A computational chemistry-driven hypothesis on the mode of action of Hipposudoric Acid and related analogues

Islam M. Araar* and Stephen P. Wren
School of Life Sciences, Pharmacy and Chemistry, Faculty of Science, Engineering and Computing, Kingston University, Penrhyn Road, Kingston upon Thames, KT1 2EE, U.K.

*Author for correspondence: islamaraar@outlook.com

Abstract

Aim: To elucidate the mode of action of the hipposudoric acid derivatives and identify hit compounds for synthesis. **Materials & Methods**: Structural fragments of known bioactive fluorenes were introduced onto the hipposudoric acid scaffold to yield novel derivatives. The binding motifs of the novel compounds were compared to the pharmacophore of DHFR co-crystallised with MTX. **Results:** Several of the novel compounds showed binding affinities that exceeded the affinity of the docked endogenous ligand (dihydrofolic acid). **Conclusion:** This study indicates compounds **3r$_{12}$, 3r$_{9}$, 1s$_{9}$ 3r$_{10}$** are promising candidate for synthesis and pharmacological evaluation.

Keywords: Hipposudoric acid, norhipposudoric acid, pigments, antimicrobials, structure-based drug design, antifolate, nucleotide biosynthesis

1. Introduction

Novel antibacterial agents must be developed in response to emerging antimicrobial resistance. [1] In 2019 alone, it is estimated that 4·95 million deaths were associated with bacterial antimicrobial resistance & 1.27 million deaths were attributable to bacterial antimicrobial resistance. [2] It has been shown by several studies that inhibition of dihydrofolate reductase (DHFR) & thymidylate synthase (TS) can be promising strategies for the development of novel antimicrobial agents. [3, 4]

DHFR has a crucial role in biological organisms during folate metabolism. It catalyses the reduction of dihydrofolic acid (DHF) to tetrahydrofolic acid (THF) with the cofactor NADPH being responsible for hydride transfer as described by the chemical equation:

$$DHF \xrightarrow{DHFR,\ NADPH,\ H^+} THF$$

THF serves an important role as a substrate for folate-dependent enzymes which are vital in DNA synthesis. Inhibition of DHFR depletes THF, thereby interrupting the synthesis of nucleic acid precursors and cell proliferation, which leads ultimately to cell death. [5, 6] DHFR is conserved among both prokaryotic and eukaryotic species, and therefore it is an attractive target to produce novel antimicrobial, anticancer, antifungal, and antiparasitic agents. [7, 8]

Among the most well-known antifolate medications include Methotrexate (MTX) which is a competitive inhibitor of the target with an affinity several orders of magnitude higher than endogenous dihydrofolic acid. [9-11]

1.1 Pharmacophore based approach for drug design

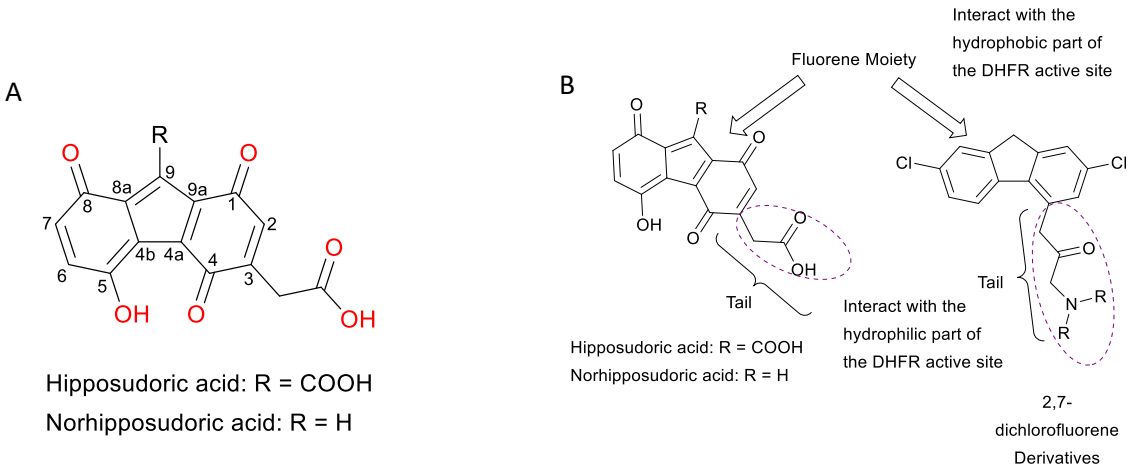
In 2009, Bennett et al obtained the crystal structure of DHFR complexed with MTX and NADH. The key interactions of the substrate and cofactor were analysed in order to use a knowledge-based approach for drug design of molecules that mimic the interactions. [11]
On the basis of understanding the MTX binding interactions, several studies by Hussein designed novel DHFR inhibitors bearing key structural features to mimic the DHFR inhibitor. Hussein et al designed several novel classes including the N^4-substituted sulfonamides and many fluorene-based bioactive compounds. [12-16] Novel 2,7-dichlorofluorenes derivatives were prepared and

tested for biological activity. They exhibited antimicrobial, antifungal, and antitumor activity, and were tested *in-silico* to analyse their binding within the DHFR binding pocket in order to rationalise their biological activity.

1.2 Design of novel Bioactives from Fluorene-based Natural Products

The natural world has historically been an excellent source of medicines and has contributed considerably to the development of modern pharmaceuticals. Nature is crucial in drug discovery as it provides a rich diversity of compounds and starting scaffolds for derivatisation. Yoko Saikawa et al. (Kyoto Pharmaceutical University) reported the isolation and synthesis of hipposudoric and norhipposudoric acids which exhibited *in vitro* antimicrobial activity. [17] As a consequence, these compounds and derivatives show great potential for the development of a novel series of therapeutic agents. However, there is a current gap in the research for derivatives of hipposudoric acids, and to date, there has only been reports of the natural products norhipposudoric and hipposudoric acids. In this study, we used a logic-based approach for drug design to broaden this class antimicrobials. Novel hipposudoric acids are designed with varying substituents chosen to be similar those of the bioactive 2,7-dichlorofluorenes investigated by Hussain et al which showed antimicrobial, anticancer, and antifungal activity

Figure. 1 (A) Structures of hipposudoric and norhipposudoric acids (B) Comparison of hipposudoric acids and 2,7-dichlorofluorenes with Key Structural elements for DHFR inhibition in the enzymatic active site



Chemical Structures and Activity

The hipposudoric and norhipposudoric acids consist of a tricyclic fluorene framework (Figure 1A) containing the key motifs of a 5-hydroxy, a 1,4,8-trioxo unit, and the 3-alkyl substituent bearing a terminal carboxylic acid. The distinction between the two natural products is the substitution on the central five-membered ring, namely the hydrogen atom at the R position (norhipposudoric acid) versus the carboxyl moiety on the hipposudoric acid structure (see Figure 1A).

The hipposudoric acids share great structural resemblance to the bioactive 2,7-dichlorofluorene derivatives developed by Hussain et al., both containing a fluorene core with a polar side tail (see Figure 1B) [13, 14]. In light of the reported activity, we envisioned that designing novel derivatives bearing this structural feature holds promise for the delivery of effective antimicrobials.

1.3 Mechanism of Action

In light of their recent discovery and relative dearth of research literature in this field, the precise mechanisms of action of these fluorene-based antimicrobials are not yet known. The 2,7-dichlorofluorenes derivatives were designed to conform to the pharmacophoric template of known DHFR inhibitors, and we therefore speculated that perhaps the 2,7-dichlorofluorenes & the hipposudoric acids exhibit their antimicrobial activity in a similar manner (through DHFR inhibition).

To elucidate a rational hypothesis on the mode of action for the natural products and derivatives, herein, we used computational chemistry to evaluate the interaction of hipposudoric acids within the binding site of a known antimicrobial pharmacophore. Through analysis of DHFR co-crystallised with MTX (PDB ID: 4DFR) [18], the topography of the binding site was defined, thereby enabling a structural template for the design of novel compounds. We identified novel hipposudoric acids bearing the key structural moieties to conform to the defined pharmacophore.

2. Experimental

The principal aim of this study was to elucidate the mode of action of the hipposudoric acid derivatives and to identify hit compounds for synthesis. This was performed *via* molecular docking simulation of the compounds and analysis of the various interactions within the active site of DHFR.

