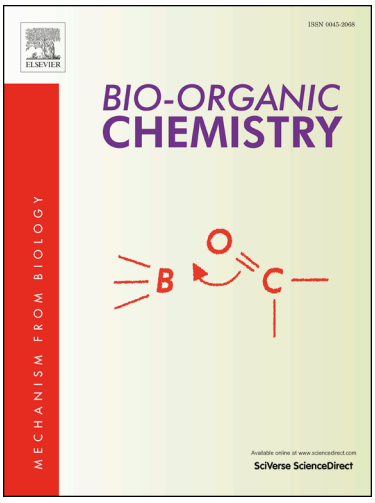


Structural Comparison of Mtb-DHFR and h-DHFR for Design, Synthesis and Evaluation of Selective Non-Pteridine Analogues as Antitubercular Agents

Kalicharan Sharma, Omprakash Tanwar, Shweta Sharma, Shakir Ali, M.M. Alam, M.S. Zaman, Mymoona Akhter

PII: S0045-2068(18)30141-X
DOI: https://doi.org/10.1016/j.bioorg.2018.04.022
Reference: YBIOO 2345

To appear in: *Bioorganic Chemistry*

Received Date: 14 February 2018
Revised Date: 18 April 2018
Accepted Date: 24 April 2018

Please cite this article as: K. Sharma, O. Tanwar, S. Sharma, S. Ali, M.M. Alam, M.S. Zaman, M. Akhter, Structural Comparison of Mtb-DHFR and h-DHFR for Design, Synthesis and Evaluation of Selective Non-Pteridine Analogues as Antitubercular Agents, *Bioorganic Chemistry* (2018), doi: https://doi.org/10.1016/j.bioorg.2018.04.022

Structural Comparison of Mtb-DHFR and h-DHFR for Design, Synthesis and Evaluation of Selective Non-Pteridine Analogues as Antitubercular Agents

Kalicharan Sharma[1], Omprakash Tanwar[1], Shweta Sharma[1], Shakir Ali[2], M. M. Alam[1], M. S. Zaman[1], Mymoona Akhter*[1, 3]

[1]Department of Pharmaceutical Chemistry, SPER, Jamia Hamdard, New Delhi, 110062, India

[2] Department of Biochemistry, School of Chemical and Life Sciences, Jamia Hamdard, New Delhi-110062

[3] Bioinformatics Infrastructure Facility Lab, Jamia Hamdard, New Delhi-110062

✉Dr.Mymoona Akhter

Address: Drug Design and Medicinal Chemistry Lab, Dept. of Pharmaceutical Chemistry, School of Pharmaceutical Education & Research, Jamia Hamdard, New Delhi, India, 110062.

Ph.: +91 9891369482

Mail ID: mymoonaakhter@gmail.com

Abstract: Tuberculosis is an infectious disease that affects millions of population every year. Mtb-DHFR is a validated target that is vital for nucleic acids biosynthesis and therefore DNA formation and cell replication. This paper report identification and synthesis of novel compounds for selective inhibition of Mtb-DHFR and unleash the selective structural features necessary to inhibit the same. Virtual screening of databases was carried out to identify novel compounds on the basis of difference between the

binding pockets of the two proteins. Consensus docking was performed to improve upon the results and best ten hits were selected. Hit 1 was subjected to analogues design and the analogues were docked against Mtb-DHFR. From the docking results 11 compounds were selected for synthesis and biological assay against $H_{37}Rv$. Most potent compound (**IND-07**) was tested for selectivity using enzymatic assay against Mtb-DHFR and h-DHFR. The compounds were found to have good inhibitory activity (25-200µM) against $H_{37}Rv$ and in enzyme assay against Mtb-DHFR and h-DHFR the compound was found selective towards Mtb-DHFR with selectivity index of 6.53. This work helped to identify indole moiety as novel scaffold for development of novel selective Mtb-DHFR inhibitors as antimycobacterial agents.

Keywords: Mtb-DHFR inhibitors, antimycobacterial, Indole moiety, 1DF7, 1OHJ, antitubercular, consensus docking

1. Introduction

DHFR is an important enzyme in both mammals and microorganism including *Mycobacterium tuberculosis* (Mtb). It catalyzes folic acid conversion to dihydro and tetrahydro folic acid which is a crucial step in folate pathway[1]. Inhibition of folate pathway leads to interruption of thymidine supply resulting cell death. Folate metabolism plays an essential role in the biosynthesis of purines and pyrimidines and therefore in the nucleic acid biosynthesis for all the living organisms and is thus directly or indirectly involved in the processes of cell reproduction[2]. Tuberculosis (TB) is a top infectious disease killer worldwide. In 2016 global statistics showed that 10.4 million suffered with TB and 1.7 million died from the disease. TB is a leading killer of HIV-positive people: in 2016, 1 in 3 HIV deaths was due to TB. Globally in 2016, an estimated 490,000 people developed multidrug-resistant TB (MDR-TB) [3]. New forms of MDR-TB and extensively drug resistant TB (XDR-TB), pose new challenge for medicinal chemist. To address new threat, there is an urgent need to develop new anti-TB drugs on novel targets having minimum cross resistance with existing drugs. Thus, new anti-tuberculosis agents which could act via unique mechanism is need of the hour. DHFR being crucial for survival of microorganism has been explored extensively as drug target for example methotrexate[4], trimethoprim[5] and Br-WR99210[6] an analogue of the antimalarial agent R99210[7] (*Figure 1*) have been found effective clinically against tuberculosis. However, all clinically effective DHFR inhibitors are non-selective, and inhibit both human and pathogenic DHFR more or less to a similar extent.

Figure 1: Structure of some reported DHFR inhibitors.

Mycobacterium *(*Mtb*)* and human (h) DHFRs show 26% structure similarity, but their active sites are not identical[8]. Dissimilarities in the enzymatic constitution of the microorganisms and mammals, forms the basis for design of selective inhibitors devoid of toxicity to human cells[9]. Therefore a comparative computational study will help design selective Mtb-DHFR inhibitors. The structural based virtual screening (SBVS) has gained enormous attention and success in pharmaceutical industry over the past few decades for discovering novel drugs[10]. The aim of this study was to develop a novel series of selective Mtb-DHFR inhibitors as antitubercular agents using SBVS and synthetic approach (*Figure 2*).

Figure 2: Identification of potential hit through virtual screening protocol and modification of the same at three different sites to obtain the best compound IND-07

2. Experimental

2.1 In silico Work:

2.1.1 NCI and Drug Bank database: The National Cancer Institute compounds library (NCI release 265,242 compounds, published in May 2012 and drug bank database, 8709 compounds) were used for virtual screening. The 3D conformations of all the molecules were generated using the maestro 9.6 (Glide version) virtual screening tools.

*2.1.2 Software tools:*Protein structure preparation was done by using Prepwizard (Maestro 9.6) and ligands databases were prepared by using Ligprep tool (Maestro 9.6).

2.1.3 Protein Selection and preparation: Two proteinsMtb-DHFR(PDB ID: 1DF7) and h-DHFR (PDBID: 1OHJ) were selected for VS workflow and docking protocols for identification of selective Mtb-DHFR inhibitors. The selection of PDB's was done on the basis of species and resolution. The target proteins were then imported from RCSB Protein Data Bank saved in folder and was prepared. The structure was preprocessed by adding hydrogens and check for missing residues etc. The selection of water molecules was set at default and protein was optimized followed by energy minimization at default setting. The grid was prepared and docking was performed. The docking was carried on 1DF7 without glycerol moiety.

2.1.4 Structural comparison of Mtb-DHFR *to* h-DHFR*and their binding pocket:*

Till date thirteen crystal structures of DHFR have been reported, out thirteen, ten structures are of Mtb-DHFR and three are of h-DHFR with NADP and different inhibitors. The resolution of the structures deposited in PDB range from 1.7 to 2.0 Å, thereby making them excellent for *insilico* studies. The bound inhibitors include methotrexate, trimethoprim and Br-WR99210 [11-14]. Mtb-DHFR and h-DHFR are quite similar in structure; Mtb-DHFR contains 159 amino acid residues compared to 187 of the human protein with active site located at the C-terminal side of the sheet. The structural comparison of two complexes revealed 26% amino acid sequence similarity and also similarity between overall protein folds. However the environment of the inhibitor binding site of the two enzymes has interesting differences like a glycerol molecule can

be seen in pocket of the Mtb-DHFR: MTX complex, while the same is absent in h-DHFR-COP complex as the site is filled with hydrophobic side-chains. These differences between the pathogen enzyme and host enzyme provide opportunities for designing new selective inhibitors of Mtb-DHFR[12].

2.1.5 Study of key difference in binding regions of Mtb-DHFR*and* h-DHFR*:*

In the ternary complex of MTX with Mtb-DHFR, the side of the aminopterin ring is accessible to the solvent which is not in case of *h-DHFR* (*Figure 3*). Also in Mtb-DHFR a glycerol molecule is found in a depression near MTX, Trp22, Asp27 and Gln28 whereas in case ofh-DHFRs (complexed folate or COP) the site is well packed with three hydrophobic residue side chains, Leu22, Pro26 and Phe31 which correspond to Leu20, Arg23 and Gln28 in Mtb-DHFR(*Figure 3a and 3b*). Therefore if a compound is designed which could replace the glycerol from the pocket and interact with the site in addition to other important interactions may result in development of specific Mtb-DHFR inhibitor. On this assumption the present study was undertaken to design novel selective inhibitors for Mtb-DHFR.[12]

Figure 3: Comparison of inhibitor binding site in Mtb-DHFR and h-DHFR *at* catalytic domain. In Mtb DHFR inhibitor binding site a glycerol molecule bound close to the inhibitor MTX is visible. (Li R. *et al*[12])

2.2 Virtual Screening against Mtb-DHFR *and* h-DHFR:

Virtual screening is one of the most widely used approaches for identifying novel hits and for removing non complementary compounds, which can lower the selectivity of compound for targets. Since virtual screening workflow is based on docking of ligands on protein active site, validation of docking protocol become important[15]. The validation of docking procedure was carried out by extracting the co-crystallized ligands MTX and COP from the respective protein complexes obtained from protein data bank and re-docked on the active site of Mtb-DHFR and h-DHFR respectively.

Compounds were collected from NCI and drug bank database and merged to create a single database (2, 73,951 compounds). This database was subjected to ADME and Lipinski rule screening followed by virtual screening workflow against PDB 1DF7. In virtual screening

workflow compounds were subjected to HTVS, SP and XP docking processes. In each one of these steps, by default top 5% of the compounds were selected for the next step. On the basis of score top 100 hits were selected from the XP screening and visually analysis and then subjected to screening against h-DHFR (PDBID: 1OHJ) for selectivity. Fifty compounds which showed lowest docking score against h-DHFR and high score against Mtb-DHFR were selected for further study. Virtual screening work flow is depicted in *Figure 4.*

Figure 4. Virtual Screening Workflow

2.3 Consensus Docking:

compounds selective towards Mtb-DHFR were also subjected to consensus docking using GLIDE module of Schrödinger 9.4 and GOLD suite 5.2.2[16-17].

2.3.1 Docking by using Glide module of Maestro 9.4: Docking of the selected fifty compounds was carried out on Glide module of Maestro 9.4 as per the Glide protocol given in Schrödinger. All the default parameters were used. For ligand preparation the pH was 7.0±2.0, force field was OPLS3 and ionization was done using Epik. For protein preparation the pH was 7.0±2.0, force field was OPLS3, ionization was done using Epikand the water molecules within 5 Å were kept and rest were deleted.

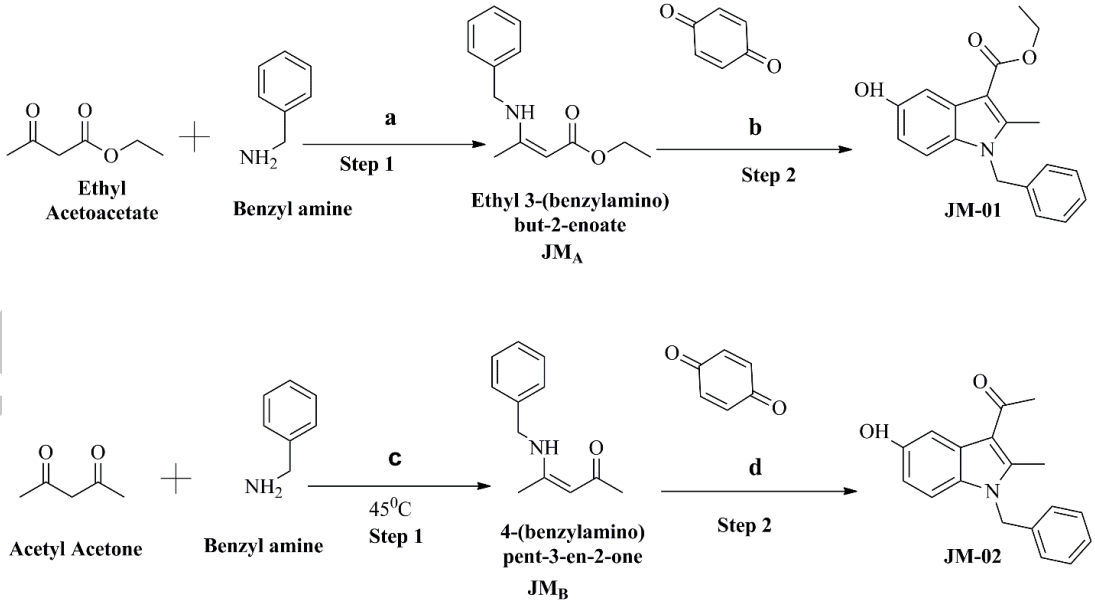
2.3.2 Docking by using GOLD Suite 5.2.2: GOLD suite 5.2.2 was used for consensus docking of final 50 compounds selected using the wizard with default parameters population size (100); selection- pressure (1.1); number of operations (10,000); number of islands (1); niche size (2); and operator weights for migrate (0), mutate (100), and crossover (100) were applied. The active site with a 10 Å radius sphere was defined by selecting an active site residue of protein. Default Genetic Algorithm settings were used for all calculations and a set of 10 solutions were saved for each ligand.

Table 1: Structure of final hits selected after virtual screening of NCI and drug bank followed by consensus docking using Gold suite

On the basis of results of consensus docking ten (10) hits were selected out of which hit 1was selected for synthesis.

