids are in contact with the bridge region. Another hydrogen bond was recognized to be formed with Try^{121} by the NH_4 group at the end terminal which is not recognized in the generated complex of MTX.

The pteridine ring of the top ligand for seed B, interacted with the same amino acids as this moiety did in the top selected ligand for seed A, with amino acids Phe^{34}, Glu^{30}, Val^{115} and Ile^{7} respectively. The bridge region did not interact significantly to binding. Growth at end terminal generated a pyridine ring fused with a benzene ring with various electronegative carbonyl groups attached to the pyridine ring. Asp^{21} interacted via a hydrogen bond with the hydroxyl group substituent to pyridine ring, as donated by the dashed line. The benzene ring interacts via π bonds with Phe^{31}, while the end terminal occupies a hydrophobic pocket formed byPro^{61}, Phe^{34}, Ile^{60} and Phe^{31} as donated by the green spline segment. This gave the highest predicted binding affinity and energy as illustrated in Table 2.

SeedC generated only 4 Lipinski's rule compliant molecules, deriving from two different families. Result_163 deriving from Family 5 was selected as the top ligand for seed C, with molecular weight of 476, *log* P of 4.05, *pKd* of 9.73 and a total of 4 and 9 hydrogen bond donors and acceptors in total.

The pteridine ring of Result_163, illustrated in Figure 10 interacted with Glu^{30}, Ile^{115}, Ile^{7} and Phe^{34} as described for the top ligands selected for seeds A and B. Binding differences arose at the binding interactions for the bridge and end terminal region, which occupied the hydrophobic pocket formed Pro^{61}, Phe^{34}, Ile^{60} and Phe^{31} as donated by the green spline segment.

Seed D gave rise to only 3 Lipinski rule complaint molecules. Result_179 illustrated in Figure 10 was selected originating from Family 3 was selected as the top seed, with a molecular mass of 494, *log*P of 3.67, *pKd* of 9.93 and a total of 4 H-bond donors and 9 H-Bond acceptors. This gave the highest predicted binding affinity and energy with values of 7.10 and -9.69 respectively.

Table 1: Illustrating the number of accepted and number rejected molecules and families for each 4of the 7 seeds, depicted in the second column. Figures of seeds were generated with UCSF Chimera v1.8 (Pettersen et al., 2004).
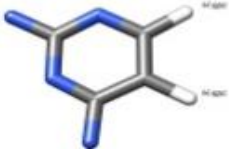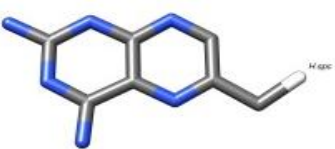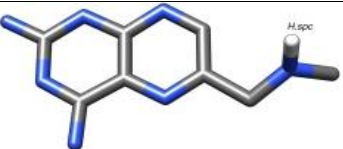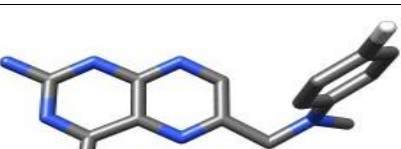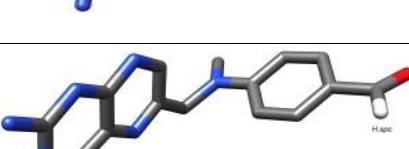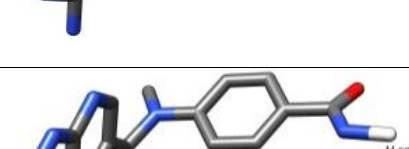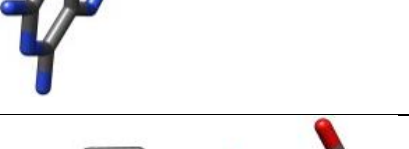
	Seed Structure	Accepted		Rejected	
		Number of molecules	Different families	Number of molecules	Different families
A		89	10	110	0
B		16	4	184	3
C		4	3	195	7
D		3	2	197	6
E		28	7	172	8
F		3	5	192	10
G		41	7	160	10