2.1. Materials

All computational chemistry studies were performed on a Dell XPS 15 9530 with an IntelI CITM) i7-4702HQ CPU @ 2.20 GHz, 2201 Mhz, 4 Core(s), 8 Logical Processor, with 16.0 GB of Physical Memory (RAM), x64-based PC system type, and Microsoft Windows 10 Pro operating system (Version 10.0.18363 Build 18363). All docking and molecular modelling calculations were performed using the computational software Molecular Operating Environment (MOE® 2015.10; Chemical Computing Group (CCG) Inc., Montreal, Canada. http://www.chemcomp.com. All software was written in python (2021), [19] all 2-Dimensional chemical structures were drawn with ChemDraw (2020, Version 20.0.0.41) all chemical data conversion was performed with OpenBabel (2006, Version 2.3.1).

2.2. General Methodology

2.2.a. Protein Preparation

The X-ray crystal structure of DHFR co-crystallised with MTX (PDB ID: 4DFR) [18] was obtained from the Protein Data Bank [https://www.rcsb.org] with the resolution of 1.70 Å. The enzyme structure was checked for missing atoms, bonds and contacts, and all water molecules and ions were removed. Hydrogen atoms were added to the enzyme structure and water molecules and ions were manually deleted. The X-ray crystal structure contained a dimer, the A polypeptide chain and ligand was chosen, while the remaining B polypeptide chain and ligand were deleted. The partial charges on the enzyme were calculated.

2.2.b. Novel Ligand Generation

The fragment library was manually drawn with ChemDraw® (2020) and saved as a single file in cdx format. The python [19] software code executed OpenBabel [20] *via* a command line to convert the fragment library from the cdx format to the SMILES format which allowed manipulation of the chemical data. The core hipposudoric acid scaffolds 1, 2, and 3 were used as the template. Through successive iterations, the fragments were systematically added to the framework one-by-one to generate the corresponding novel compounds. The resulting novel compounds were categorised into 3 libraries based on the core scaffold, each library containing 64 compounds (total: 192 compounds). The novel compounds were converted using OpenBabel[20] from the smile notation to the PDB format and 3D coordinates were generated. Molecular Operating Environment (MOE) [21] was the software package used for docking and evaluating ligand-protein interactions – the chemical data was imported into three separate MOE databases [21] (based on chemical scaffold). Energy minimisation and partial charge calculations were performed for all compounds utilising MOE with the MMFF94X molecular mechanics force field. [22]

2.2.c. Docking in the Active Site of DHFR

The topography of the binding site of DHFR (PDB ID: 4DFR) was determined by the atoms of the co-crystallised ligand (Methotrexate), which was used for docking of all compounds. The docking protocol validation used the Triangle matcher placement method [23] which resulted in a RMSD value of 0.2888 which provided a confidence in the reliability of the docking protocol. The databases of ligands (total: 192 compounds) were subjected to docking simulations using MOE-Dock. [21] The docking method used for the placement stage was the Triangle matcher placement method with the London dG scoring function, the refinement stage used the Rigid receptor refinement method. [23] This used the GBvi/WSA dG force-field based scoring function which estimates the free energy of binding of the ligand from a given pose. [23] The resulting output data was combined into a single file by selecting the highest scoring pose for each compound.

2.2.d. Molecular Dynamics Simulations

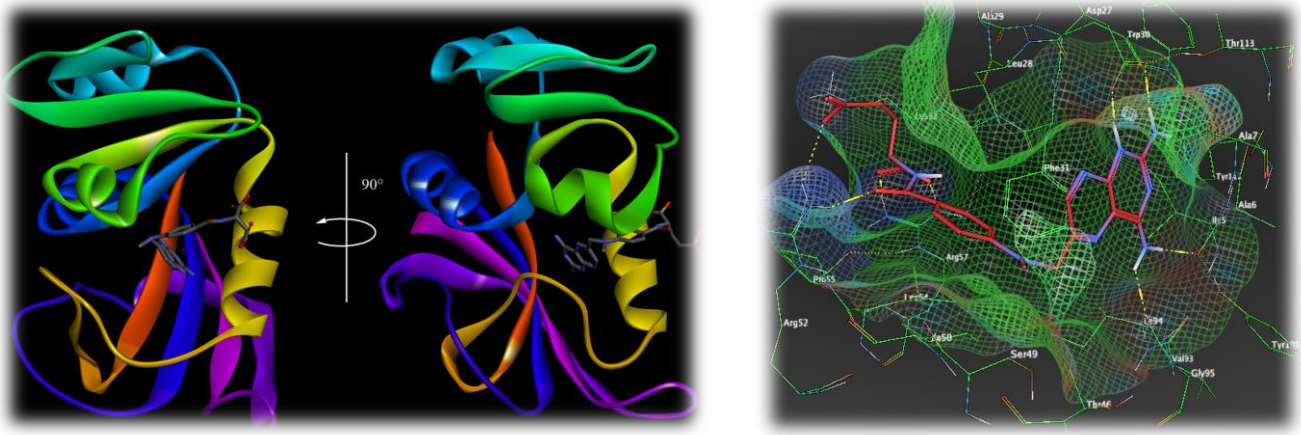
Molecular dynamics simulations were carried out using MOE [21] for the highest scoring compound **3r**$_{12}$ (see supplementary information). The protein–ligand complex had been prepared from prior steps and was checked to ensure there were no missing atoms, bonds or hydrogen atoms. Using the droplet mode, the system was solvated in water with a sphere shape, partial charges were calculated and the energy of the system was minimized. The atoms of the ligand were selected in MOE and assigned to a named set utilising the SVL command line with the function 'oSetCollection ['MD INTERACTION', SelectedAtoms []]'[24] The simulations were performed using the MMFF94X molecular mechanics force field, [22] with an algorithm utilising the Nosé-Poincaré-Andersen equations of motion to generate true ensemble trajectories. [25] The resulting output data was used to generate the time-dependent ligand-target interactions (see supplementary information).

3. Results & Discussion

3.1 Native Ligand-Protein complex

In this investigation, we determined the pharmacophore of a known DHFR inhibitor by analysis of the protein-ligand crystal structure obtained from the protein data bank. In order to obtain a reliable model for docking simulation, the drug target X-ray structure must have a resolution of at least 2.4 Å simulation with resolution of less than 2 Å being ideal. [11, 26] The DHFR protein co-crystallised with methotrexate as the ligand was chosen and this was downloaded from the protein data bank (PDB ID: 4DFR) [18] with a resolution of 1.70 Å. MOE is a software package used for docking and evaluating ligand-protein interactions. [21] It has numerous tools capable to analyse ligand docking interactions and to elucidate a pharmacophore, and it allows for rapid serial docking of a library of compounds which we deemed to be advantageous for our study. Moreover, it is the software used in the study by Hussein et al., [12-16] therefore we chose to use the same software and docking parameters.

Figure 2a : (A) DHFR X-Ray Structure visualised in Discovery Studios 2020 [27] (B) DHFR binding site visualised in MOE. [21] The H-bonds are represented with yellow lines with a cylinder. The ionic interactions are represented with yellow dotted lines. The H-pi bonds are represented by the violet dotted lines.



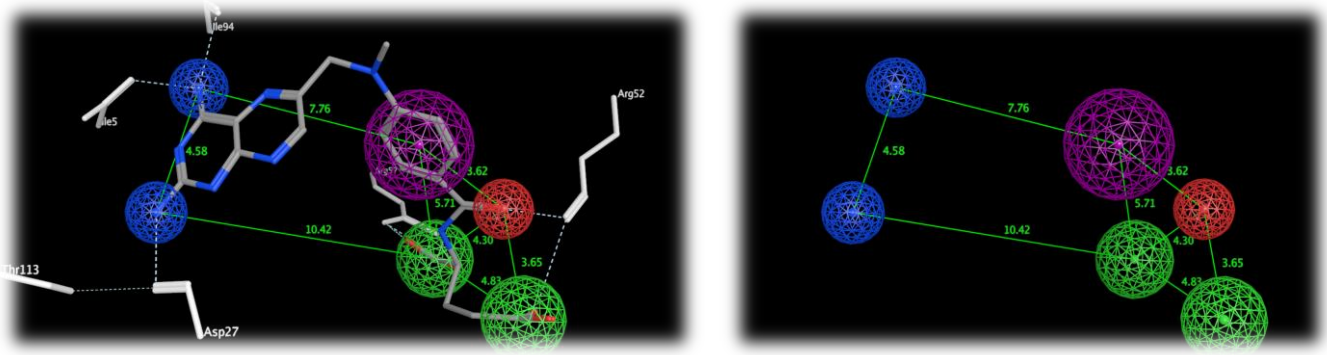


Figure 2b : (A) Generation of pharmacophore model from X-ray structure (B) 3D spatial arrangement of features & Inter-feature distances. Spheres Represent the chemical features: Blue is Hydrogen Bond Donor (HBD), Red is Hydrogen Bond Acceptor (HBA), Purple is Ring Aromatic (RA), Green is Negative Ionisable (NI)

3.2 Pharmacophore discovery

The spatial arrangement and key interactions of the ligand-protein complex (PDB ID: 4DFR) was used to generate a structure-based pharmacophore (See Figure 2b) and this was, in turn, used as a template for the subsequent study of novel, virtual derivatives. The X-ray structure was analysed using MOE [21] (, and this study showed that the ligand interacts with the active site of DHFR *via* several hydrophobic and hydrogen bonding interactions.