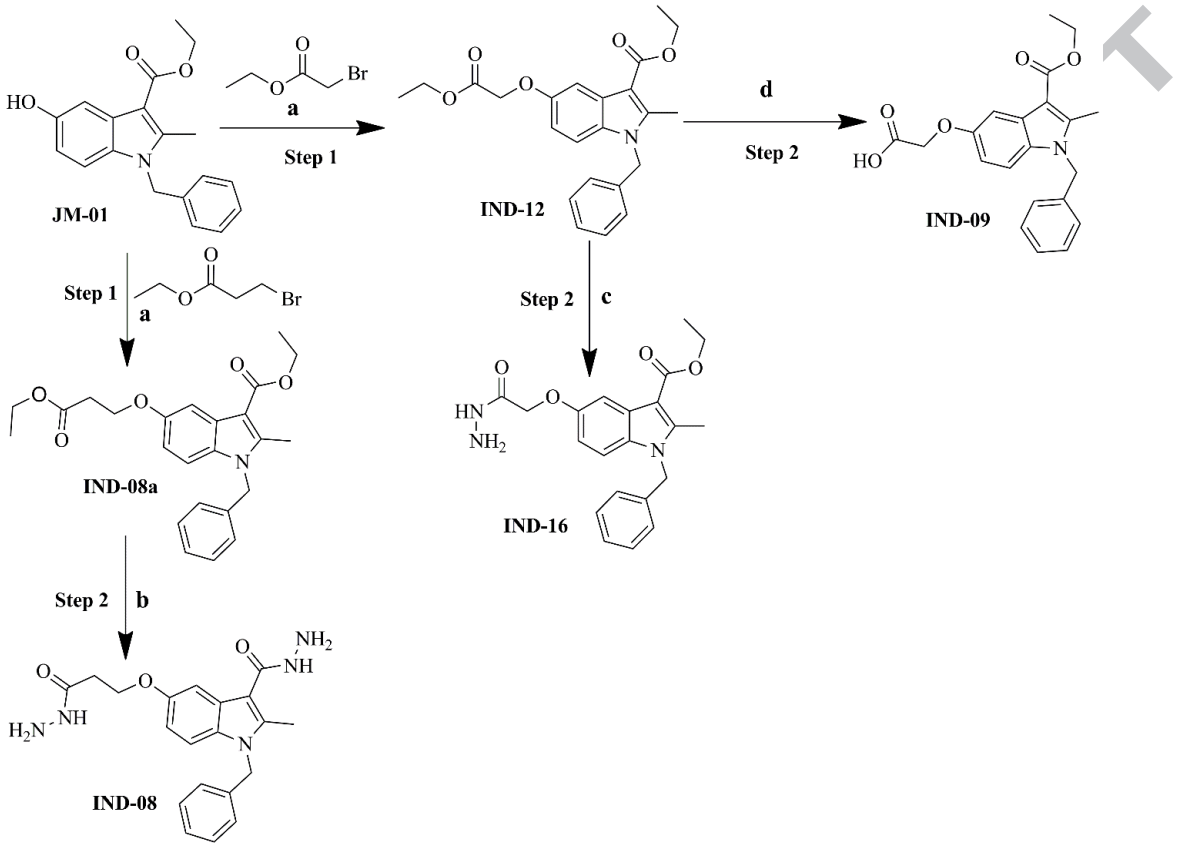
2.4 Chemistry:

All chemicals were purchased from local commercial suppliers and used as such or otherwise mentioned. Melting points were determined by open capillary method by digital melting point apparatus by 'VEEGO' and are uncorrected. ^{1}H NMR and ^{13}C NMR spectra were recorded on a BrukerAvance II 400 spectrometer, using DMSO-d_6 as solvent and TMS as an internal standard. Mass spectral analysis was carried out using Applied BiosystemQtrap 3200 MS/MS system in ESI mode. CHN analysis was done by Elementar, Germany. Infrared (IR) spectra were recorded on a Shimadzu FT-IR 8400S infrared spectrophotometer using the ATR accessory. Reactions were monitored by TLC using pre-coated silica gel aluminium plates (Kiesel gel 60, 254, E. Merck, Germany); zones were detected visually under ultraviolet irradiation or exposing to I_2 vapour. Analytical reversed-phase high performance liquid chromatography (HPLC) was performed on a Shimadzu CLASS-VP system equipped with an auto sampler and a photodiode array detector. Purity of the final compounds was determined using chromatographic systems; column, Princeton SPHER- 100, C18 (particle size = 5 mm, pore size = 10 nm, dimensions = 50mmx 4.6 mm); mobile phase A, methanol; mobile phase B, acetonitrile. Using a flow rate of 1.0 mL/min, gradient elution was performed from 10% B to 90% B over 6 min. In every case, 10 mL of a 20 µl solution was injected.



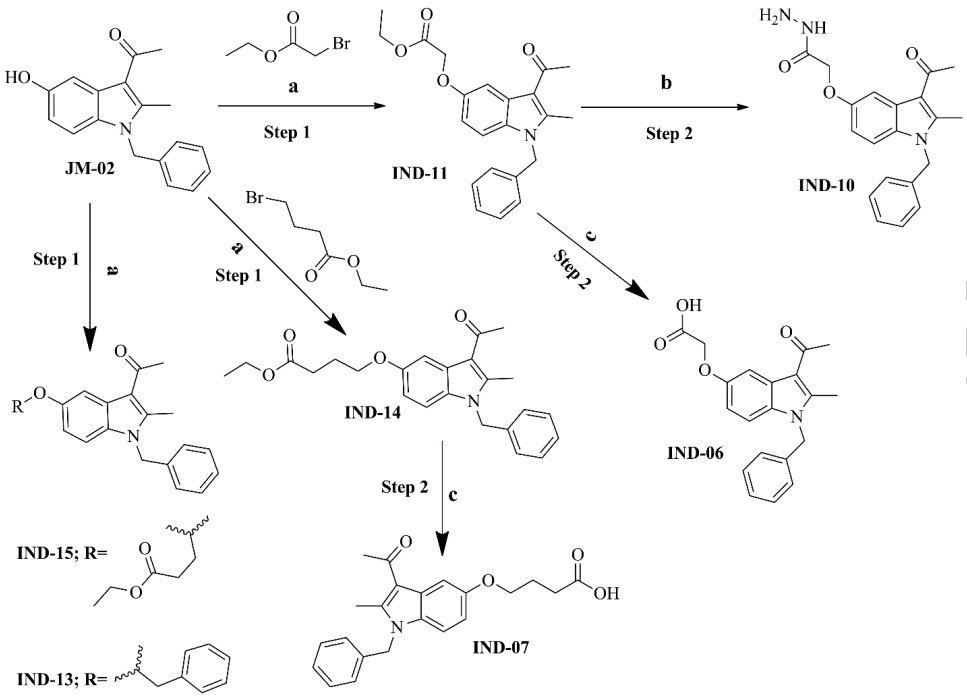
Reagents: a) P-TsOH, MeOH, 45^{0}C, b) $ZnCl_2$, DCM, 45^{0}C, 1h, c) P-TsOH, MeOH, d) $ZnCl_2$, DCM, 45^{0}C, 1h

***Scheme1*:** Synthetic route for generation of starting maaterialindole Scaffold



Reagents: a) K_2CO_3, Acetone, Reflux, 6h, b) Hydrazine Hydrate, Ethanol, 80°c, 16h, c) Hydrazine Hydrate, Ethanol, 80°c, 6h, d) LiOH, THF:H_2O, O°C to 80°C, 1h

***Scheme 2*:** Synthetic route of indoleester derivatives



Reagents: a) K_2CO_3, Acetone, Reflux, 6h, b) Hydrazine Hydrate, Ethanol, 80^0C, 12h, c) LiOH, THF:H_2O, 80^0C, 1h

***Scheme 3*:** Synthesis route of indoleketo derivatives

*2.4.1 Ethyl 1-benzyl-5-hydroxy-2-methyl-1H-indole-3-carboxylate (**JM-01**):*

A mixture of benzyl amine (10.70g, 100 mmol), ethyl acetoacetate (12.98g, 100 mmol) and *p*-TsOH (0.86g, 5 mmol) in MeOH in a 100 mL conical flask were refluxed for 2h at 45 °C. The reaction was followed by TLC and after completion of the reaction (~2h) the reaction mixture was evaporated and quenched with cold water 200 mL and extracted with dichloromethane (3x500 mL). The organic layer was washed successively with saturated sodium bicarbonate, brine and finally with water, dried over anhydrous Na_2SO_4 and then concentrated. The intermediate (**JM_A**) obtained was directly used for next step without purification.

To a solution of 1, 4-benzoquinone (17.75 g, 0.1643 mol) in CH_2Cl_2 (300mL) was added $ZnCl_2$ (14.68 g, 0.1095 mol). The resultant mixture was heated to boiling and a solution of **JM_A**(30.0 g, 0.1369 mol) in CH_2Cl_2 (100 mL) was added drop wise with stirring in 10 minutes. The mixture was stirred under boiling for additional 40 minutes[19]. After completion of the reaction, the reaction mixture was kept at 0-5°C for 4-5 hours. The crystals precipitated were filtered off and washed with CH_2Cl_2 (3 x 1 mL) and acetone (2 x 1 mL) to afford 30.5 g (82%) of **JM-01**, m.p.

212-214 °C,^{1}H NMR (DMSO- d_6): 1.33 (3H, t, CH_3,*J*= 6.8), 2.62 (3H,s, CH_3), 4.245 (2H, q, CH_2, *J*= 6.8), 5.41 (2H, s, CH_2), 6.15 (1H, dd, ArCH, *J*= 8.4), 6.981 (1H, d, ArCH, *J*= 6.8), 7.23 (4H, m, ArCH), 7.39 (1H, s, ArCH), 8.93(1H, s, OH); ^{13}C NMR (100 MHz, DMSO-d_6) δ 166.11, 153.712, 145.559, 136.308, 131.895, 128.958(2C), 127.696(3C), 125.916, 112.587, 110.312, 105.519, 104.480, 46.666, 14.644, 12.032; ESI-MS (*m/z)* : 310.20 $(M+1)^+$.

*2.4.2 1-(1-benzyl-5-hydroxy-2-methyl-1H-indol-3-yl)ethanone (**JM-02**):*

Synthesis of **JM-02** was carried out by following the same procedure as in **JM-01** except that acetyl acetone(100 mmol) was used in place of ethyl acetoacetate (100 mmol) to afford 28 g (79%) of **JM-02**, mp, 206-208°C, ^{1}H NMR (DMSO-d_6): 2.48 (3H, s,CH_3), 2.63 (3H, s, CH_3), 5.43 (2H, s, CH_2), 6.39 (1H, dd,ArCH, *J*=8.8), 7.23-7.35 (5H, m, ArCH), 7.48 (1H, d, ArCH, *J*=8.8), 7.56 (1H, s, ArCH), 8.88(1H, s, OH); ^{13}C NMR (100 MHz, DMSO-d_6) δ 194.443, 153.867, 145.231, 136.044, 132.096, 129.018, 127.696(2C), 125.896, 114.722, 112.147, 110.576, 105.899, 46.666, 31.590, 12.948;ESI-MS (*m/z)* : 280.30 $(M+1)^+$.

*2.4.3 2-((3-acetyl-1-benzyl-2-methyl-1H-indol-5-yl) oxy) acetic acid (**IND-06**)*

To a solution of **IND-11** (1eq) in THF: H_2O (1:1) was added LiOH (3eq) at 0 °C to 80 °C for 1h. After completion of reaction, reaction mixture was acidified with HCl and extracted with ethyl acetate followed by drying of organic layer over sodium sulphate. The evaporation of solvent under vacuum gave white powdered compound.

White solid, Yield= 72%, HPLC purity 100%, R_f = 0.26, mp, 172-174 °C; ^{1}H NMR (400 MHz, DMSO-d_6), δ 2.526 (3H, s, CH_3), 2.642 (3H, s, CH_3), 4.627 (2H, s, CH_2), 5.477 (2H, s, CH_2), 6.805 (2H, dd, Ar-CH, *J*=8.8), 6.999 (2H, dd, Ar-CH, *J*= 6.8), 7.251 (3H, m, Ar-CH), 7.403 (2H, d, Ar-CH, *J*= 8.8), 7.508(1H, s, Ar-CH); ^{13}C NMR (100 MHz, DMSO-d_6) δ 193, 171.18, 154.47, 145.37, 137.61, 131.84, 129.24, 127.20(3C), 114.27, 111.71, 105.05, 65.98, 46.26, 31.80, 13.24; IR (cm^{-1}): 3743.13 (OH), 1699.59 (C=O), 1528.98 (C-O, Bend); elemental analysis (%) calculated/found: C(71.20/71.04); H(5.68/5.55); N(4.15/4.03); O(18.97/18.69); ESI-MS (*m/z)*: 338.17$(M+1)^+$.

*2.4.4 4-((3-acetyl-1-benzyl-2-methyl-1H-indol-5-yl) oxy) butanoic acid (**IND-07**)*

To a solution of **IND-14** (1eq) in THF: H_2O (1:1) mixture was added LiOH (3eq) at 0°C to 80°C for 1h. After completion of reaction, mixture was acidified with HCl, extracted with ethyl acetate and organic layer was dried over sodium sulphate. After removing the solvent under vacuum white powdered compound was obtained which was characterized.

White solid, Yield=69%, HPLC purity 100%, R_f = 0.35, mp, 140-142 °C; ^{1}H NMR (400 MHz, DMSO-d_6), δ 1.938(2H, q, CH_2, *J*=13.2), 2.338 (2H, t, CH_2, *J*=14.4), 2.540 (3H, s, CH_3), 2.646 (3H, s, CH_3), 3.994 (2H, q, CH_2, *J*= 11.6), 5.486 (2H, s, CH_2), 6.795 (2H, d, Ar-CH, *J*=8.8), 6.996 (2H, d, Ar-CH, *J*= 6.8), 7.253 (3H, m, Ar-CH), 7.398 (1H, d, Ar-CH, *J*=8.8), 7.535 (1H, s, Ar-CH); ^{13}C NMR (100 MHz, DMSO-d_6) δ 193.67, 174.72, 155.10, 145.24, 137.65, 131.65, 129.22, 127.20 (3C), 114.32, 111.84, 104.91, 67.49, 46.25, 30.75, 24.96; IR (cm^{-1}): 3558.99 (OH), 1696.95 (C=O), 1410.71 (C-O, Bend); elemental analysis (%) calculated/found: C(72.31/72.04); H(6.34/6.19); N(3.83/3.73); O(17.51/17.49); ESI-MS (*m/z)* :366.12(M+1).