Figure 8: Complex of top ligand for seed A, Reulst_052 in apo-hDHFR, generated by PoseView (Stierand and Gastreich, 2011)
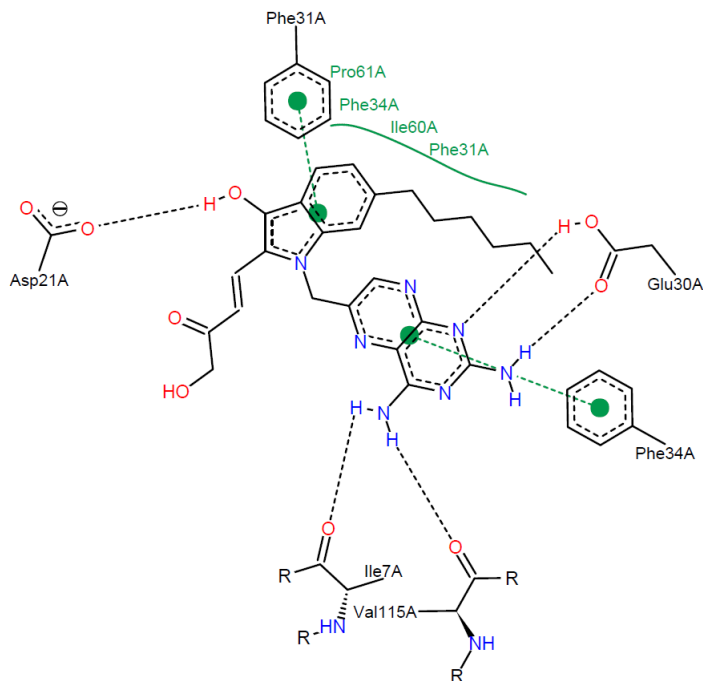

Figure 9: Complex of top ligand for seed B, Result_125 in apo-hDHFR, generated by PoseView (Stierand and Gastreich, 2011)
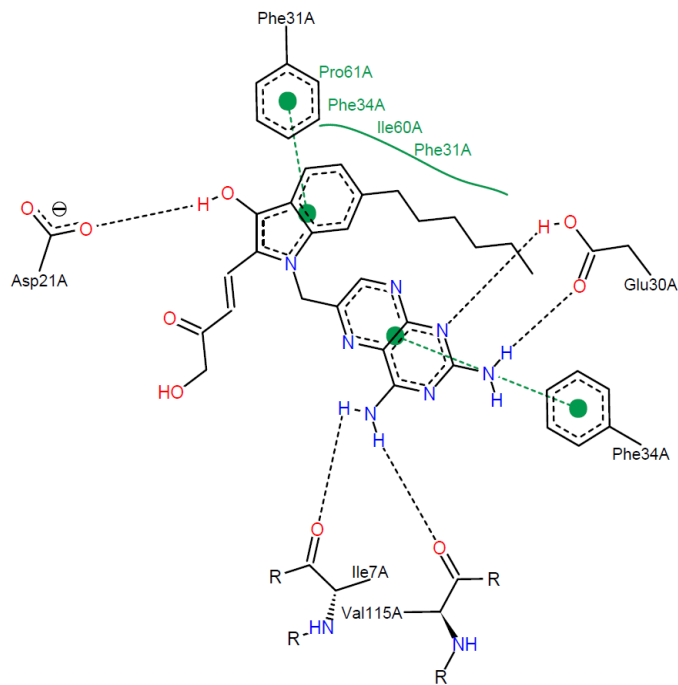

Figure 10: Complex of top ligand for seed C, Result_163 in apo-hDHFR, generated by PoseView (Stierand and Gastreich, 2011).
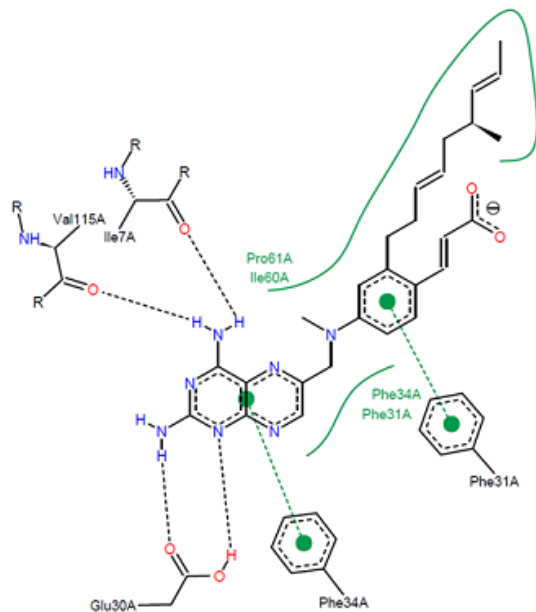
The top ligand selected for seed D generated similar binding modalities in comparison with the top selected ligands for seeds A, B and C. The only difference was the absence a bond with Ile[60] that was part of the hydrophobic pocket as illustrated in Figure 11.