Within the polyglutamic acid moiety (see Figure 3A), the carboxyl O1 & O2 made hydrogen bonding interactions with Arg57 (distance: 2.57 Å and 2.72 Å, respectively), while OE2 interacted with Arg52 *via* an ionic interaction (distance: 3.43 Å). Within the p-methylaminobenzoyl moiety, the carbonyl oxygen interacted with Arg52 *via* hydrogen bonding (distance: 2.88 Å). Within the pteridine moiety, there was an arene-H interaction between the ring and Ala7 (distance: 4.24 Å), while N1 & NA2 made hydrogen bonding interactions with Asp27 (distance: 2.48 Å and 2.88 Å, respectively), and NA4 formed hydrogen bonds with Ile94 & Ile5 (distance: 2.84 Å and 2.73 Å, respectively).

3.3 Docking protocol validation

The molecular interaction between the ligand and the target was established through the crystal structure (PDB ID: 4DFR). Therefore, this was utilised to validate the docking protocol. All simulations were conducted *via* MOE [21], which is a valuable tool in drug discovery to assess potential hit compounds by performing ligand-protein docking simulations. MTX was docked within the vicinity of the original binding pocket and the result was evaluated against the original crystal structure (PDB ID: 4DFR).

The resulting binding pose had an affinity of -9.3551 kcal/mol and had high similarity to the original crystal structure with a low value for the small root mean standard deviation (RMSD) of 0.2888, which provides a confidence in the reliability of the docking protocol. This can be seen with the superimposition of the two structures in Figure 3B.

3.4 Analysis of docking of MTX into the DHFR binding pocket

The docking simulation was analysed using MOE[21], which showed similar binding motifs as the co-crystallised ligand structure in 4DFR (see Figure 3A). The ligand interacts with the active site of DHFR *via* several hydrophobic interactions, with exceptional overlapping of the p-methylaminobenzoyl moiety (Figure 3B).

As shown in Figure 3A, there were several hydrogen bonds within the polyglutamic acid moiety; the carboxyl O1 & O2 made hydrogen bonding interactions with Arg57 (distance: 2.83 Å and 2.86 Å, respectively), while OE1 was a hydrogen bond acceptor with Lys32 (distance: 2.95 Å), and OE2 with ARG52 (distance: 2.97 Å). Within the p-methylaminobenzoyl moiety, the carbonyl oxygen interacted with Arg52 via hydrogen bonding (distance: 2.82 Å), and there was an arene-H interaction between the ring and Ile50 (distance: 4.36 Å).

Relative to the co-crystallised ligand (MTX), the pteridine moiety exhibited exceptional overlapping (Figure 3B), with an arene-H interaction between the ring and Ala7 (distance: 4.22 Å), while N1 & NA2 made hydrogen bonding interactions with Asp27 (distance: 2.65 Å and 3.04 Å, respectively), and NA4 formed hydrogen bonds with Ile94 & Ile5 (distance: 2.75 Å and 2.96 Å, respectively) shown in Figure 3A.

.

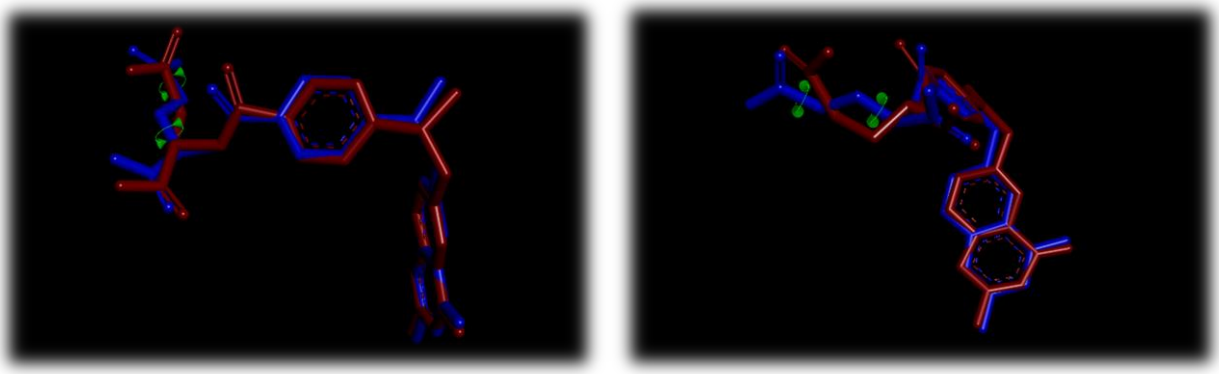


Figure 3a : Comparisons between the conformation of MTX in the original X-ray structure (red), and MOE's best prediction (blue). The green arrow indicates the rotational differences between the two structures. The figures display two different views, rotated to the right by 90°.

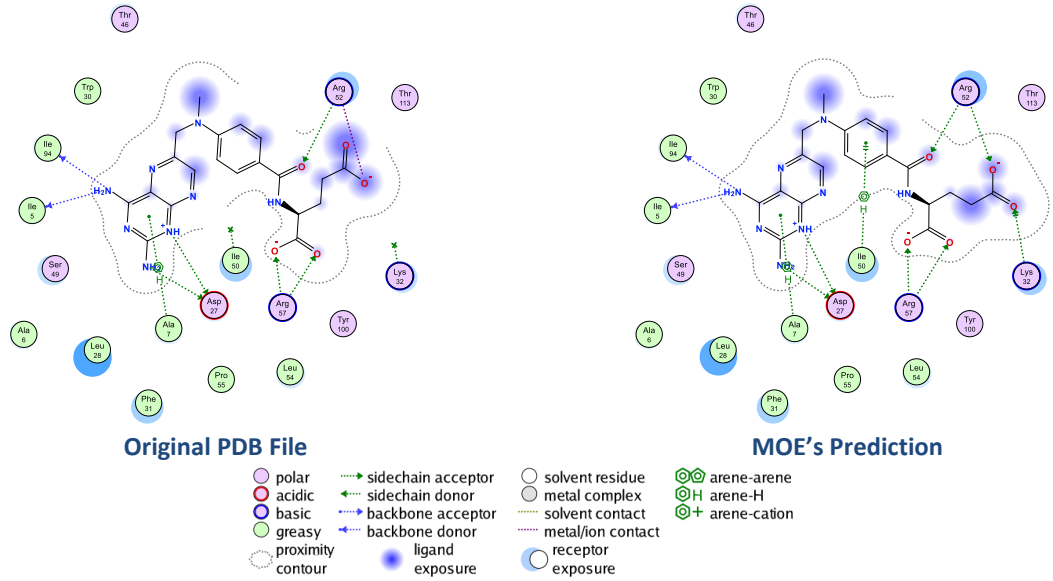


Figure 3b: Ligand-protein interactions diagram of MTX (MOE): Comparison between the original PDB file, and MOE's prediction.