*2.4.5 1-benzyl-5-(3-hydrazinyl-3-oxopropoxy)-2-methyl-1H-indole-3-carbohydrazide (**IND-08**)*

To a solution of **IND-08a** (1eq) in ethanol was added hydrazine hydrate (5eq) at 0 °C to 80 °C for 16h. After the completion of reaction, mixture was poured onto ice, extracted with ethyl acetate and organic layer was dried over sodium sulphate. After removing the solvent under vacuum white powdered compound was obtained which was characterized.

White solid, Yield=82%, HPLC purity 98.62%, R_f = 0.25, mp, 152-154 °C; ^{1}H NMR (400 MHz, DMSO-d_6), δ 1.356 *(3H, t, NH, J=6.8)*, 2.648 *(3H, s, CH_3)*, 4.282 (4H, m, CH_2,NH), 4.478 (2H, s, CH_2), 6.856 (1H, d, Ar-CH, *J*=8.8), 6.974 (2H, d, Ar-CH, *J=7.6*), 7.249 (3H, m, Ar-CH), 7.394 (1H, d, Ar-CH, *J*= 8.8), 7.517(1H, s, Ar-CH); ^{13}C NMR (100 MHz, DMSO-d_6) δ 167.42, 165.43, 154.17, 145.98, 137.63, 131.88, 129.23, 127.33(3C), 114.92, 105.07, 103.62, 67.31, 59.46, 46.40, 14.90, 12.30; IR (cm^{-1}): 3741.06 (NH), 1678.59 (C=O), 1526.45 (C-N, Bend), 1177.90 (C-O, Bend); elemental analysis (%) calculated/found: C(62.98/62.84); H(6.08/6.01); N(18.36/18.03); O(12.58/12.49); ESI-MS (*m/z*) : 380.31 $(M-1)^-$.

*2.4.6 2-((1-benzyl-3-(ethoxycarbonyl)-2-methyl-1H-indol-5-yl) oxy) acetic acid (**IND-09**)*

To a solution of **IND-12** (1eq) in THF: H_2O (1:1) was added LiOH (3eq) at 0 °C to 80 °C for 1h. After completion of reaction, reaction mixture was acidified with HCl and extracted with ethyl

acetate followed by drying of organic layer over sodium sulphate. The evaporation of solvent under vacuum gave white powdered compound.

White solid, Yield=65%, HPLC purity 98.28%, R_f = 0.32, mp, 160-162 ^{o}C; ^{1}H NMR (400 MHz, DMSO-d_6), δ 1.338 (3H, t, CH_3, J=7.2), 2.638 (3H, s, CH_3), 4.248 (2H, q, Ar-CH, J= 6.8), 4.566 (2H, s, CH_2), 5.452 (2H, s, CH_2), 6.781 (1H, d, Ar-CH, J= 8.8), 6.972 (2H, d, Ar-CH, J= 7.2), 7.254 (3H, m, Ar-CH), 7.365 (1H, d, Ar-CH, J= 8.8), 7.448 (1H, s, Ar-CH); ^{13}C NMR (100 MHz, DMSO-d_6) δ 171.33, 165.65, 154.33, 145.87, 137.68, 131.68, 129.22, 127.20, 111.89, 104.59, 66.08, 59.42, 46.83, 14.82, 12.28;IR (cm^{-1}): 3649.86 (OH), 1689.59 (C=O), 1184.76 (OH, Bend); elemental analysis (%) calculated/found: C(68.65/70.62); H(5.76/5.45); N(3.81/3.78); O(21.77/21.49); ESI-MS (*m/z) :* 366.19 $(M-1)^{-}$.

*2.4.7 2-((3-acetyl-1-benzyl-2-methyl-1H-indol-5-yl) oxy) acetohydrazide (**IND-10**)*

To a solution of **IND-11** (1eq) in ethanol was added hydrazine hydrate (1.2eq) at 80 ^{o}C for 12h. After the completion of reaction, mixture was poured onto ice, extracted with ethyl acetate and organic layer was dried over sodium sulphate. After removing the solvent under vacuum white powdered compound was obtained which was characterized.

White solid, Yield=77%, HPLC purity 100%, R_f = 0.22, mp, 148-150 ^{o}C; ^{1}H NMR (400 MHz, DMSO-d_6), δ 2.548 (3H, s, CH_3), 2.651 (3H, s, CH_3), 4.331 (2H, s, NH_2), 4.454 (2H, s, CH_2), 5.493 (2H, s, CH_2), 6.871 (1H, dd, Ar-CH, *J=8.8*), 6.974 (2H, m, Ar-CH), 7.255 (3H, m, Ar-CH), 7.425(1H, d, Ar-CH, J=8.8), 7.572 (1H, s, Ar-CH), 9.348(1H, s, NH); ^{13}C NMR (100 MHz, DMSO-d_6) δ,193.65, 167.59(2C), 154.37, 145.48, 137.67, 131.97, 129.24(2c), 127.12(3c), 114.27, 111.80(2C), 105.48, 67.45, 46.25, 31.89, 13.23; IR (cm^{-1}): 3632.40 (OH), 1674.57 (C=O), 1614.55(C=O), 1164.81 (OH, Bend); elemental analysis (%) calculated/found: C(71.20/71.09); H(5.68/5.57); N(4.15/3.99); O(18.97/18.79); ESI-MS (*m/z) :* 352.22 $(M+1)^{+}$.

*2.4.8 Ethyl 2-((3-acetyl-1-benzyl-2-methyl-1H-indol-5-yl) oxy) acetate (**IND-11**)*

To a mixture of JM-02 (1eq) and ethylbromoacetate (1.2eq) in 50 ml dry acetone was treated with anhydrous K_2CO_3 (3eq) and allowed to reflux for 6 h. The reaction mixture was then filtered and washed thoroughly with acetone. The solvent was evaporated, the product thus separated was purified by recrystallization from ethanol to afford the white crystalline compound.

White solid, Yield=78%,HPLC purity 97.35%, R_f = 0.56, mp, 240-242 ^{o}C; ^{1}H NMR (400 MHz, DMSO-d_6), δ 1.203 (3H, t, CH_3, J=7.2), 2.534 (3H, s, CH_3), 2.650(3H, s, CH_3), 4.160(2H, q, CH_2, J=14.4), 4.802 (2H, s, CH_2), 5.492 (2H, s, CH_2), 6.827 (1H, dd, Ar-CH, J= 8.8), 7.000 (2H, d, Ar-CH, J=7.2), 7.255 (3H, m, Ar-CH), 7.429(1H, d, Ar-CH, J=8.8), 7.512(1H, d, Ar-CH, J=2.4); ^{13}C NMR(100 MHz, DMSO-d_6) δ, 193.61, 169.55, 154.18, 145.50, 137.59, 132.01, 129.24, 127.12(3c), 114.31, 111.75(2C), 105.24, 65.94, 61.02, 46.28, 31.79, 14.57, 13.24; IR (cm^{-1}): 3103.28 (Ar-Str-CH), 1621.07 (C=O), 1449.49(C-O); elemental analysis (%) calculated/found: C(72.31/67.66); H(6.34/6.22); N(3.83/3.77); O(17.51/17.45); ESI-MS (*m/z)* :366.17$(M+1)^{+}$.

*2.4.9 Ethyl 1-benzyl-5-(2-ethoxy-2-oxoethoxy)-2-methyl-1H-indole-3-carboxylate (**IND-12**)*

To a mixture of JM-01 (1eq) and ethylbromoacetate (1.2eq) in 50 ml dry acetone was treated with anhydrous K_2CO_3 (3eq) and allowed to reflux for 6 h. The reaction mixture was then filtered and washed thoroughly with acetone. The solvent was evaporated, the product thus separated was purified by recrystallization from ethanol to afford the white crystalline compound.

White solid, Yield=85%, HPLC purity 92.43%, R_f = 0.36, mp, 120-122 ^{o}C; ^{1}H NMR (400 MHz, DMSO-d_6), δ 1.185 (3H, t, CH_3, J=7.2), 1.343 (3H, t, CH_3, J=7.2), 2.649 (3H, s, CH_3), 4.156 (2H, Q, CH_2, J=14.4), 4.262 (2H, Q, CH_2, J=14.4), 4.757 (2H, S, CH_2, J=6.4), 5.469 (2H, S, CH_2), 6.811 (1H, q, Ar-CH, J=8.8), 6.978 (2H, d, Ar-CH, J=7.2), 7.250 (3H, m, Ar-CH), 7.414 (2H, m, Ar-CH, J=8.8); ^{13}C NMR(100 MHz, DMSO-d_6) δ, 169.61, 165.40, 153.97, 146.11, 137.61, 131.90, 129.23, 127.12(3c), 112.02(2C), 104.72, 65.77, 61.04, 59.43, 46.41, 31.79, 14.65, 12.25; IR (cm^{-1}): 3103.28 (Ar-Str-CH), 1752.53 (C=O), 1682.58(C=O), 1527.93(C-O); elemental analysis (%) calculated/found: C(69.86/69.77); H(6.37/6.19); N(3.54/3.51); O(20.23/19.99); ESI-MS (*m/z)* :396.22 $(M+1)^{+}$.

*2.4.10 1-(1-benzyl-5-(benzyloxy)-2-methyl-1H-indol-3-yl) ethanone (**IND-13**)*

To a mixture of JM-02 (1eq) and benzylbromide (1.2eq) in 50 ml dry acetone was treated with anhydrous K_2CO_3 (3eq) and allowed to reflux for 6 h. The reaction mixture was then filtered and washed thoroughly with acetone. The solvent was evaporated, the product thus separated was purified by recrystallization from ethanol to afford the white crystalline compound.

White solid, Yield=86%, HPLC purity 99.93%, R_f = 0.58, mp, 180-182 °C; ^{1}H NMR (400 MHz, DMSO-d_6), δ 2.532 (3H, s, CH_3), 2.464 (3H, s, CH_3), 5.122 (2H, s, CH_2), 5.478 (2H, s, CH_2), 6.888 (1H, d, Ar-CH, *J*=8.8), 7.000(2H, s, Ar-CH), 7.329 (9H, m, Ar-CH), 7.638 (1H, s, Ar-CH);^{13}C NMR(100 MHz, DMSO-d_6) δ 193.64, 154.92, 145.31, 138.01, 137.62, 131.75, 128.76(5C), 114.31, 112.19, 105.42, 70.20, 46.26, 31.86, 13.24; IR (cm^{-1}): 3030.28 (Ar-Str-CH), 1617.83 (C=O), 1485.53 (C-O); elemental analysis (%) calculated/found: C(81.27/81.06); H(6.27/6.21); N(3.79/3.59); O(8.66/8.58); ESI-MS (*m/z):* 370.21 $(M+1)^+$.

*2.4.11 Ethyl 4-((3-acetyl-1-benzyl-2-methyl-1H-indol-5-yl) oxy) butanoate (**IND-14**)*

To a mixture of JM-02 (1eq) and ethylbromobutyrate (1.2eq) in 50 ml dry acetone was treated with anhydrous K_2CO_3 (3eq) and allowed to reflux for 6 h. The reaction mixture was then filtered and washed thoroughly with acetone. The solvent was evaporated, the product thus separated was purified by recrystallization from ethanol to afford the white crystalline compound.

White solid, Yield=69%, HPLC purity 100%, R_f = 0.68, mp, 232-234 ^{o}C; ^{1}H NMR (400 MHz, DMSO-d_6), δ 1.159 (3H, t, CH_3, *J*=7.2), 1.975 (2H, t, CH_2, *J*=6.8), 2.463 (2H, t, CH_2, *J*=7.6), 2.538 (3H, s, CH_3), 2.646 (3H, s, CH_3), 4.027 (4H, m, CH_2), 5.481 (2H, s, CH_2), 6.784 (1H, d, Ar-CH, *J*=8.8), 7.001 (2H, d, Ar-CH, *J*=7.2), 7.252 (3H, m, Ar-CH), 7.396 (1H, d, Ar-CH, *J*=8.8), 7.537 (1H, s, Ar-CH); ^{13}C NMR(100 MHz, DMSO-d_6) δ 193.64, 173.11, 155.06, 145.24, 137.64, 131.61, 129.21, 127.31(3C), 114.33, 104.94, 67.38, 60.33, 46.25, 31.84, 30.74, 24.93, 14.58, 13.26;IR (cm^{-1}): 2903.25 (Ar-Str-CH), 1703.76 (C=O), 1526.28 (C-O); elemental analysis (%) calculated/found: C(73.26/73.21); H(6.92/6.83); N(3.56/3.51); O(16.26/15.99); ESI-MS (*m/z) :*416.25 $(M+Na)^+$.

*2.4.12 Ethyl 3-((3-acetyl-1-benzyl-2-methyl-1H-indol-5-yl) oxy) propanoate (**IND-15**)*

To a mixture of JM-02 (1eq) and ethylbromopropionate (1.2eq) in 50 ml dry acetone was treated with anhydrous K_2CO_3 (3eq) and allowed to reflux for 6 h. The reaction mixture was then filtered and washed thoroughly with acetone. The solvent was evaporated, the product thus separated was purified by recrystallization from ethanol to afford the light brown compound.

Light brown solid, Yield=55%, R_f = 0.62, mp, 215-217 ^{o}C; ^{1}H NMR (400 MHz, DMSO-d_6), δ 2.487 (3H, s, CH_3), 2.629 (3H, s, CH_3), 5.436 (2H, s, CH_2), 6.641 (1H, dd, Ar-CH, *J*=8.8), 6.702

(2H, dd, Ar-CH, *J*=7.2), 7.265 (4H, m, Ar-CH), 7.428 (1H, d, Ar-CH, *J*=2.4); ^{13}C NMR(100 MHz, DMSO-d_6) δ 193.45, 153.61, 144.93, 137.72, 130.87, 129.20, 127.60(3C), 113.95, 111.95, 106.03, 46.18, 31.77, 13.23; IR (cm^{-1}): 2903.25 (Ar-Str-CH), 1703.76 (C=O), 1526.28 (C-O); elemental analysis (%) calculated/found: C(72.80/72.71); H(6.64/6.63); N(3.69/3.59); O(16.87/16.81); ESI-MS(*m/z):* 402.20 $(M+Na)^+$.