Figure 11: Complex of top ligand for seed D, Result_179 in apo-hDHFR, generated by PoseView (Stierand and Gastreich, 2011)
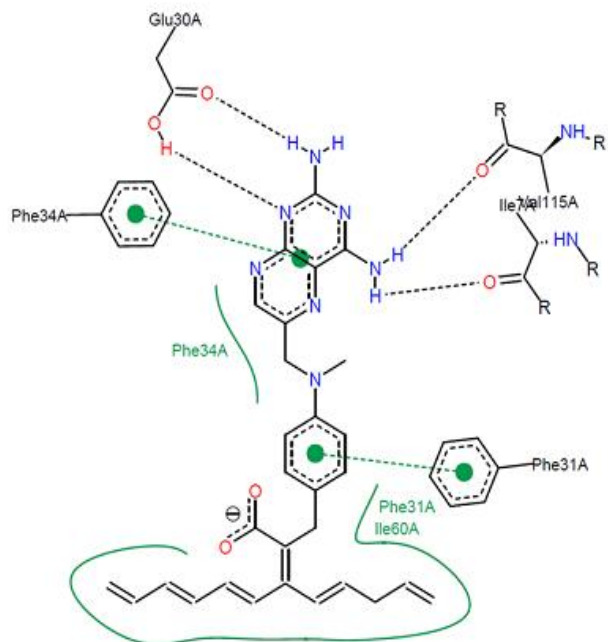
Seed E generated a total of 28 Lipinski rule compliant molecules (Lipinski *et al*., 2001).

Result_132 from Family 6 was selected as the top ligand with molecular weight of 495, *log*P of 4.19, *pKd* of 9.97, 3 and 9 H-bond donors. This ligand interacted with the receptor throughGlu[30], Ile[115], Ile[7] and Phe[34]amino acids that formed hydrogen bonds with the pteridine ring. Phe[34], Ile[60] and Phe[31] formed a hydrophobic pocket that was occupied with the bridge region and Arg[70]formed hydrogen bonds with the end terminal,as donated by dashed lines in Figure 12.

Table 2: Illustrates the predicted *pKd* and binding energy for MTX and the top selected ligand for seeds A to G respectively

Ligand	Predicted binding affinity -log (Kd)	Predicted binding energy kcal/mol
MTX	6.73	-9.19
Top selected ligand for Seed A	6.83	-9.32
Top selected ligand for Seed B	7.16	-9.77
Top selected ligand for Seed C	7.10	-9.69
Top selected ligand for Seed D	7.49	-10.22
Top selected ligand for Seed E	7.54	-10.28
Top selected ligand for Seed F	7.02	-9.57
Top selected ligand for Seed G	7.82	-10.67

Figure 12: Complex of top ligand for seed E, Result_132 in apo-hDHFR, generated by PoseView (Stierand and Gastreich, 2011)