3.5 Generation of Hipposudoric acid analogues

A substituent database was used to systematically generate the virtual compounds by modifying the group at R position of the scaffolds (see Figure 4a). The substituents were chosen to be similar those of the bioactive 2,7-dichlorofluorenes investigated by Hussain et al which showed activity against human carcinoma (lung carcinoma: A-549 and breast carcinoma: MCF-7), fungal species (*A. fumigatus* and *C. albicans*), and bacterial species (Gram +ve bacteria*: S. aureus* and *B. subtilis*; Gram -ve: *E. coli*, and *P. vulgaris*). [13, 14]

The library of substituents was drawn using ChemDraw® (Version 2020) and saved as a single CDX format file. We used the programming language Python [19] to write a software to automate the generation of the compound library (Figure 4b). The script converted the subsistent library from CDX format into the SMILES format; the SMILES notation was used to easily manipulate the data to generate novel compounds.

The core hipposudoric acid scaffolds had the substituent systematically added in order to generate the library of compounds (Figure 4b) in the SMILES notation. Utilising OpenBabel, [20] 3D coordinates were generated by conversion of the compounds into individual PDB files. The compounds were imported into an MOE database (MDB format), the energies were minimised, and the charges were calculated using MOE. A total of 192 novel hipposudoric acid derivatives were generated (Figure 4b) and subjected to docking simulations within the known DHFR binding site (PDB ID: 4DFR).

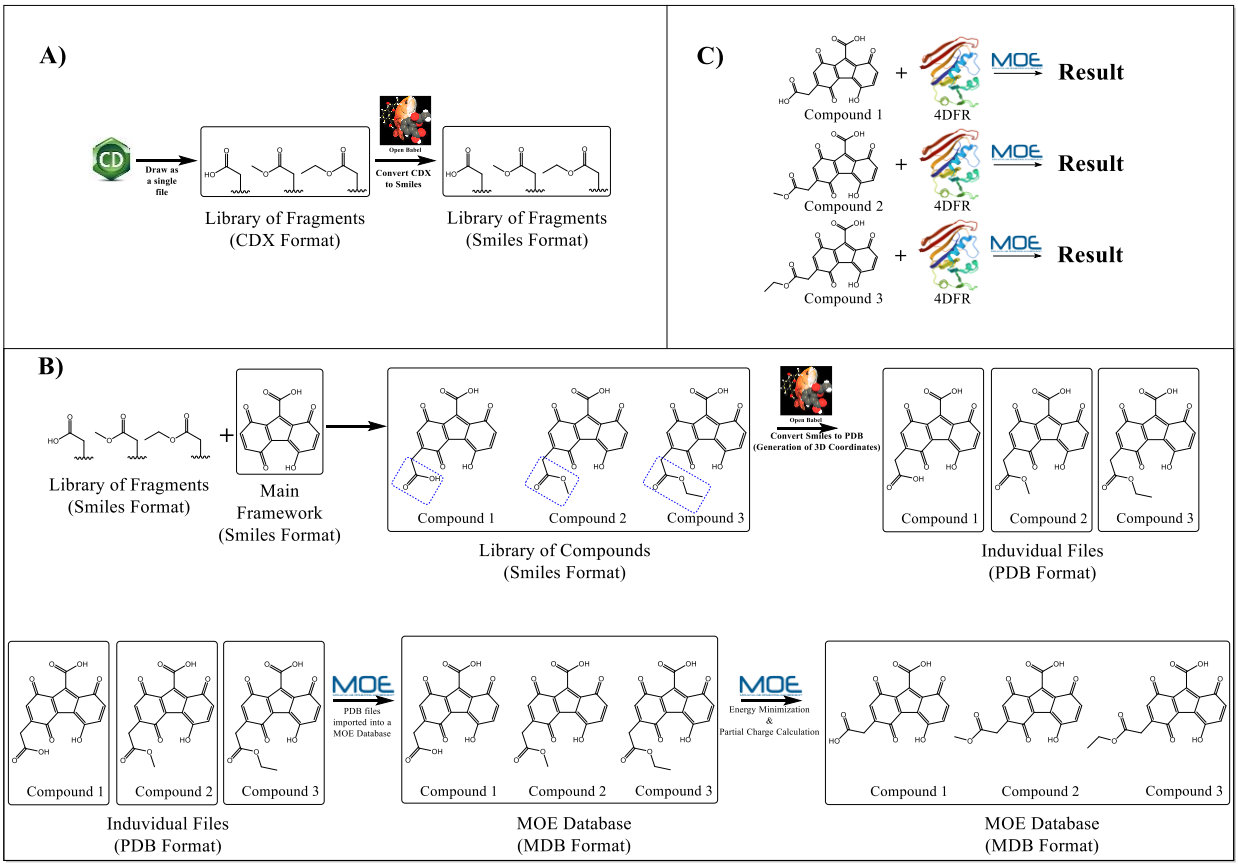


Figure 4a: General method used to create novel Hipposudoric acids. (A) Generation of fragment library and conversion of CDX to SMILES notation (B) Software automation to construct library of novel compounds and ligand preparation for docking (C) Docking of Library of compounds in 4DFR binding site using MOE.

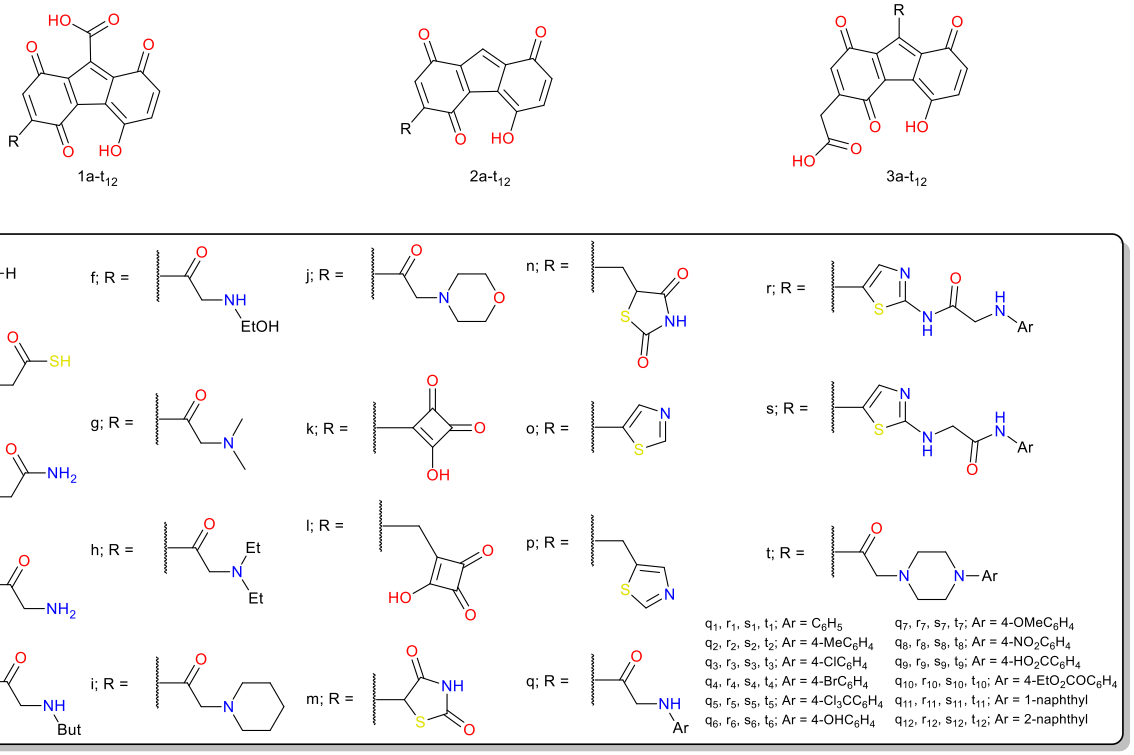


Figure 4b: Library of compounds subjected to docking simulation

3.6. Molecular docking of Hipposudoric acids in the active site of DHFR

Utilising MOE, the novel hipposudoric acids were subjected to docking simulations in order to explore the potential binding mode. The simulations were performed three times for each compound to ensure the reliability of the results. The binding site of the co-crystallised ligand was used to dock the novel derivatives. All docking simulations were conducted using the Merck molecular force field 94x (MMFF94x). [28] During the placement stage, the Triangle placement method was used to generate poses by superposing triplets of ligand atoms and triplets of receptor site points. A total of thirty poses were generated and for each pose the ligand binding free energy was estimated using the London dG scoring function. [23] In the refinement phase, five poses were generated using a rigid receptor method and scored according to the force field-based scoring function GBVI/WSA dG, which represents the predicted ligand-protein binding free energy for each spatial poses. [23]