*2.4.12 Ethyl 1-benzyl-5-(2-hydrazinyl-2-oxoethoxy)-2-methyl-1H-indole-3-carboxylate (**IND-16**)*

To a solution of **IND-12** (1eq) in ethanol was added hydrazine hydrate (1eq) at 80 ^{o}C for 6h. After the completion of reaction, mixture was poured onto ice, extracted with ethyl acetate and organic layer was dried over sodium sulphate. After removing the solvent under vacuum white powdered compound was obtained which was characterized.

White solid, Yield=77%, HPLC purity 100%, R_f = 0.55, mp, 198-200^{o}C; ^{1}H NMR (400 MHz, DMSO-d_6), δ 1.358 *(3H, t, CH_3, J=7.2),* 2.649 (3H, s, CH_3), 4.285 (4H, m, CH_2, NH), 4.481 (2H, S, CH_2), 5.468 (2H, s, CH_2), 6.854 (1H, dd, Ar-CH, *J*=8.8), 6.974 (2H, d, Ar-CH, *J*=7.2), 7.259 (3H, m, Ar-CH), 7.395 (1H, d, Ar-CH, *J*=8.8) 7.517 (1H, d, Ar-CH, *J*=2.4); ^{13}C NMR(100 MHz, DMSO-d6) δ 167.42, 165.43 154.18, 145.98, 137.63, 131.89, 129.23, 127.23(3C), 111.83(2C), 105.10, 103.63, 67.31, 59.46, 46.40, 14.90, 12.29; IR (cm^{-1}): 3283.58 (NH_2), 2963.25 (Ar-Str-CH), 1682.34 (C=O), 1532.95 (C-O); elemental analysis (%) calculated/found: C(72.80/72.71); H(6.64/6.63); N(3.69/3.59); O(16.87/16.81); ESI-MS (*m/z)*: 404.23 $(M+Na)^+$.

2.5 Biological Assay:

2.5.1 MIC against H_{37}Rv assay:

The biological assay was performed by TAACF center, NIAD, USA. The *MIC* of compound was determined by measuring bacterial growth after 5 days in the presence of test compounds. Ten–point two-fold serial dilutions of compounds were prepared in DMSO and diluted with 7H9-Tw-OADC medium in 96-well plates with a final DMSO concentration of 2%. The highest concentration of compound was 200 µM where compounds were soluble in DMSO at 10 mM. For compounds with limited solubility, the highest concentration was 50x less than the stock concentration e.g. 100 µM for 5 mM DMSO stock, 20 µM for 1 mM DMSO stock. For potent compounds, assays were repeated at lower starting concentrations. Each plate included assay

controls for background (medium/DMSO only, no bacterial cells), zero growth (100 µM rifampicin) and maximum growth (DMSO only), as well as a rifampicin dose response curve.OD_{590} and fluorescence (Ex 560/Em 590) were measured using BioTek™ Synergy 4 plate reader for *M. tuberculosis* inoculated plates which successively incubated for 5 days. Growth was calculated separately for OD_{590} and RFU. To calculate the MIC, the 10-point dose response curve was plotted as % growth and fitted to the Gompertz model using Graph Pad Prism 5. (*Figure 5A*). In addition, dose response curves were generated using the Levenberg-Marquardt algorithm and the concentrations that resulted in 50% and 90% inhibition of growth were determined (IC_{50} and IC_{90} respectively) (*Figure 5B*)[20].

***Figure 5*:** Dose response curve used to calculate *MIC*, IC_{50}, and IC_{90}. Data points obtained from a dose response growth inhibition assay are curve fitted using (A) the Gompertz model to calculate *MIC* and (B) the Levenberg-Marquardt algorithm to calculate IC_{50} and IC_{90}.

2.5.2 In vitro Mtb-DHFR *and* h-DHFR *enzyme assay:*

DHFR enzyme was expressed and purified as described in literature (30). Enzyme assays were performed in 100 mM HEPES, 50 mMKCl, pH 7.0 at 25 ^{0}C. The decrease in absorbance at 340 nm representing the oxidation of NADPH was monitored with a spectrophotometer. Solution of test compound (**IND-07**) was added to a 1 mL cuvette at various concentrations, with the individual *DHFR* s (20 nM), and 40 µM of NADPH, and the reaction was initiated by the addition of 40 µM of dihydrofolate. For IC_{50} determinations, 20 nM of *DHFR* s was incubated with 40 µM of cofactor NADPH and seven serially diluted concentrations of the inhibitors for 1 minute. The reaction was initiated by the addition of 40 µM dihydrofolate[21-22].

2.6 Molecular dynamic simulations:

The synthesized compound (**IND-07**) which showed highest activity was selected for MD simulation. MD simulations of **IND-07** at the binding site of Mtb-DHFR (PDBID: 1DF7) were performed using the Desmond package incorporated in the Maestro. The system was built by applying OPLS-AA force field in an explicit solvent with the single point charges (SPC) water model (OPLS-AA/SPC). The initial coordinates for the MD calculations were taken from the docking experiments. The SPC water molecules were then added and system was neutralized by adding Na^{+} counter-ion to balance the net charges of the system. After the construction of the

solvent environment, the complex system was composed of approximately 25241 atoms. Before equilibration and long production MD simulations, the systems were minimized and pre-equilibrated using the default relaxation routine implemented in Desmond. The MD simulations were run for 10 ns and during the MD simulations, the equations of motion were integrated with a 2 fs time step in the NVT ensemble. The SHAKE algorithm was applied to all hydrogen atoms; the van der Waals (VDW) cutoff was set to 9Å. The temperature was maintained at 300 K, employing the Nosé–Hoover thermostat method with a relaxation time of 1ps. The trajectory recording interval was kept for every 10 ps during entire MD runs[23].

2.7 AMDET and MM/GBSA Study:

Eleven compound synthesized were subjected to ADMET analysis by QikProp (Qikprop, Version 3.5). It provides ranges for comparing the properties of a particular molecule with those of 95% of known drugs. The descriptors calculated were partition coefficient, human oral absorption, CNS activity and gut-blood barrier permeability.

The free binding energy studies of complexes of final 11 compounds with 1DF7, was carried out by running MM/GBSA energy calculations implemented in Prime module of the Schrödinger molecular modeling package. Complex structures from last 10 ns (i.e., 1 trajectory frame in each 2ns) of the trajectory frames were selected and binding free energies of corresponding structures were calculated using VSGB 2.0 solvation model and OPLS3 force field to predict binding free energies[24].

3. Results and discussion:

3.1 Comparative study of binding pocket of Mtb-DHFR *and* h-DHFR*:*

This intention of this study was to identify the selective structural features required to inhibit Mtb-DHFR and to develop novel compounds for selective inhibition of the same, through computational studies and wet lab techniques. Since DHFR is present in both humans as well as bacteria, selectivity becomes an important criterion for design and development of the ligands[25]. For developing selective inhibitors for Mtb-DHFR, VS protocol was run with two target proteins Mtb-DHFR (PDB Id: 1DF7) and h-DHFR (PDB Id: 1OHJ). A database of 2,73,951 compounds

(2,65,242 molecule from NCI and 8709 from drug bank) was screened against Mtb-DHFR and the results obtained were screened against h-DHFR for selectivity.

It is very important to understand the composition of active site of the two target proteins to explore and use the difference therein for development of selective ligands. The comparative analysis of the two target proteins Mtb-DHFR (PDB: 1DF7) and h-DHFR (PDB: 1OHJ) was carried out and difference in the type of amino acid making catalytic triad and other important amino acids involved in interaction with ligand was noted. The results (**Table 2**) obtained were similar to the experimental results presented by Li R. *et al*[12] but what could be inferred from *in silico* studies was not possible from the experimental data. As per the X-ray crystallographic data two interactions are important for Mtb-DHFR inhibition i.e. H-bond interaction with Asp27 and Arg32 and for h-DHFR inhibition the two important interactions are Val115 andGln35. This was also observed in cross dockingstudies of the two ligands MTX and COP with the two target proteins. What caught our attention was that for inhibition of Mtb-DHFR, formation of H-bond with Asp27 and Arg32 was crucial as it could be seen in docking results of both the ligands (MTX and COP) and similarly the docking results of MTX and COP with 1OHJ revealed formation of H-bond with Val115 and Gln35 is crucial for h-DHFR inhibition. This supports the fact that MTX and COP are non-selective inhibitors as both of them bind to crucial amino acids in the two proteins. Second observation from the study was π stacking of MTX and COP with Phe31 and Phe34 in Mtb-DHFR and h-DHFR active site respectively. These results provide very important insight for the development of selective inhibitors of Mtb-DHFR and that is ligands must forms H-bond with Asp27 and Arg32 and not with Val115 andGln35 and must shows π-stacking with Phe31 only.

Table 2*:* Important interaction of Methotrexate, COP and Hit1 identified with the active site of Mtb-DHFR and h-DHFR

For the same purpose the binding pocket of the two proteins were critically studied for size and type of small molecules that could bind in the active site. Studies have revealed that in addition to the MTX in active site of Mtb-DHFR, a glycerol molecule is also present whereas in h-DHFRpocket is filled with COP only (*Figure 3a and 3b*). The

glycerol binding site near DHFR binding site is relatively small in size and may be treated as extension of the DHFR site. In h-DHFRs, the active site is well packed with three hydrophobic residue side chains, Leu22, Pro26 and Phe31 and no extension could be observed suggesting a compounds with side chain which could mimic the binding mode of glycerol to protein, may bind to Mtb-DHFRselectively as also suggested by Wei H. et al.[26] Such a ligand will be sterically and chemically hindered from forming a complex with h-DHFR. These hypothesis have been used as basis for present study and for identifying novel selective hits a virtual screening protocol was run using two databases viz. NCI and Drug Bank database (2,73,951 compounds) against the two selected targets.

3.2 Virtual Screening and MD Simulation study:

The virtual screening protocols use docking to knock out the hits which are not of interest i.e. don't show good interactions with the target. Therefore it is crucial to use correct docking strategy and validate the procedure before proceeding to VS[27]. The docking study on Mtb-DHFR was carried out without bound glycerol in the pocket. The docking procedure was evaluated by re-docking the co- crystal ligand and reproducing the crystal orientation of the co-crystal ligand. The methotrexate and COP were extracted, refined and docked into their respective protein active sites using the similar parameters which were set for virtual screening and RMSD were calculated between their docked poses and co-crystal conformations. The RMSD value of 1.2Å for methotrexate and COP docked pose with their corresponding crystal orientation was obtained indicating the correctness of the docking procedure. Also the performance of the virtual screening is evaluated by its ability to pick the active compounds for a particular target present in the database. To check the same 20 actives were added into the database and docked compounds was performed. The docking experiment resulted in 20 actives as top scoring ligands among the top 100 compounds. Moreover, all the low energy conformers of the co-crystalized ligands were on the top of the docking scores.

In the VS protocol the selected databases (2,73,951) were screened against Mtb-DHFR (1DF7) using virtual screening protocol (HTVS, SP and XP) of GLIDE module (Schrödinger). The first screening which was high throughput in nature resulted in 17877

compounds which were subjected to SP screening (1225 compounds) followed by XP screening which resulted in 122 compounds. From these results top 100 hits (72 from NCI and 28 from Drug Bank) were selected on the basis of docking score, binding interactions, visual analysis and removal of duplicates and other molecules reported with Mtb-DHFR activity. The hits selected were inspected manually to cover maximum amount of chemical space and screened against h-DHFR (PDB: 1OHJ). To work towards selectivity, 50 hits were shortlisted after analyzing the binding interaction and visual inspection but this time the compounds with lowest docking score on 1OHJ and high score with Mtb-DHFR were selected assuming that these will be selective towards Mtb-DHFR. For reassuring the results obtained and selectivity consensus docking was performed. Since different docking software's use different algorithms, so compounds which show results in consensus with each other might be potential hits. Therefore, consensus docking was carried out by using GLIDE and GOLD 5.2.2. Top ten hits were selected and analyzed visually and for synthetic feasibility. On the basis of synthetic feasibility **hit-1** was selected, synthesized, derivatized and evaluated for biological activity. To improve upon the chemical space and interaction of the hits, 30 derivatives of **hit 1** were designed and docked on the Mtb-DHFR and h-DHFR. On the basis of docking ratio and synthetic feasibility 11 compounds were finally synthesized. The structure of synthesized compounds was confirmed by spectral analysis. All the synthesized compounds were screened for *in vitro* antimycobacterial assay on $H_{37}R_V$. From the literature, we could be observed that most of the compounds reported for inhibitors of DHFR are either pteridine or pyrimidine based with an exception of one or two. The advantage of our study was identification of completely new class of Mtb-DHFR inhibitors which are indole based.

3.3 Classification of Mtb-DHFR *catalytic domain:*

Hit-1 exhibit four sites (***Figure 6***) of modifications which could be explored to design novel selective Mtb-DHFR inhibitors. Here are presented some assumptions based on the *insilico* studies of the targets for the selectivity against Mtb-DHFR inhibition.

***Figure 6*:** Basic indole moiety and sites of modification

3.3.1 Site-1 Acidic side chain:

The acidic group of this side chain is crucial for forming salt bridge between two positively charged arginine residues (Arg32 and Arg60). This salt bridge interaction is very important as can be seen with methotrexate also. The benzene ring along with the side chain of hit -1 corresponds to the *p*-amino benzoic acid part of methotrexate. Replacement of acidic group with sulphate, sulphone, sulphonamide and phosphate was found to decrease glide score which may be due to their inability to form the proper salt bridge. Position of this side chain is also very crucial for the better binding of hit-1 to the receptor. Substitution at 4 and 5^{th} position decrease binding energy as reflected by decreased glide score while substitution at 6 and 7^{th} position are favourable for binding. Optimum binding was observed with substitution at 6^{th} position because the side chain length is optimum for forming the salt bridge with Arg32 and Arg60.

3.3.2 Site-2 Acid amide group modification:

This part corresponds to the N12 of methotrexate which faces the NADPH. At this site methotrexate does not form any hydrogen bond whereas hit-1 forms H-bond with Ser49. This part may help in attaining the selectivity by properly positioning the ligands in the pocket. Replacing this acid amide group with acetic acid, propionic acid, propionamideand diketoamide groups reduced the binding scores and also the selectivity as it failed to form hydrogen bond with Ser49. Thus the optimum group at this position is ketone group.