Figure 13: Complex of top ligand for seed F, Result_031 in apo-hDHFR, generated by PoseView (Stierand and Gastreich, 2011)
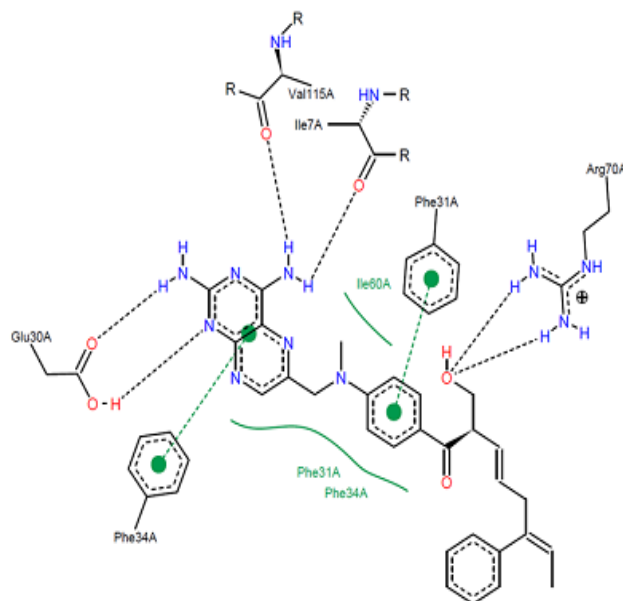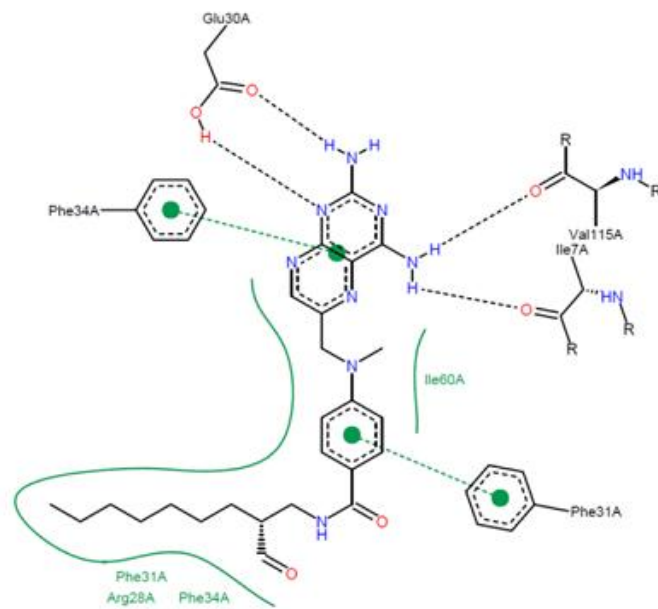

On the other hand, seed F, resulted in only 8 Lipinski rule compliant molecules. Result_031, illustrated in Figure 14 was chosen from Family 2 as the top selected ligands, with molecular mass of 463, *log* P of 4.26, *pKd* of 9.98, 3 H-bond donors and 9 H-Bond acceptors. Result_031, as depicted in Figure 13, made contact with a hitherto unutilized amino acid, specifically Arg[28], which was noted

to form part of the hydrophobic pocket which formed part of a hydrophobic pocket occupied by the end terminal of the ligand. Seed F, the parent fragment of this specific molecule, was the most structurally similar to MTX from the seed population.

Seed G was the only seed that represented the terminal glutamate tail. This generated a substantial amount of Lipinski rule complaint molecules. Result_017 was selected as the top ligand, with molecular weight of 463, *log*P of 3.49, *pKd* of 9.98, H-bond donor and acceptor of 6 and 7 respectively.

Result_017 illustrated in Figure 14, differed from the other top selected ligands because it lacked of pteridine ring. This was replaced by a two aromatic ring system (naphthalene ring), with substitutions by an ammonium group, a carbonyl group and a hydrophilic chain.
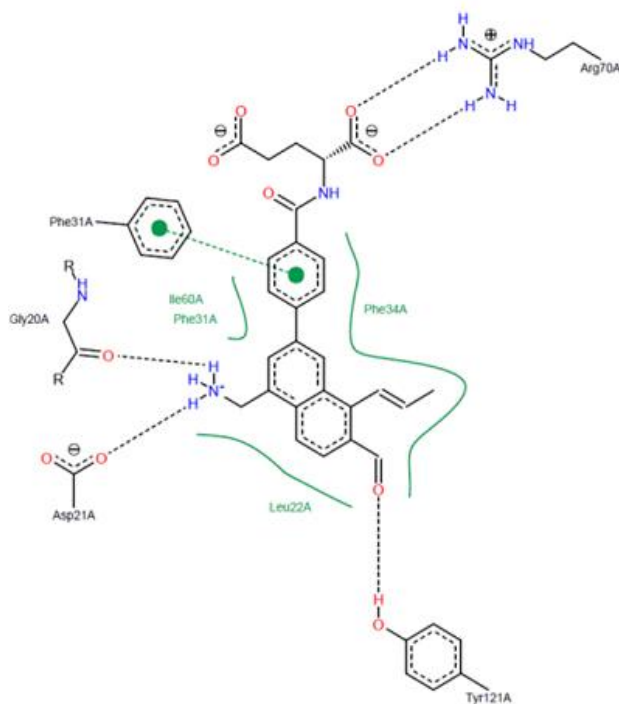
Figure 14: Complex of top ligand for seed G, Result_017 in apo-hDHFR, generated by PoseView (Stierand and Gastreich, 2011)