The binding free energy was calculated by MOE using the following equation.

$$\Delta G \approx c + \alpha \left[\frac{2}{3} (\Delta E_{coul} + \Delta E_{sol}) + \Delta E_{vdw} + \beta \Delta SA_{weighted} \right]$$

The term *G* represents the binding free energy, the term *c* is defined as the gain or loss of translational and rotational entropy, which is calculated during the training stage, in addition to calculation of the forcefield-dependent constants α and β. The term E_{coul} is the energy pertaining to the electrostatic forces of the system, and the term E_{sol} is defined as the energy of electrostatic solvation. E_{vdw} is defined as the energy of the van der Waals potential and $SA_{weighted}$ represents the surface area as a function of exposure. [23]

3.7. Analysis of the binding interactions

We determined the binding scores of 192 compounds (see supplementary information) and the endogenous ligand dihydrofolic acid was docked and used as a reference to compare the novel derivatives (see Table 1 and 2). The results showed that the novel derivatives interacted with the binding site to varying extents, with many showing a high affinity towards the binding site (with binding scores that exceeded that of the endogenous ligand, dihydrofolic acid). Many of the top scoring compounds often showed similar binding motifs as the co-crystallised ligand (MTX). Namely, a carbonyl moiety interacting *via* hydrogen bonding with Arg52, while an adjacent carboxyl moiety often was observed to interact *via* hydrogen bonding with Arg57, and the hydrophobic moiety of the compounds was often observed to be oriented to fit within the central hydrophobic portion of the binding site. The criteria for inclusion for further analysis had been a threshold for compounds whose binding energy was less than 7.3 kcal/mol (Table 1).

Table 1. Docking results of the novel derivatives with a binding energy less than -7.3 kcal/mol

Compound	Score	E-conf	E-place	E-score 1	E-score 2	E-refine
Control	-7.0240	-134.5073	-122.5660	-13.9134	-7.0240	-22.9774
$1q_{10}$	-7.3885	-17.9818	-143.7074	-12.2540	-7.3885	-24.8540
$1r_1$	-7.3737	-47.7705	-96.1423	-10.6129	-7.3737	-27.9134
$1r_2$	-7.3642	-45.3535	-133.3829	-10.6955	-7.3642	-28.3800
$1r_4$	-7.3909	-47.5473	-120.1658	-11.2292	-7.3909	-26.6898
$1r_5$	-7.3908	-25.3870	-124.2776	-11.3704	-7.3908	-25.8514
$1r_7$	-7.4595	-43.5138	-117.4223	-10.8874	-7.4595	-28.0293
$1r_9$	-7.5625	-116.8867	-82.3729	-13.0620	-7.5625	-28.6826
$1r_{10}$	-7.4198	-37.1094	-137.5452	-11.3460	-7.4198	-27.2029
$1r_{12}$	-7.3370	-31.2836	-117.8592	-11.6769	-7.3370	-23.5309
$1s_5$	-7.5117	-24.9577	-128.5559	-10.8948	-7.5117	-27.3029
$1s_7$	-7.3283	-55.4630	-87.6263	-11.1527	-7.3283	-23.9863
$1s_9$	-7.9411	-108.9729	-110.1848	-14.7684	-7.9411	-28.3428
$1s_{10}$	-7.4933	-49.2555	-87.7975	-10.1013	-7.4933	-25.5681
$1s_{11}$	-7.5140	-19.5688	-108.1233	-10.5178	-7.5140	-28.2015
$1s_{12}$	-7.4194	-40.5653	-103.9443	-10.7196	-7.4194	-26.0706
$1t_9$	-7.5780	-61.7238	-102.0943	-12.8393	-7.5780	-23.9289
$2r_9$	-7.5130	-27.2106	-68.5316	-12.8351	-7.5130	-26.6575
$2r_{12}$	-7.3739	50.1143	-96.8856	-9.9964	-7.3739	-30.0885
$2s_3$	-7.3662	28.1790	-94.9236	-10.6071	-7.3662	-24.2617
$2s_6$	-7.5601	7.1974	-88.2140	-12.5934	-7.5601	-28.0972
$2s_{10}$	-7.5420	40.3911	-117.5318	-10.5110	-7.5420	-26.8777

$2s_{11}$	-7.3393	58.7246	-100.5986	-10.2598	-7.3393	-28.0729
$2s_{12}$	-7.4290	42.5740	-72.5539	-10.6578	-7.4290	-27.3171
$3r_2$	-7.3810	-43.3422	-92.7516	-10.6101	-7.3810	-23.5110
$3r_3$	-7.3017	-41.8045	-92.6147	-11.0896	-7.3017	-28.2470
$3r_4$	-7.7087	-35.8706	-100.6932	-11.6789	-7.7087	-29.4758
$3r_6$	-7.9436	-94.9114	-162.1009	-13.3420	-7.9436	-34.8005
$3r_8$	-7.3368	-10.2033	-95.9956	-12.6000	-7.3368	-28.9827
$3r_9$	-7.6526	-32.3202	-111.8695	-12.1529	-7.6526	-25.9100
$3r_{11}$	-7.6197	-36.1043	-108.0134	-11.2174	-7.6197	-29.7067
$3r_{12}$	-7.3238	-47.4253	-86.4283	-11.3619	-7.3238	-28.5288
$3s_2$	-7.6265	-3.4102	-70.3488	-11.6978	-7.6265	-26.8128
$3s_3$	-7.4036	-46.6258	-132.8253	-12.0879	-7.4036	-24.2250
$3s_4$	-7.6751	-32.9135	-81.9703	-11.2243	-7.6751	-27.6446
$3s_5$	-7.3122	-76.7697	-74.6664	-12.1465	-7.3122	-18.5681
$3s_6$	-7.4560	-109.2485	-98.9647	-12.1597	-7.4560	-31.8795
$3s_7$	-7.9295	-26.6449	-76.1882	-11.7596	-7.9295	-23.4576
$3s_9$	-8.0037	-13.8617	-86.1811	-11.9679	-8.0037	-31.9681
$3t_6$	-7.7548	-42.0611	-105.9601	-13.5710	-7.7548	-25.2458
$3t_7$	-7.6893	34.2088	-85.8691	-11.1160	-7.6893	-24.3398
$3t_8$	-7.7325	51.3889	-98.3328	-11.9189	-7.7325	-27.0017
$3t_9$	-7.6645	36.3645	-70.6199	-11.4962	-7.6645	-23.8272

Control: Dihydrofolic acid. Score: lower scores are considered favourable. All values are expressed in kcal/mol. E-conf: Refers to the Energy of the conformer. E-place: Score resulting from placement. E-score 1: Score resulting from the first rescoring stage. E-score 2: Score resulting from the second rescoring stage. E-refine: Score resulting from the refinement stage.