3.3.3 Site-3 Modification:

Site 3 is sterically hindered site as the adjacent sites bear large substitutions. Therefore modification at this site is not favorable. Changing the methyl group with bigger substituents e.g. ethyl, propyl, butyl or phenyl does not allow acid amide group to form H-bond with Ser49 and thus reducing the interaction of the compounds with receptor.

3.3.4 Site-4 Modification:

One of the critical differences between Mtb-DHFR and h-DHFR structure is glycerol binding site present in Mtb-DHFR active site. The docked pose of Mtx in Mtb-DHFRrevealed that N7 and C8 atom of methotrexate are accessible to solvent and are facing hydrophilic side chains. At this site an L type of narrow cavity formed by residues including Phe31, Asp27, Gln28, Ile94, Ile5, Tyr100, Ala7, Trp6, His30, Gly95 NADP and glycerol. Therefore the opportunity to design selective Mtb-DHFR inhibitors lies at this site also i.e. to design the molecules which possess site chain characteristics of glycerol and could bind to the glycerol binding pocket. So the molecules which could accommodate in this cavity can be selective Mtb-DHFR inhibitors because this cavity is different in h-DHFR enzyme. The best virtual hit (Hit 1) found to occupy the same pose and interaction as methotrexate as well as the benzyl part is aligned towards the negatively charged Asp27 which is catalytic amino acid and responsible for reductase activity. This interaction is most important for Mtb-DHFR selectivity as glycerol also forms H-bond with acidic group of Asp27 and is absolutely absent in h-DHFR. So the functional groups which are forming hydrogen bond with this amino acid will definitely have the Mtb-DHFRactivity.

Another interesting finding of the docking analysis is that the ligands which could from π-stacking interaction with Phe31 in Mtb-DHFRcan be selective Mtb-DHFRinhibitors as it is absent in h-DHFR(π-stacking with Phe34). Phe31 residue forms π-stacking interaction with pteridine ring of methotrexate while in h-DHFRit forms π-stacking with Phe34 so it active in both human and mycobacterium. But the Hit 1 ligand forms π-stacking with Phe31 in Mtb-DHFR and not with Phe34. Thus this finding suggests that the structure which forms π-stacking with Phe31 in Mtb-DHFR and not with Phe34 in h-DHFR will provide the selective inhibition towards Mtb-DHFR.

Therefore modification at this site provides further opportunities for developing selective Mtb-DHFR inhibitors. The modification at site 4 were done in order to optimization the activity using hydrophobic and hydrophilic substitutions like aliphatic moieties, benzyl ring with or without substitution etc. I*n silico* studies were carried out to observe the effect of these groups on the selectivity of these agents towards Mtb-DHFR. It was observed that hydrophilic groups like hydroxyl and methoxy substitutions were more selectivity towards Mtb-DHFR compared to hydrophobic substituents like halogen. This

may be due to the fact that in Mtb-DHFR the amino acids facing pteridine (Trp22, Asp27 and Gln28) are hydrophilic in nature while amino acids facing pteridine ring (Leu22, Pro26 and Phe31) in h-DHFR are hydrophobic (***Figure 3a*** *and* ***3b***). Hydrophobic groups do not form hydrogen bond with Asp27 however they form π-stacking interaction with Phe31. Most important interactions of indole ligand, methotrexate and COP are summarized in ***Table 1***.

Figure 7**:** (A) Docking pose of *Hit 1* against Mtb-DHFRin GOL catalytic domain, (B) Ligplot of *Hit 1* against Mtb-DHFR catalytic domain (C) Binding interaction of **IND-07**against Mtb-DHFR, (D) Ligplot of **IND-07**against Mtb-DHFR, (E) Docking pose of MTX against Mtb-DHFR, (F) Ligplot of MTX against Mtb-DHFR.

Table 3: Docking score of synthesized compounds against Mtb DHFR and h DHFR, their i*n vitro* antimycobacterial activity and enzymatic inhibition data

3.4 Chemistry:

As a part of our ongoing research, number of protocols were developed for synthesis of indole moiety. Intermediate **JM-01** and **JM-02** were obtained through $\mathbf{JM_A}$ and $\mathbf{JM_B}$ from benzylamine reaction with ethylacetoacte and acetyl acetone respectively.

From the two intermediates **JM-01 and JM-02** rest of the compounds were synthesized by condensation reactions with bromoethylacetate/ethyl-4-bromo butyric ester followed by either hydrolysis with lithium hydroxide or treatment with hydrazine hydrate[28].
The final compounds were purified by crystallization from ethanol and the structures were confirmed by IR, ^{1}H NMR, Mass and elemental analysis. ^{1}H NMR of all the compounds showed a singlet corresponding to CH_2 proton between δ 5.4-5.6 and CH_3 protons could be located around δ 2.5-2.7. The aromatic protons were observed as triplet, doublet and multiples range δ 6.66 to 7.60. The IR spectra showed characteristic peaks at 3200 to 3400 (NH str.), 1670 to 1700 (C=O str.), 2925 to 3050 (Ar C-H str.), 1570 to 1590 (C=C Str.). The yield of synthesized compounds was between 60-85%.

3.5 Biological Evaluation:

3.5.1 MIC Antimycobacterial assay:

The *MIC* of the compounds was determined by measuring bacterial growth after 5 days in the presence of test compounds against *M. tuberculosis* $H_{37}Rv$ colonies[29]. For potent compounds, assays were repeated at lower concentrations. All the 11 compounds showed significant inhibition against $H_{37}Rv$ cell lines between ranges (25 to 200µM). Out of 11 compounds **IND-07** showed highest activity and was selected for *in vitro* testing against Mtb-DHFR and h-DHFR.

Figure 8: Dose response curve for **IND-07**

3.5.2 In vitro Mtb-DHFR *and* h-DHFR *enzyme assay:*

The selectivity of the compound towards Mtb-DHFR over human DHFR was done by performing enzyme assay using h-DHFR. The identified compound was found selective towards Mtb-DHFR. Most potent derivative (**IND-07**) inhibited the two enzymes at an IC_{50} of 150 µM and 980 µM respectively. The selectivity index was found to be 6.53 (***Table 3***).

3.6 MD Simulation Studies:

The stability of docked pose at the binding site of Mtb-DHFR (PDBID: 1DF7) was evaluated by molecular dynamics simulation. The RMSD of protein backbone and heavy atoms during the entire 10 ns MD simulation are given in ***Figure 9***. The MD simulation studies revealed that the compound (**IND-07**) is highly stable within the active pocket of the enzyme. The studies displayed that the compound interacts profoundly with important amino acids viz. Arg60, Arg32, Ser49 and Phe31. The comparison of the MD simulation studies of compound (**IND-07**) and the reference ligand (MTX) displayed similar interaction.

***Figure 9*: (A)** Time dependence of the total energy and protein backbone, (**B**) RMSD relative to the initial minimized complex of molecule during MD simulations.

The evaluation of MD trajectories reveals that the protein-ligand complex was stable after initial 1 ns period. Throughout the simulation studies, the molecule **IND-07** maintained

its interaction with active site residues of protein. Both RMSD values remained within the range 1 Å for the system after reaching equilibrium, which further proved the conformational stabilities of the protein-ligand complex.

3.7 ADMET and MM/GBSA Studies:

The ADME properties of the 11 compounds were analyzed using Qikprop. The result of Qikprop analysis of most potent compound, **IND-07** is presented ***Table 4***. MM/GBSA studies of **IND-07** revealed that its binding free energy is less than that of MTX which is -84.83 and -23.58 respectively indicating that the compound **IND-07** may be stable in the receptor pocket (***Table 4***).

Table 4: ADMET analysis and MM/GBSA binding score

4. Conclusion:

By using *in silico* approach, the glycerol binding site of Mtb-DHFR binding pocket was exploited successfully to identify and develop novel inhibitors for same. Based on the virtual screening **hit 1** was obtained, and a series of molecules was designed, by using different substituents at 3, 5 position of indole moiety of hit 1. Compound **IND-07** was developed as novel selective inhibitors of Mtb-DHFR. It contains a novel central core (1-(1-benzyl-5-hydroxy-2-methyl-1H-indol-3-yl) ethanone), which will significantly expand the chemical space of novel Mtb-DHFR inhibitors. This compound did show selectivity against Mtb-DHFR and can be used as a lead compound for further optimization. Further, the present findings will help in designing novel Mtb-DHFR inhibitors in future to improve the selectivity of compounds against Mtb-DHFR.

Software used: (a) Schrödinger Suite 2017 Protein Preparation Wizard; LLC, New York, NY, Glide 7.1 Schrödinger, LLC; New York, NY, 2017, LigPrep, version 4.6, Schrödinger, LLC, New York, NY, 2017. (b) Molinspiron property calculation program.

Acknowledgements

The authors wish to acknowledge the Department of Biotechnology (DBT), Govt of India, for providing Bioinformatics infrastructure facility (via letterno.AS/MP (RES.)/JH-5/2013). The

author KC Sharma thanks to C. Karthikeyan for his valuable support for this study. We gratefully acknowledge the TAACF (Tuberculosis Antimicrobial Acquisition Coordinating Facility, National Institute of Allergy and Infectious Diseases (Contract No. HHSN272201100009I), USA for carrying out the biological activity. One of the authors (KC) also thanks Dr. G S Deora, Monash Institute of Pharmaceutical Sciences, Monash University, and Melbourne, Australia for his help in MD Simulation studies.

Conflict of interest:

The authors report no declarations of interest.

† **Supplementary data:** The supporting information includes spectral data details (NMR, IR, HRMS, HPLC and Mass), MD simulation and molecular modelling study of the compounds. This material is available free of charge.

References:

1. Argyrou, A.; Vetting, M. W.; Aladegbami, B.; Blanchard, J. S. *Nature Structural & Molecular Biology* **2006**, *13*, 408.
2. Hagner, N.; Joerger, M. *Cancer Management and Research* **2010**, *2*, 293.
3. World Health Organisation, Tuberculosis fact sheet, Available at; http://www.who.int/mediacentre/factsheets/fs104/en, data accessed: December 2017.
4. Robson, C.; Meek, M. A.; Grunwaldt, J. D.; Lambert, P. A.; Queener, S. F.; Schmidt, D.; Griffin, R. J. *Journal of Medicinal Chemistry* **1997**, *40*, 3040.
5. Russell, P. B.; Hitchings, G. H. *Journal of the American Chemical Society* **1951**, *73*, 3763.
6. Hekmat-Nejad, M.; Rathod, P. K. *Experimental Parasitology* **1997**, *87*, 222.
7. Yuthavong, Y.; Tarnchompoo, B.; Vilaivan, T.; Chitnumsub, P.; Kamchonwongpaisan, S.; Charman, S. A.; McLennan, D. N.; White, K. L.; Vivas, L.; Bongard, E.; Thongphanchang, C.; Taweechai, S.; Vanichtanankul, J.; Rattanajak, R.; Arwon, U.; Fantauzzi, P.; Yuvaniyama, J.; Charman, W. N.; Matthews, D. *Proceedings of the National Academy of Sciences of the United States of America* **2012**, *109*, 16823.
8. Kobayashi, M.; Kinjo, T.; Koseki, Y.; Bourne, C. R.; Barrow, W. W.; Aoki, S. *Journal of Chemical Information and Modeling* **2014**, *54*, 1242.
9. Lele, A. C.; Raju, A.; Khambete, M. P.; Ray, M. K.; Rajan, M. G. R.; Arkile, M. A.; Jadhav, N. J.; Sarkar, D.; Degani, M. S. *ACS Medicinal Chemistry Letters* **2015**, *6*, 1140.

10. Anderson, A. C. *Chemistry & Biology* **2003**, *10*, 787.
11. Cody, V.; Galitsky, N.; Luft, J. R.; Pangborn, W.; Rosowsky, A.; Blakley, R. L. *Biochemistry* **1997**, *36*, 13897.
12. Li, R.; Sirawaraporn, R.; Chitnumsub, P.; Sirawaraporn, W.; Wooden, J.; Athappilly, F.; Turley, S.; Hol, W. G. J. *Journal of Molecular Biology* **2000**, *295*, 307.
13. Dias, Marcio Vinicius B.; Tyrakis, P.; Domingues, Romenia R.; Leme, Adriana Franco P.; Blundell, Tom L. *Structure* **2014**, *22*, 94.
14. Hajian, B.; Scocchera, E.; Keshipeddy, S.; G-Dayanandan, N.; Shoen, C.; Krucinska, J.; Reeve, S.; Cynamon, M.; Anderson, A. C.; Wright, D. L. *PLOS One* **2016**, *11*, e0161740.
15. Katsila, T.; Spyroulias, G. A.; Patrinos, G. P.; Matsoukas, M.-T. *Computational and Structural Biotechnology Journal* **2016**, *14*, 177.
16. Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. *Journal of Molecular Biology* **1997**, *267*, 727.
17. Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; Mainz, D. T. *Journal of Medicinal Chemistry* **2006**, *49*, 6177.
18. Schrödinger Release 2017-4: LigPrep, Schrödinger, LLC, New York, NY, 2017.
19. Pawlak, J. M.; Khau, V. V.; Hutchison, D. R.; Martinelli, M. J. *The Journal of Organic Chemistry* **1996**, *61*, 9055.
20. Ollinger, J.; Bailey, M. A.; Moraski, G. C.; Casey, A.; Florio, S.; Alling, T.; Miller, M. J.; Parish, T. *PLOS One* **2013**, *8*, e60531.
21. White, E. L.; Ross, L. J.; Cunningham, A.; Escuyer, V. *FEMS Microbiology Letters* **2004**, *232*, 101.
22. Nixon, M. R.; Saionz, K. W.; Koo, M. S.; Szymonifka, M. J.; Jung, H.; Roberts, J. P.; Nandakumar, M.; Kumar, A.; Liao, R.; Rustad, T.; Sacchettini, J. C.; Rhee, K. Y.; Freundlich, J. S.; Sherman, D. R. *Chemical Biology* **2014**, *21*, 819.
23. Schrödinger Release 2017-4: Desmond Molecular Dynamics System, D. E. Shaw Research, New York, NY, 2017. Maestro-Desmond Interoperability Tools, Schrödinger, New York, NY, **2017**.
24. Schrödinger Release 2017-4: QikProp, Schrödinger, LLC, New York, NY, 2017.
25. Gond, D. S.; Meshram, R. J.; Jadhav, S. G.; Wadhwa, G.; Gacche, R. N. *Drug Invention Today* **2013**, *5*, 182.