Four hitherto unaccessed amino acids interacted with this ligand; Gly[20] and Asp[21] interacted via hydrogen bonds with the ammonium group, Tyr[121] interacted via hydrogen bonds with the carbonyl group, while Leu[22] formed part of a hydrophobic pocket which was occupied by the ring system. All other interactions, specifically with Ile[60], Phe[31], Phe[34] and Arg[70] had previously been observed with the top ranking molecules deriving from the other seed structures.

Conclusion

This rational drug design study was constructed in such a way that allowed molecular growth to occur within the confines of the LBP in a staggered fashion, implying different degrees of freedom. This was done in order to ensure the casting of as wide as possible of a pharmacophoric

net that would ensure a cohort of molecules that was as structurally diverse as possible, which is an asset in subsequent rounds of structure optimisation.

In fact this study has achieved pharmacophoric diversity with a total of 82 structurally distinct families (n=10 for seed A, n=7 for seed B, n=10 for seed C, n=8 for seed D, n=15 for seed E, n= 15 for seed F and n= 17 for seed G).
Common to all *in silico* studies, this project will at its terminus provide a hypothesis; in this case, the hypothesis is that a preliminary cohort of high affinity molecules whose structure is based on a bioactive antagonist conformation of MTX and which have the propensity to oral bioavailability, has been designed *de novo.*

An issue that must be raised is the robustness of this hypothesis. This study must be considered as a first step in a rational, logically well executed trajectory, towards the identification of a novel clinically useful entity. The caveat consequently, is 'first step' and it is in this context that the robustness of this hypothesis must be evaluated.
This study was carried out under a number of constraints the most pertinent of which was time. This study was also carried out in a static environment in which both ligand and receptor are considered as rigid entities, which clearly is not representing the *in vivo* scenario. However, irrespective of this, a number of precautions were taken in order to lay the foundations for the exercise of time consuming and computationally intense molecular dynamics simulations, which would impart *in silico* notion to both entities according to Newtonian physics. These precautions included molecular simplifications prior to drug design commencement. Specifically, redundant water molecules and ligands not considered essential to ligand binding were computationally edited from the pdb crystallographic deposition 1U72 (Cody *et al.*, 2005)[3].

In a scenario in which this study was to be furthered, the initial hypothesis made would be validated through an *in silico* molecular dynamic simulating study, before being confirmed through *in vitro* assay techniques

Reference:
1. Accelrys Software Inc., Discovery Studio User Guide, June 2007, Accelrys Software Inc., San Diego.
2. Schweitzer B, Dicker A, Bertino J. Dihydrofolate reductase as a therapeutic target. [Internet] 2012 [cited: 2012 Apr 4]. The FASEB Journal. 2441-2452. Available from: http://www.fasebj.org/content/4/8/2441
3. Cody V, Luft JR, Pangborn W. Understanding the role of Leu22 variants in methotrexate resistance: comparison of wild-type and Leu22Arg variant mouse and human dihydrofolate reductase ternary crystal complexes with methotrexate and NADPH. Acta Crystal*log*r D Biol Crystal*log*r.2004;61(2): p. 147-55.
4. Oefner C, D'Arcy A, Winkler FK. Crystal structure of human Dihydrofolate reductase complexed with folate. Eur J Biochem.1988; 174(2): p. 377-85.
5. Pettersen, E.F. et al. UCSF Chimera. Journal of Computational Chemistry. 2004 Oct;25 (13):1605-12.
6. Stierand K, Gastreich M. PoseView 2D Sketches of Protein-Ligand Complexes version 1.1. BioSolveIT GmbH, 2011.
7. Wang. R., Gao. Y., Lai. L., (2000). Ligbuilder: A Multi-Purpose program for Structure-Based Drug Design. *J.Mol.Model* 6: p. 498-51