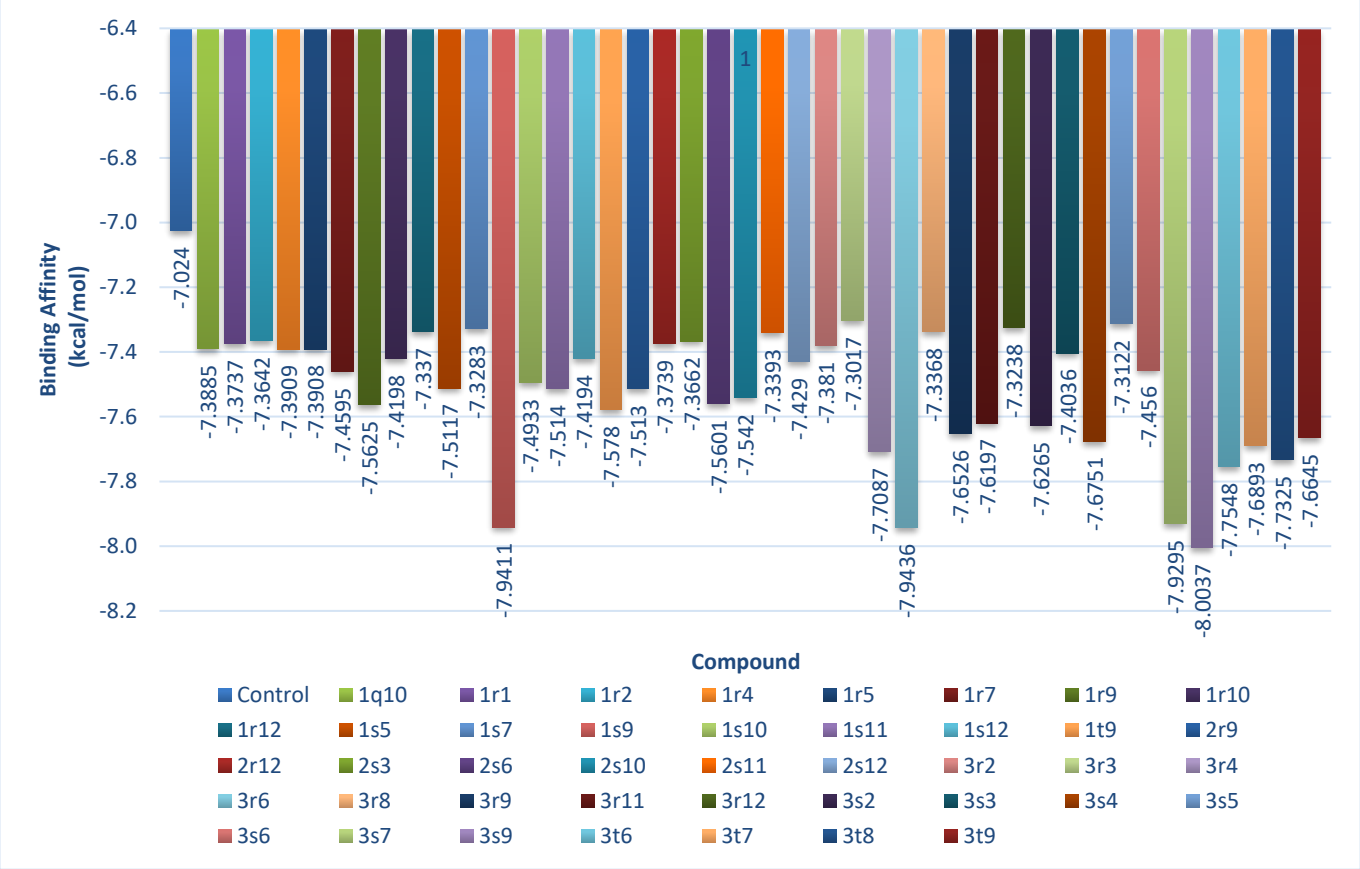


Figure 5. Binding affinity of compounds with a binding energy less than -7.3 kcal/mol. Control: Dihydrofolic acid.

3.7 Analysis of the best binders

Among the compounds with binding affinities less than -7.3 kcal/mol (Figure 5), there was great variety in binding interactions within the DHFR binding site. The top best binders were selected for analysis of the ligand receptor interactions (Table 2). All of these exhibited a higher predicted binding affinity than the docked endogenous ligand (dihydrofolic acid).

In terms of binding interactions, the novel derivatives exhibited several different intermolecular interactions; however, by analysis of the top ten compounds with the highest binding score we identified commonalities between their binding interactions (see table 2). These compounds ranked according to their binding affinity to DHFR are the following: $3r_{12} > 3r_9 > 1s_9 > 3r_{10} > 3t_{10} > 3r_{11} > 3r_4 > 3t_{10} > 3r_7 > 3t_{12}$ (see Table 2). The four best predicted binders ($3r_{12}$, $3r_9$, $1s_9$, and $3r_{10}$) are visualized in Figure 6a.

The results showed compound $3r_{12}$ exhibited the highest score and displayed a higher score than the docked endogenous ligand (dihydrofolic acid). Using SciFinder^n CAS, we have searched for compound $3r_{12}$ and have not found any results. [29] Therefore, we have determined this compound to be a novel structure. Upon analysis of the binding interaction of this novel derivative, it contained several similar binding motifs to the co-crystallised ligand, as demonstrated by the overlay of the two structures (see Figure 7a). The main hipposudoric skeleton was involved in a hydrogen bonding interaction between the O1 and Arg52 (distance: 2.75 Å), the carboxyl side-chain/tail exhibited hydrogen bonding between O2 and Arg57 (distance: 3.06 Å), an arene–hydrogen interaction existed between the hipposudoric acid core and Ile50 (distance: 4.34 Å), and an arene–hydrogen interaction existed between the anthracene moiety and Gly15 (distance: 3.90 Å).

There was a strong hydrogen bond interaction between a carboxyl moiety and Arg57 in many of the compounds who scored highly, suggesting that there may be an advantage to carboxyl interactions with this residue. In contrast to the co-crystallised ligand, many new derivatives displayed an additional hydrogen bonding interaction, namely an interaction with Lys32 which is positioned in close proximity to Arg52 (distance: 4.569 Å from each other).

The highest scoring compound ($3r_{12}$) bears the naphthalen-2-amino acetamido thiazole moiety (see Figure 6b). In the study by Hussain's group, their reported 2,7-dichlorofluorene derivative bearing a naphthalen-2-amino acetamido thiazole moiety exhibited the greatest *in-vitro* cytotoxic activity among all of the investigated compounds against lung carcinoma (A-549) and human breast carcinoma (MCF-7), as well as outstanding activity against *A. fumigatus*. [14] Both the hipposudoric acid ($3r_{12}$) and the related 2,7-dichlorofluorene derivative (see Figure 6b) showed a high dockings score within a similar range (-8.0037 kcal/mol & -7.8543 kcal/mol, respectively). [14, 30] The similarities indicate that compound $3r_{12}$ may be a promising candidate for synthesis and testing for biological activity.

Molecular dynamics simulations were conducted (see supplementary information) on the highest scoring compound ($3r_{12}$) in order to observe the dynamic evolution of the ligand-protein system across time. Through use of a molecular mechanics force field, the simulation predicted the interatomic interactions between the ligand, protein, and solvent system. In these simulations, covalent bonds are modelled as springs to calculate preferred length, mathematical equations are used to calculate motion, and intermolecular interactions, such as electrostatic and van der Waals interactions, are considered. [31]

A variety of intermolecular interactions were observed within all conformations – at the carboxyl moiety, the carbonyl oxygen and ionise oxygen interacted *via* hydrogen bonding with Lys32 and Arg57, respectively. In the case of the hipposudoric acid scaffold, the carbonyl moiety was a hydrogen bond acceptor with Arg52, and at the naphthalene moiety, an arene-H interaction was identified between the ring and Gly15. All the aforementioned interactions were conserved across time-dependent conformations, suggesting these binding motifs are strong enough to persist over time.

In addition, transient binding occurred across conformations between residues Trp22 and Ile50 and the ligand. Several transient interactions of water molecules as hydrogen bond donor or acceptors were observed towards the ligand, as well as transient water bridges between residues of the protein and the ligand.