26. Hong, W.; Wang, Y.; Chang, Z.; Yang, Y.; Pu, J.; Sun, T.; Kaur, S.; Sacchettini, J. C.; Jung, H.; Lin Wong, W.; Fah Yap, L.; Fong Ngeow, Y.; Paterson, I. C.; Wang, H. *Scientific Reports* **2015**, *5*, 15328.
27. Lionta, E.; Spyrou, G.; Vassilatis, D. K.; Cournia, Z. *Current Topics in Medicinal Chemistry* **2014**, *14*, 1923.
28. Littell, R.; Morton, G. O.; Allen, G. R. *Journal of the American Chemical Society* **1970**, *92*, 3740.
29. Zelmer, A.; Carroll, P.; Andreu, N.; Hagens, K.; Mahlo, J.; Redinger, N.; Robertson, B. D.; Wiles, S.; Ward, T. H.; Parish, T.; Ripoll, J.; Bancroft, G. J.; Schaible, U. E. *The Journal of Antimicrobial Chemotherapy* **2012**, *67*, 1948.

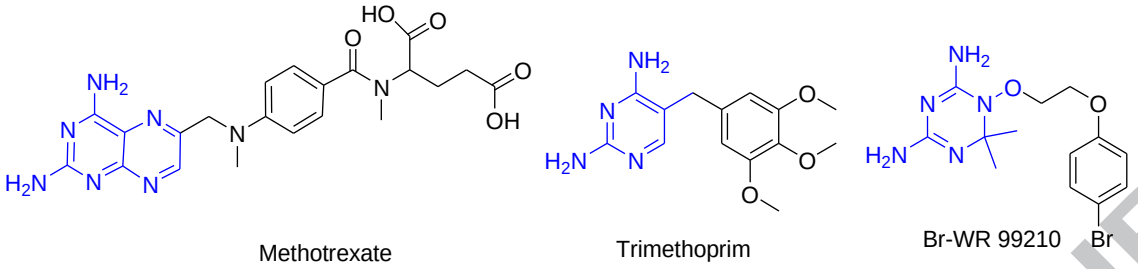


Figure 1: Structure of some reported DHFR inhibitors

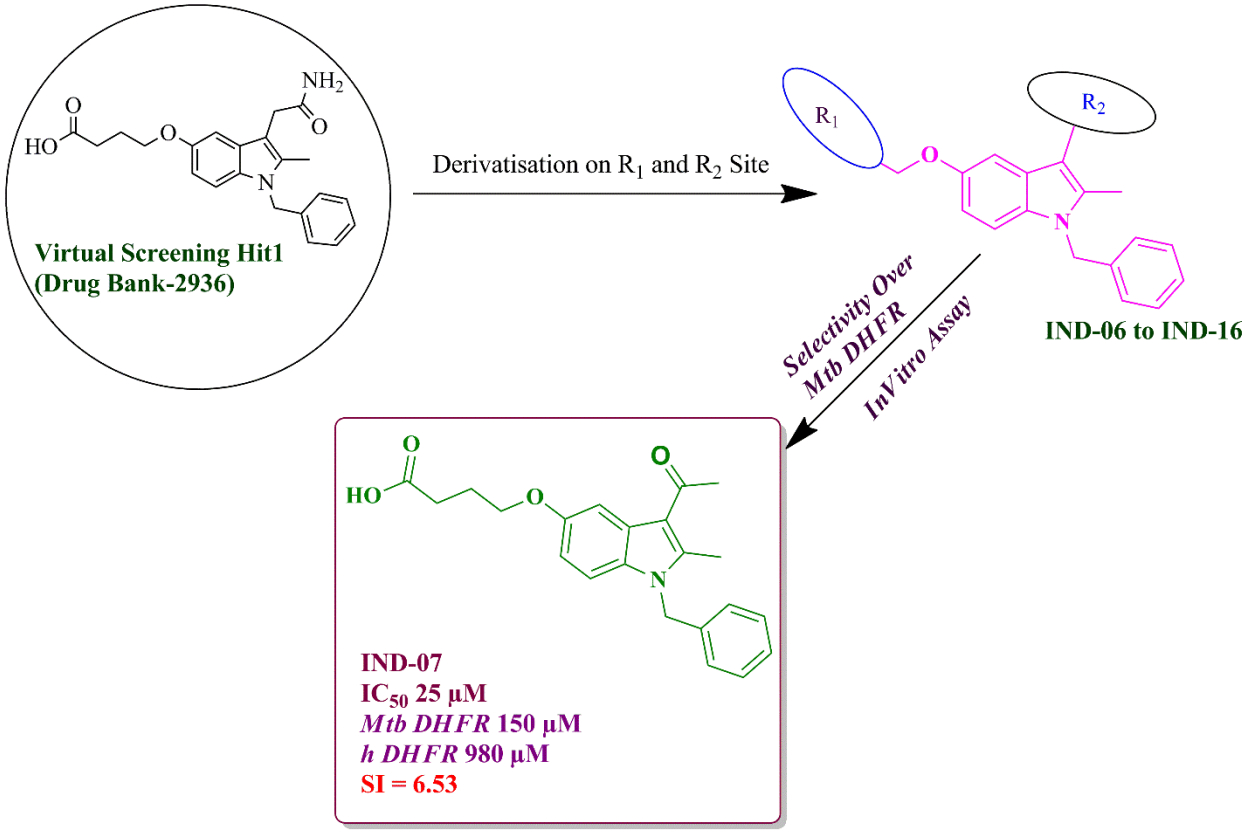


Figure 2: Identification of potential Hit through virtual screening protocol and modification of the same at three different sites to obtain the best compound **IND-07**

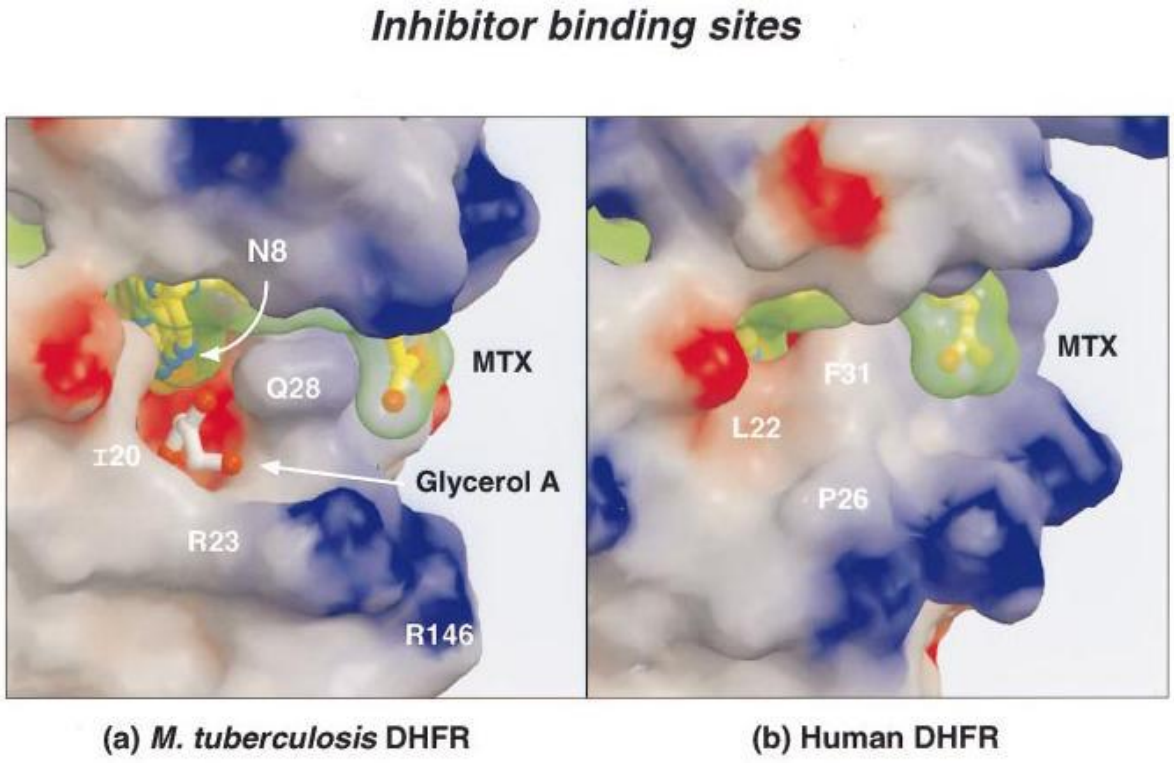


Figure 3: Comparison of inhibitor binding site in Mtb-DHFRand h-DHFRat catalytic domain. In Mtb DHFRinhibitor binding site a glycerol molecule bound close to the inhibitor MTX is visible. (Li. R. *et al*[12])

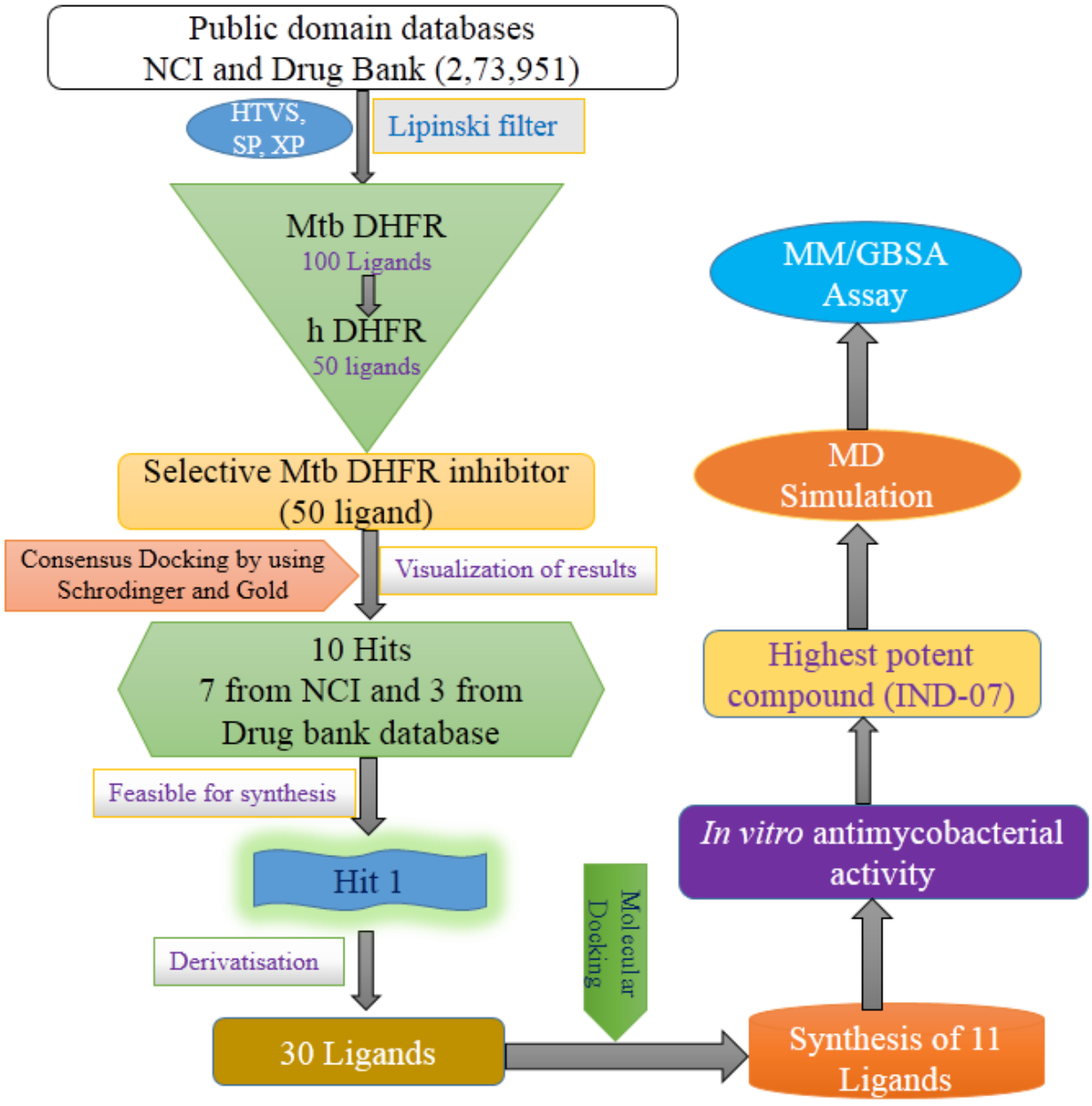


Figure 4. Virtual Screening Workflow

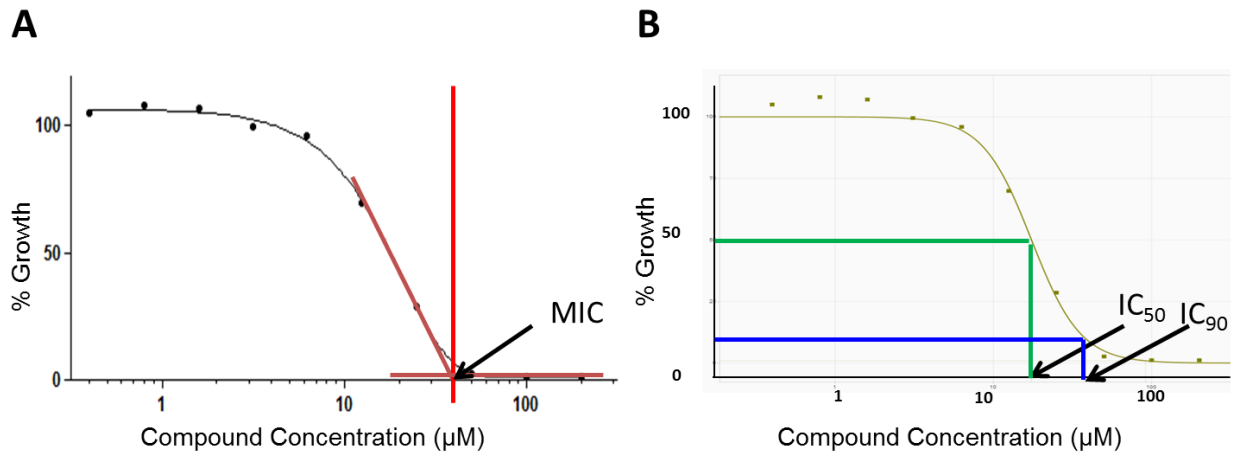


Figure 5: Dose response curve used to calculate *MIC*, IC_{50}, and IC_{90}. Data points obtained from a dose response growth inhibition assay are curve fitted using (A) the Gompertz model to calculate *MIC* and (B) the Levenberg-Marquardt algorithm to calculate IC_{50} and IC_{90}.

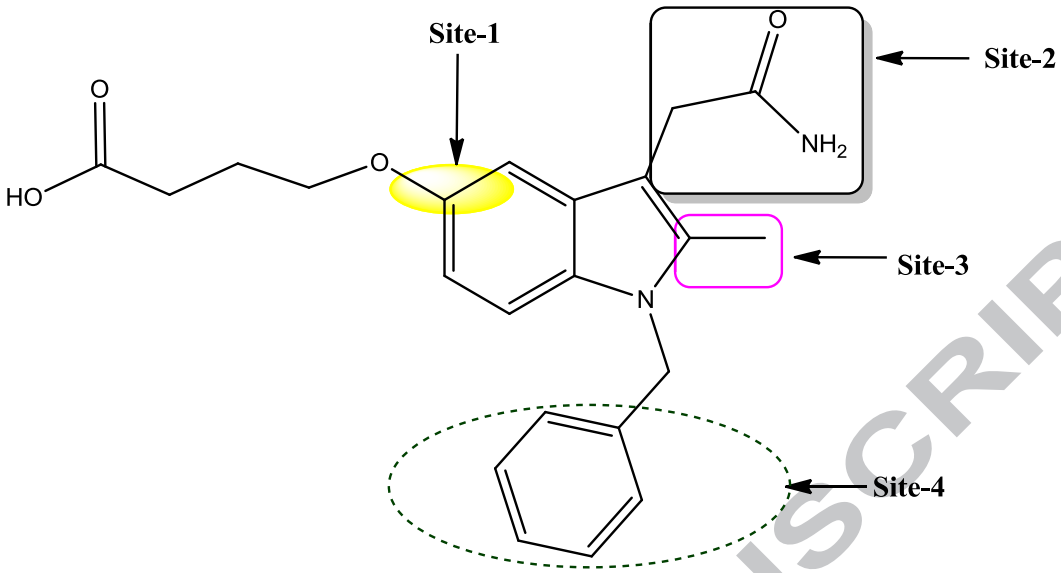
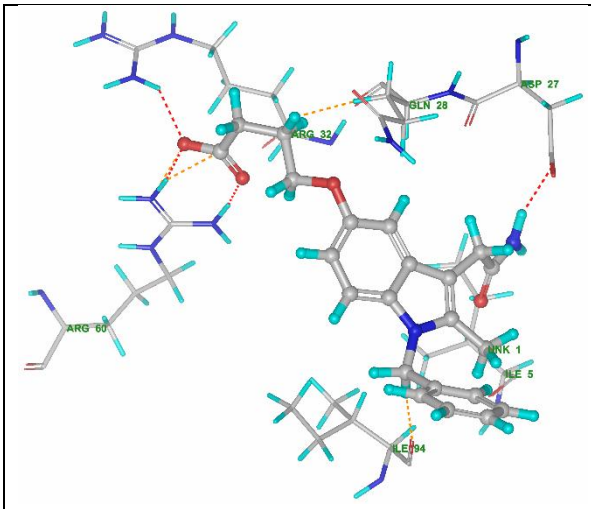
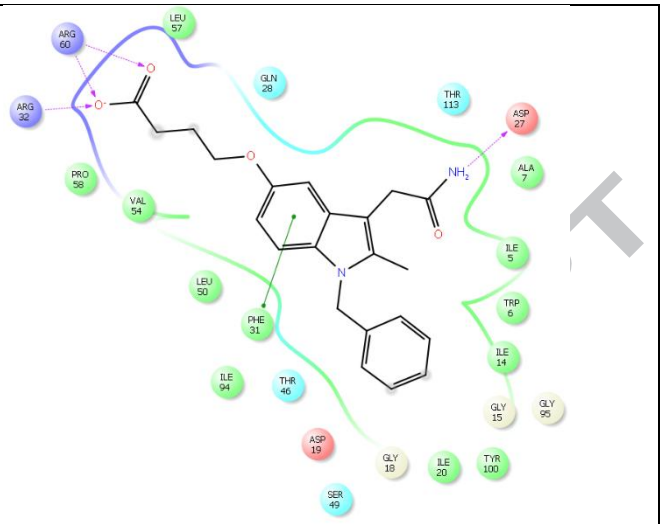


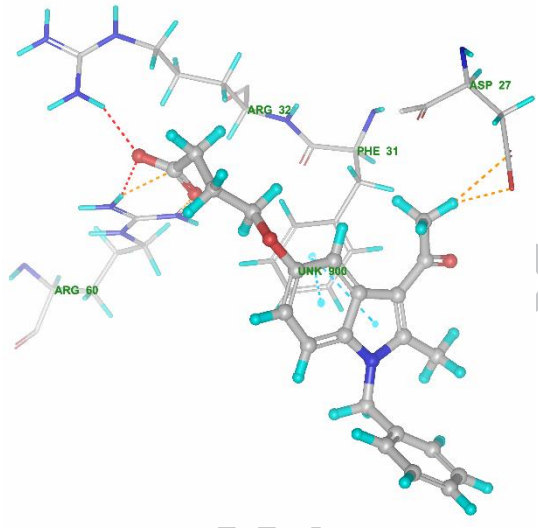
***Figure 6*:** Basic indole moiety and sites of modification



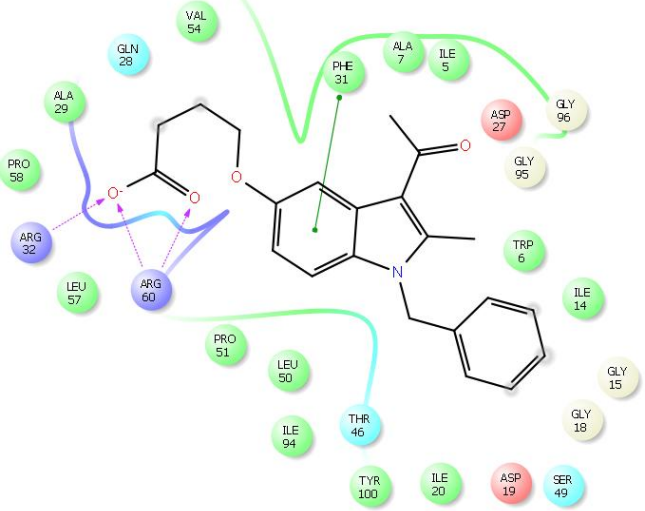
(A)



(B)



(C)



(D)

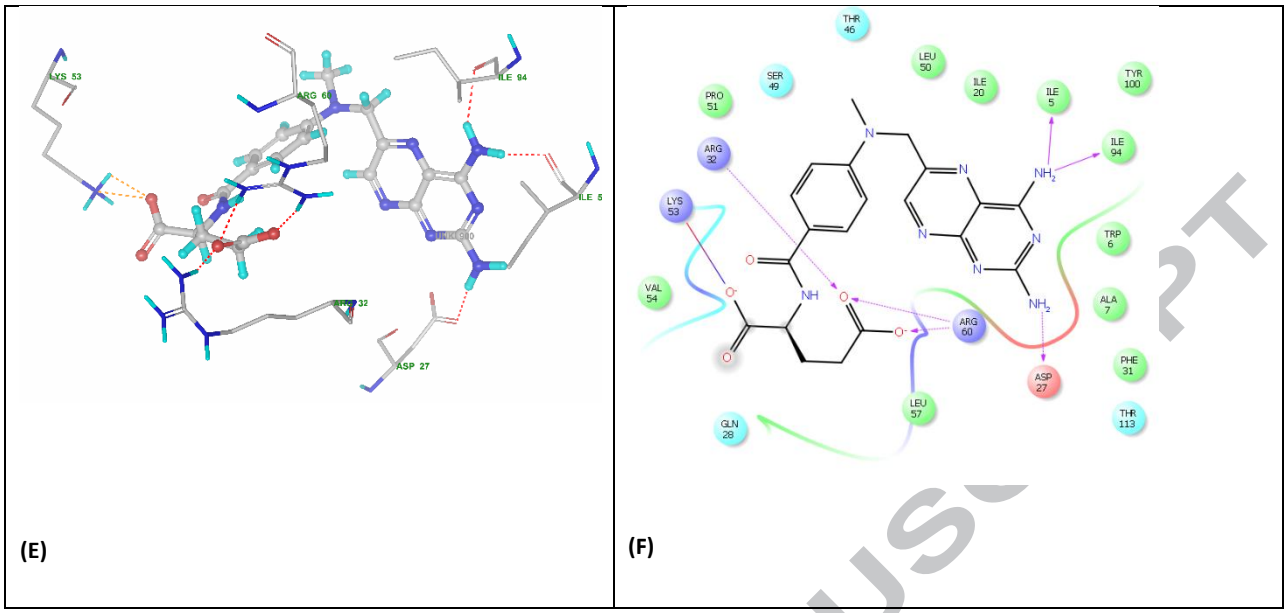


***Figure 7*:** (A) Docking pose of hit1 in Mtb-DHFR active site, (B) Ligplot of hit1 against Mtb-DHFR catalytic domain (C) Binding interaction of **IND-07** in Mtb-DHFR, (D) Ligplot of **IND-07** against Mtb-DHFR, (E) Docking pose of MTX against Mtb-DHFR, (F) Ligplot of MTX against Mtb-DHFR.

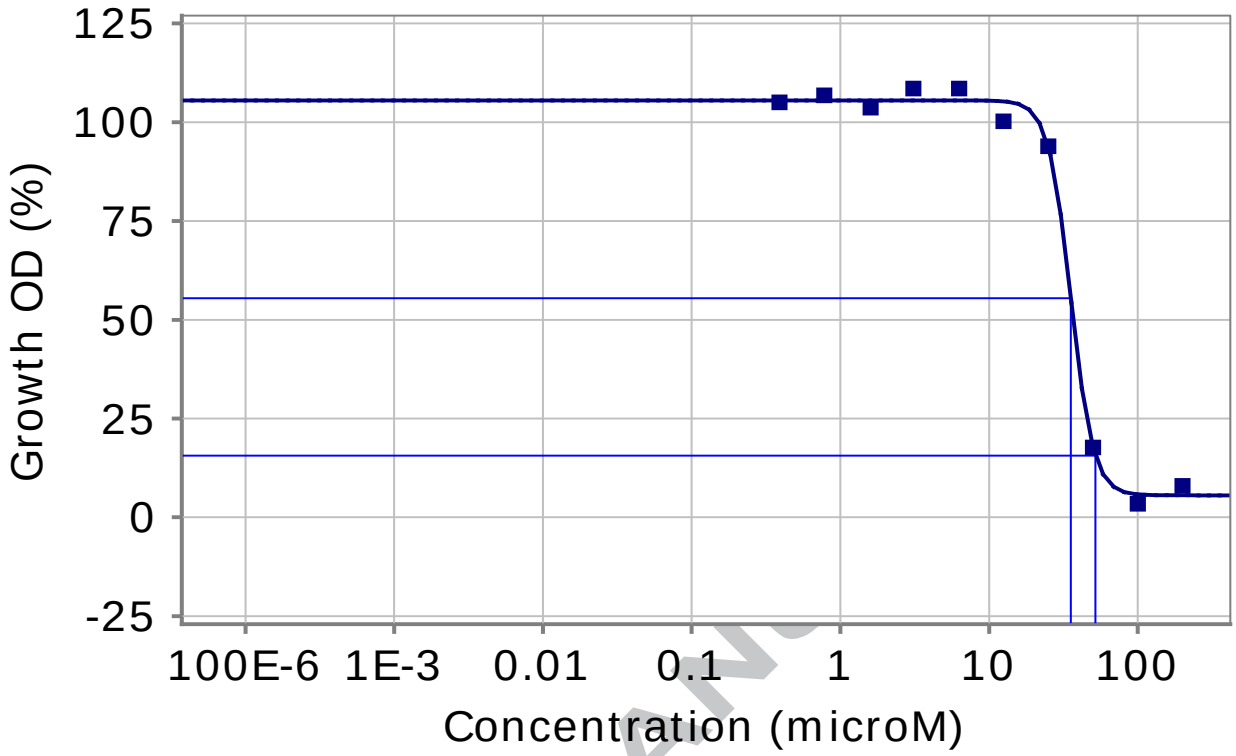
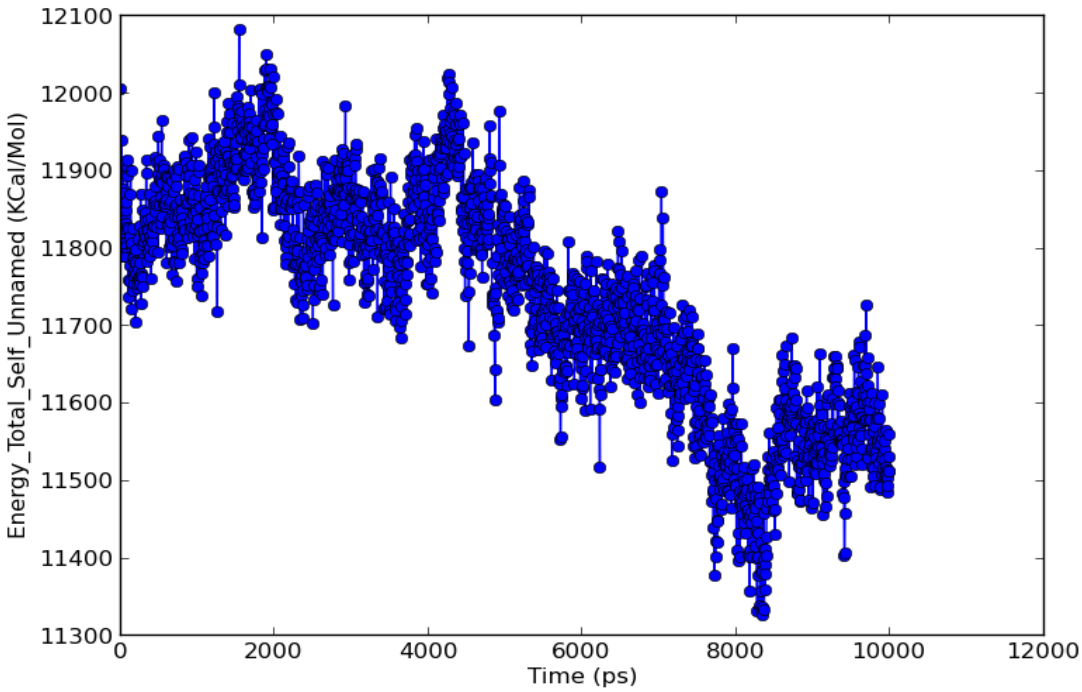
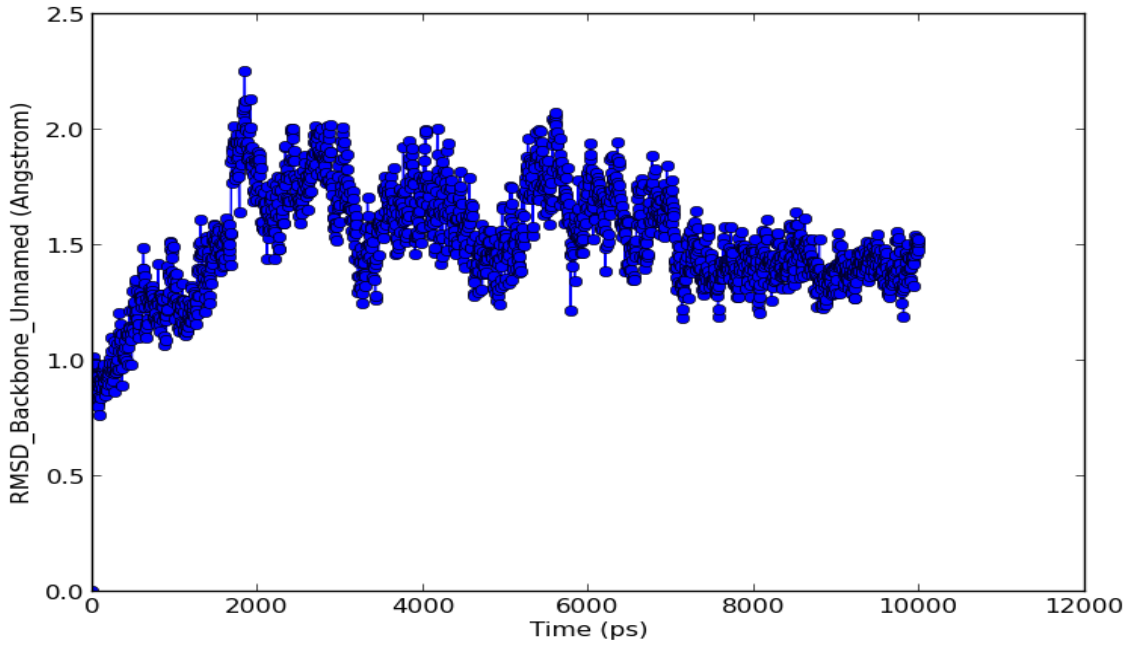