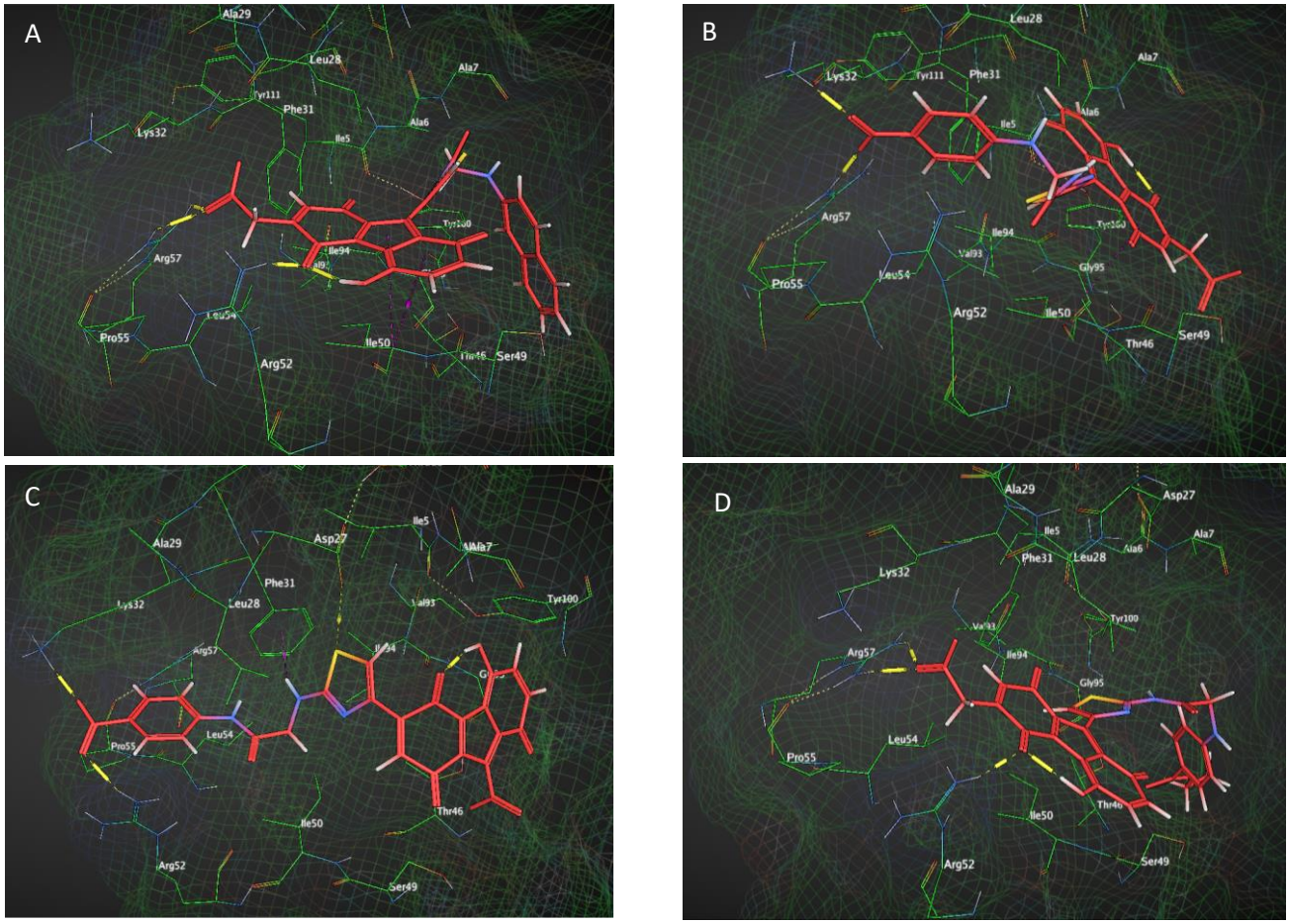


Figure 6a. Ligand-receptor binding of the best predicted binders to DHFR (A): $3r_{12}$; (B): $3r_9$; (C): $1s_9$; (D): $3r_{10}$. The H-bonds are represented with yellow lines with a cylinder. The ionic interactions are represented with yellow dotted lines. The H-pi bonds are represented by the violet dotted lines.

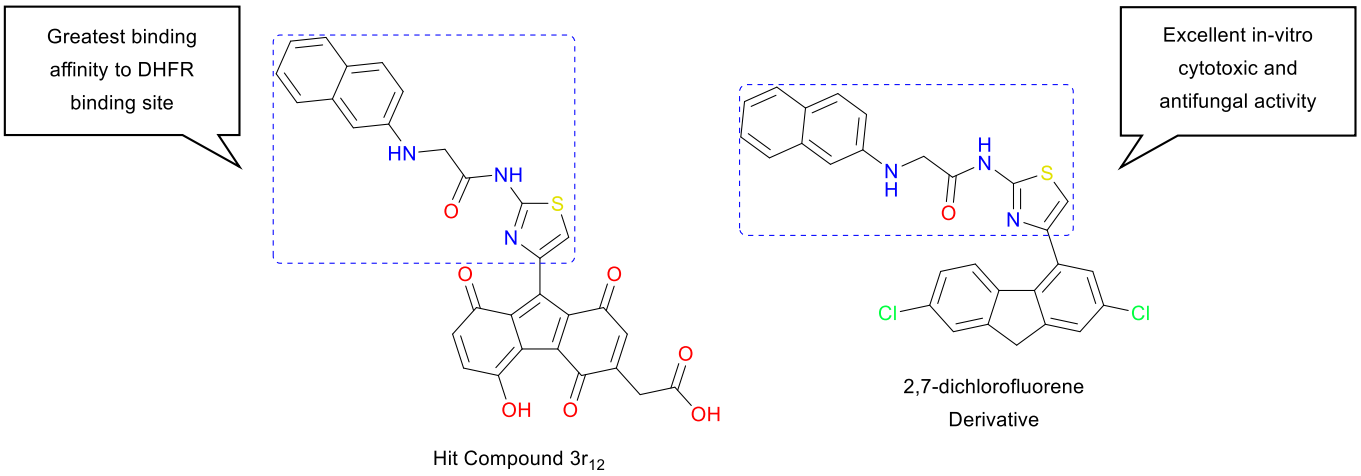


Figure 6b : Comparison of compound $3r_{12}$ which exhibited the greatest binding affinity and the corresponding 2,7-dichlorofluorene derivative which exhibited excellent *in-vitro* activity. The blue dotted line indicates the naphthalen-2-amino acetamido thiazole moiety.

Table 2. Ligand-receptor Interactions of the top 10 predicted binders ordered by score

Compound	Bonding (<5 Å)			Interaction Type	Binding Affinity (kcal/mol)	Score (kcal/mol)
	Ligand Moiety	Protein Moiety (Atom, Residue)	Hydrogen Bonds Distance (Angstrom)			
Control	NA2 3	OD1, ASP27	3.28	H-donor	-3.7	-7.0240
	O 33	NH2, ARG52	2.92	H-acceptor	-2.4	
	O2 40	NH1, ARG57	2.93	H-acceptor	-11	
	OE1 48	NZ, LYS32	2.84	H-acceptor	-17.8	
	OE2 49	NH1, ARG52	3.12	H-acceptor	-1.6	
	OE2 49	NH2, ARG52	2.77	H-acceptor	-7	
	O2 40	NH1, ARG57	2.93	ionic	-4.9	
	O2 40	NH2, ARG57	3.56	ionic	-1.7	
	OE1 48	NZ, LYS32	2.84	ionic	-5.6	
	OE2 49	NH1, ARG52	3.12	ionic	-3.7	
	OE2 49	NH2, ARG52	2.77	ionic	-6.2	
$3r_{12}$	O22	NH2, ARG52	2.75	H-acceptor	-4.9	-8.0037
	O40	NH1, ARG57	3.04	H-acceptor	-4.3	
	O40	NH2, ARG57	3.06	H-acceptor	-4.7	
	O40	NH1, ARG57	3.04	ionic	-4.2	
	O40	NH2, ARG57	3.06	ionic	-4.1	
	6-ring	CA, GLY15	3.9	pi-H	-0.9	
	6-ring	CA, ILE50	4.34	pi-H	-0.6	
$3r_9$	O35	NZ, LYS32	2.87	H-acceptor	-13.7	-7.9436
	O36	NH2, ARG57	2.93	H-acceptor	-11.8	
	O35	NZ, LYS32	2.87	ionic	-5.4	
	O36	NZ, LYS32	4	ionic	-0.5	
	O36	NH1, ARG57	3.89	ionic	-0.7	
	O36	NH2, ARG57	2.93	ionic	-5	
$1s_9$	S2	OD2, ASP27	3.71	H-donor	-0.8	-7.9410
	O35	NH2, ARG52	3.14	H-acceptor	-5.5	
	O36	NZ, LYS32	2.83	H-acceptor	-23.2	
	O35	NH1, ARG52	3.93	ionic	-0.6	
	O35	NH2, ARG52	3.14	ionic	-3.6	
	O36	NZ, LYS32	2.83	ionic	-5.7	
	N23	6-ring, PHE31	3.67	H-pi	-0.7	
	6-ring	N, GLY15	3.63	pi-H	-1.1	
$3r_{10}$	O22	NH2, ARG52	2.81	H-acceptor	-2.5	-7.9295
	O41	NH1, ARG57	3.19	H-acceptor	-2.8	
	O41	NH2, ARG57	2.93	H-acceptor	-8.7	
	O41	NH1, ARG57	3.19	ionic	-3.3	
	O41	NH2, ARG57	2.93	ionic	-4.9	
$3t_{10}$	O34	NZ, LYS32	3.22	H-acceptor	-2.7	-7.7547
	O34	NH2, ARG57	3.51	H-acceptor	-1.6	
	O35	NZ, LYS32	3.44	H-acceptor	-3	
	O34	NZ, LYS32	3.22	ionic	-3.2	
	O34	NH2, ARG57	3.51	ionic	-1.8	
	O35	NZ, LYS32	3.44	ionic	-2.1	
$3r_{11}$	O39	NZ, LYS32	2.98	H-acceptor	-15.9	-7.7324
	O40	NH2, ARG52	2.81	H-acceptor	-7.5	
	O39	NZ, LYS32	2.98	ionic	-4.6	
	O40	NH1, ARG52	3.59	ionic	-1.6	
	O40	NH2, ARG52	2.81	ionic	-5.9	
	N21	6-ring, PHE31	3.69	cation-pi	-1.4	
	6-ring	CD1, LEU28	4.42	pi-H	-0.6	
	5-ring	CD2, LEU28	4.15	pi-H	-0.7	
	6-ring	CA, ILE50	4.59	pi-H	-0.8	
$3r_4$	6-ring	NZ, LYS32	4.27	pi-cation	-0.8	-7.7087
	5-ring	CG1, ILE50	4.21	pi-H	-0.6	
$3t_{10}$	None	None	None	None	None	-7.6892
$3r_7$	O22	NH2, ARG57	2.73	H-acceptor	-5.6	-7.6750
	O39	NH1, ARG57	3.07	H-acceptor	-3.7	
	O39	NH2, ARG57	2.99	H-acceptor	-5.7	
	O39	NH1, ILE50	3.07	ionic	-4	
	O39	NH2, ARG57	2.99	ionic	-4.6	
	6-ring	CA, ILE50	4.34	pi-H	-0.6	
$3t_{12}$	6-ring	CG2, THR46	4.25	pi-H	-0.6	-7.6644
	6-ring	NH2, ARG52	3.41	pi-cation	-0.7	