Figure 8: Dose response curve for **IND-07**



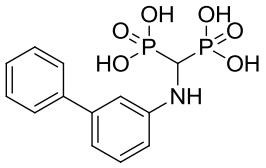
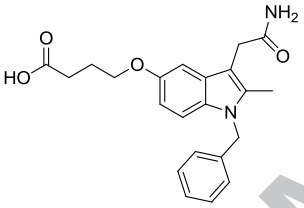
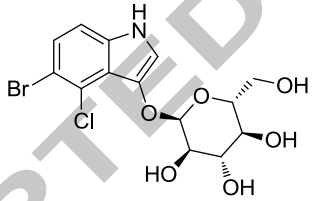
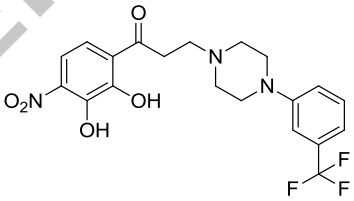
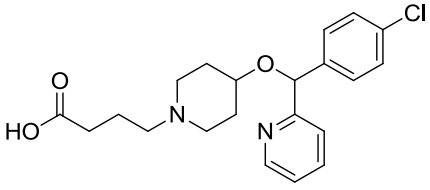
(A)



(B)

***Figure 9*:(A)**Time dependence of the total energy and protein backbone, (**B**) RMSD relative to the initial minimized complex of molecule during MD simulations.

Table 1: Structure of final hits selected after virtual screening of NCI and drug bank followed by consensus docking using Gold suite

S.N	Code	Molecules	Glide scores against Mtb-DHFR	Interactions against Mtb-DHFR	Glide scores against h-DHFR	Interactions against h-DHFR
1.	NSC725581		-11.04	Ile20, Gol502, Ser49, Phe31	-8.22	Glu30, Ile7
2	**DB2936 (Hit 1)**		**-11.03**	**H-bonds-ASP27, Arg60, Arg32, Ser49, π-stacking-Phe31**	**-6.67**	**H-bonds-Gln35, Arg70, Asp21**
3	DB448		-10.92	H-bonds-Gln28, Ser49, Ile94, π-stacking-Phe31	-7.75	H-bonds-Val115, Glu30, NADPH
4	DB328		-10.91	H-bonds-Arg60, Arg32 π-stacking-Phe31	-6.76	H-bonds-Ser59, Asn64
5	DB185		-10.72	H-bonds- -Arg60, Arg32	-4.21	H-bonds- , Arg70, Gln35

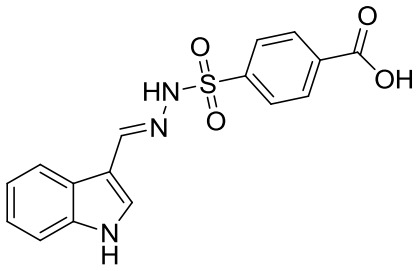
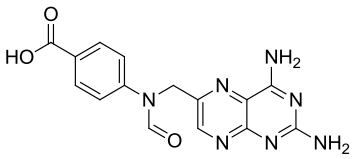
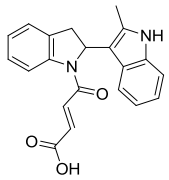
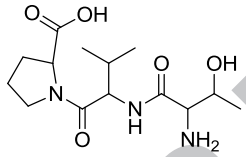
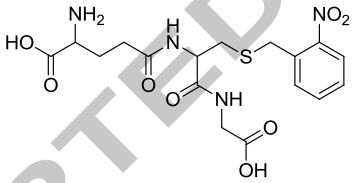
6	NSC133787		-10.44	Ile20, Ile94, Ser49, Phe31	-6.69	Glu30
7	NSC694485		-10.00	H-bonds-ASP27, Ser49, π-stacking-Phe31, Ile5, Ile20, Ile95	-8.59	Val115, Ile5, Glu30, Phe30
8	NSC631310		-9.42	Ile94, Ile20, Phe31, GOL502	-7.26	Asp21, Glu30
9	NSC85313		-9.41	Asp27, Ser49, GOL502,	-7.23	Asp21, Glu30
10	NSC131120		-9.68	Asp 27, Arg 32, Arg 60, Gln28, GOL502, Ile20, Asp19	-8.50	Asp21, Arg70, Ser59, Phe34, Gly35

Table 2: Important interaction of methotrexate, COP and hit1 identified with the active site of Mtb-DHFR and h-DHFR

Type of interaction	Methotrexate interactions			COP binding interactions			Hit1	
	Expt.	*In silico*		Expt.	*In silico*			
	Mtb-DHFR	**Mtb-DHFR**	**h-DHFR**	**h-DHFR**	**Mtb-DHFR**	**h-DHFR**	**Mtb-DHFR**	**h-DHFR**
H-Bonds	**Asp27**, Ile5, Ile-94, Asn60, Arg32, Gln28	**Asp27**, Ile5, Ile94, Arg60, Arg32, Gln28	**Val115**, Ile7, Glu30, Gln35, Arg70	**Val115**, Ile7, Glu30, Gln35, Arg70, Arg32, Asn64,	**Asp27,** Ile5, Ile94, Arg60, Arg32, Gln28, Arg27	**Val115**, Ile7, Glu30, Gln35, Arg70, Arg32, Asn64	**Asp27,** Arg60, Arg32, Ser49	Gln35, Arg70, Ser59
π-stacking	NA	Phe31	Phe34	NA	Phe31	Phe34	Phe31	No π-stacking
Glide score	NA	-12.37	-12.99	NA	-12.21	-14.04	-11.03	**-6.67**

Table 3: Docking score of synthesized compounds against Mtb-DHFR and h-DHFR, their i*n vitro* antimycobacterial activity and enzymatic inhibition data

Compound id	**Docking Score against**		**IC_{50} (µM)**			**Selectivity ratio**
	Mtb-DHFR	**h-DHFR**	***in vitro* $H_{37}RV$**	**Mtb-DHFR**	**h-DHFR**	
IND-06	-6.404	-7.041	160	nd	nd	nd
IND-07	**-8.426**	**-6.921**	**25**	**150**	**980**	**6.53**
IND-08	-6.831	-7.513	35	nd	nd	nd
IND-09	-6.500	-8.153	63	nd	nd	nd
IND-10	-8.890	-5.216	39	nd	nd	nd
IND-11	-6.076	-5.965	180	nd	nd	nd
IND-12	-6.293	-6.642	200	nd	nd	nd
IND-13	-5.629	-6.047	110	nd	nd	nd
IND-14	-7.126	-7.453	65	nd	nd	nd
IND-15	-6.874	-5.990	90	nd	nd	nd
IND-16	-6.005	-5.879	50	nd	nd	nd
Rifampicine	-	-	0.0057	nd	nd	nd
Methotrexate	-12.37	-12.99	0.58	0.00825	0.00197	0.194

nd= denotes not determined

Table 4: In silico ADMET and binding free-energy calculations for MTX and IND-07 with Mtb DHFR.

S.N.	Compound Code	QP log Po/w[a]	% HOA[b]	Rule of Five[c]	Evdw[d]	Ecoul[e]	ΔG_{bind}
1	**IND-07**	4.53	4.75	3	1.18	-47.08	-84.834
2	**MTX**	-1.80	0.00	2.0	0.193	-8.63	-23.584

Structural Comparison of Mtb-DHFR and H-DHFR for Design, Synthesis and Evaluation of Selective Non-Pteridine Analogues as Antitubercular Agents

Kalicharan Sharma[1], Omprakash Tanwar[1], Shweta Sharma[1], Shakir Ali[2], M. M. Alam[1], M. S. Zaman[1], Mymoona Akhter*[1, 3]

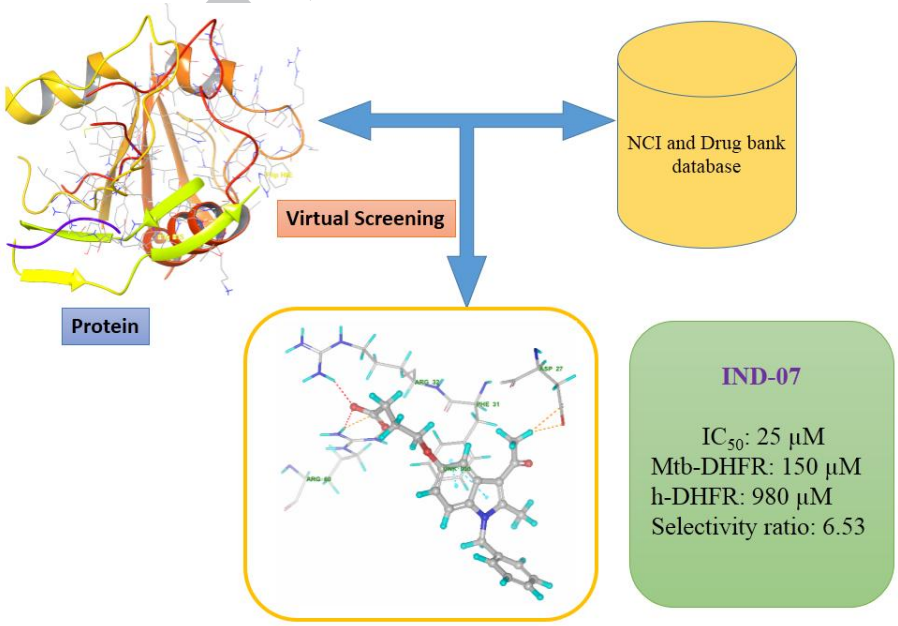
[1]Department of Pharmaceutical Chemistry, School of Pharmaceutical Education and Research, Jamia Hamdard, New Delhi, 110062, India

[2] Department of Biochemistry, School of Chemical and Life Sciences, Jamia hamdard, New Delhi-110062

[3] Bioinformatics Infrastructure Facilities, Jamia Hamdard, New Delhi-110062

Graphical abstract:

A series of indole based compounds showed excellent antitubercular activity against $H_{37}RV$. The compounds were found to have good inhibitory activity (25 - 200) µM against $H_{37}Rv$ and in enzyme assay against Mtb-DHFR and h-DHFR the compound was found selective towards Mtb-DHFR with selectivity index of 6.53. The antitubercular potential of indole based compounds was validated by molecular modelling studies. The synthesised drug-like compounds were also found non-toxic and can be optimized to get more potent antitubercular agents for future use.



Highlights

Structural Comparison of Mtb-DHFR and h-DHFR for Design, Synthesis and Evaluation of Selective Non-Pteridine Analogues as Antitubercular Agents

Kalicharan Sharma[1], Omprakash Tanwar[1], Shweta Sharma[1], Shakir Ali[2], M. M. Alam[1], M. S. Zaman[1], Mymoona Akhter*[1, 3]

[1]Department of Pharmaceutical Chemistry, SPER, Jamia Hamdard, New Delhi, 110062, India

[2] Department of Biochemistry, School of Chemical and Life Sciences, Jamia Hamdard, New Delhi-110062

[3] Bioinformatics Infrastructure Facilities, Jamia Hamdard, New Delhi-110062

- Using Structure Based Virtual Screening novel hit was identifies as selective Mtb-DHFR inhibitor
- Structural medication of hit 1 was carried out and different derivatives were designed
- 11 compounds were synthesized
- The synthesized compounds were evaluated for their antimycobacterial activity and selective Mtb- DHFR inhibition assay.
- Compound IND-07 was found to be most active with IC_{50} of 150 µM and selectivity index of 6.53 against h-DHFR
- IND-07 can be taken further as lead to developed novel selective Mtb-DHFR inhibitor as antimycobacterial agent