Control: Dihydrofolic acid. Compounds Ordered by score; Score: lower scores are considered favourable

Interestingly, the pharmacophore features of the co-crystallised ligand (MTX) and the hit compound ($3r_{12}$, binding energy: -8.0037 kcal/mol) were found to be similar (see Figure 7b). In both cases, a negatively ionisable carboxyl group hydrogen bonded to Arg57; the carbonyl moiety was a hydrogen bond acceptor with Arg52, and we identified an arene interaction with Ile50. Many of the high scoring hipposudoric acid derivatives revealed an interaction with Arg57 and/or Arg52, suggesting that these binding motifs may be important for DHFR inhibitory activity. It is speculated that the pharmacophoric similarities between the known DHFR inhibitor (MTX) and the hipposudoric acids could explain its discovered antimicrobial activity. [17]

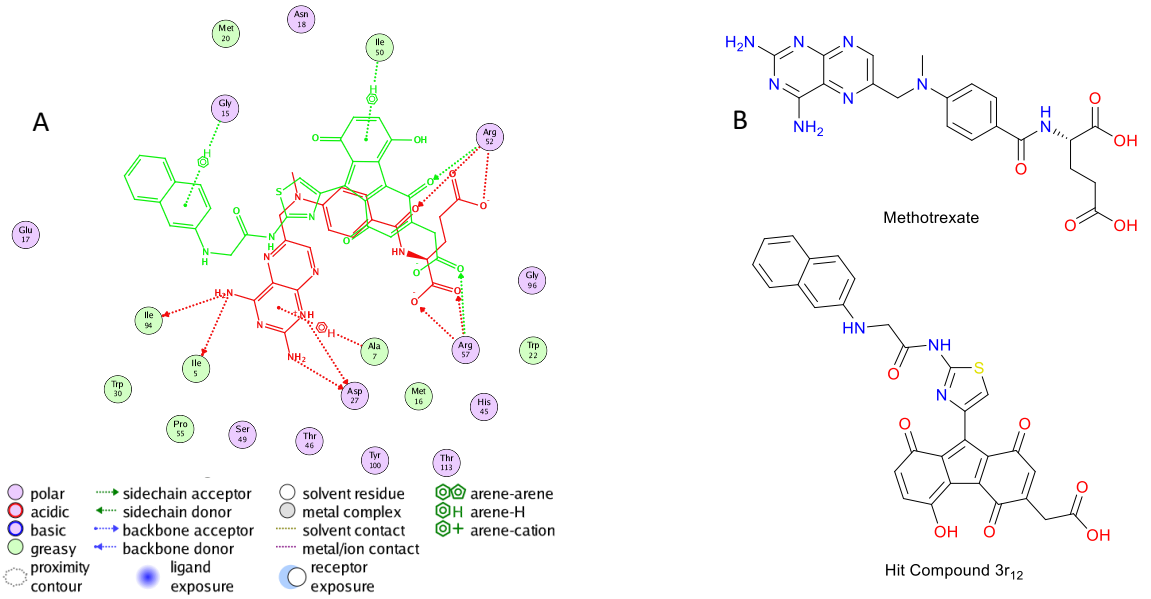


Figure 7a. : (A) Overlay between the co-crystalized MTX (red), and the hit compound $3r_{12}$ (green), and their respective binding interactions (B) 2D structures of MTX and the hit compound $3r_{12}$.

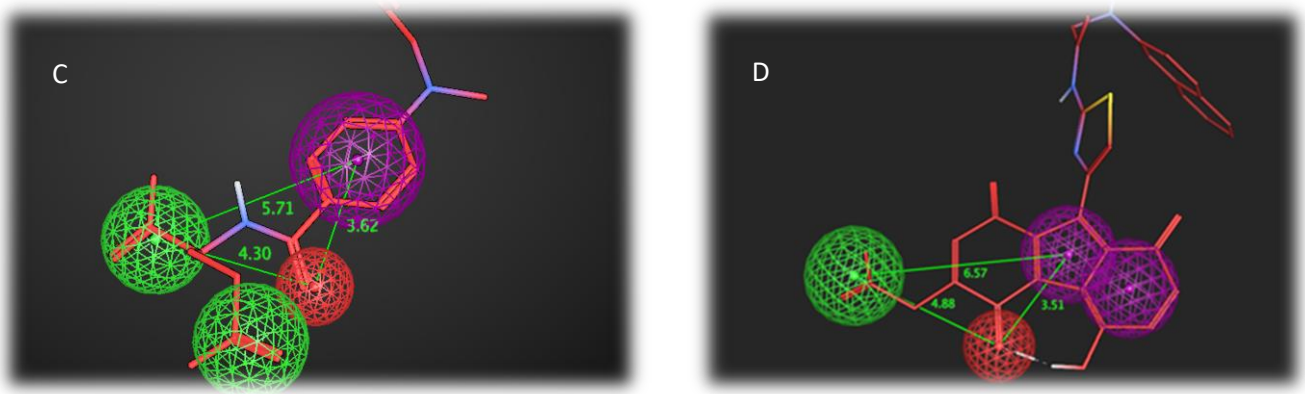


Figure 7b. Pharmacophoric similarities between the co-crystallised ligand MTX (C) and the hit compound $3r_{12}$ (D). Spheres Represent the chemical features: Red is Hydrogen Bond Acceptor (HBA), Purple Ring Aromatic (RA), Green represents a Negative Ionisable (NI) region

3.8 ADME and drug-likeness predictions

We subjected the novel derivatives to further analysis to predict their ADME and drug-likeness characteristics. For these studies, we used SwissADME, [33] which employs Lipinski's rule of 5 to determine which compounds are likely to be good candidates as orally administered drugs. Developed in 1997, these guidelines state that a molecule should not violate more than one of the following rules: it should possess a molecular mass less than 500 g/mol, a partition coefficient log P ranging from −0.4 to +5.6, no more than 5 hydrogen bond donors, and no more than 10 hydrogen bond acceptors. [34]

We subjected all 192 compounds to analysis with Swiss ADME (see supplementary information) and the top 10 best binders (Table 2) were examined further in terms of their molecular properties (see Table 3). The screened compounds $3r_4$, $3r_{11}$, $3r_{12}$, and $3t_{12}$ all

adhered to Lipinski's rule of 5 with only one violation (molecular weight > 500 g/mol), among these, compounds $3r_4$ and $3r_{12}$ were predicted to have a low absorption from the gastrointestinal tract, while the predicted gastrointestinal tract absorptions for compounds $3r_{11}$ and $3t_{12}$ were high, suggesting these may be promising candidates for an oral bioavailable drug (see Figure 8).

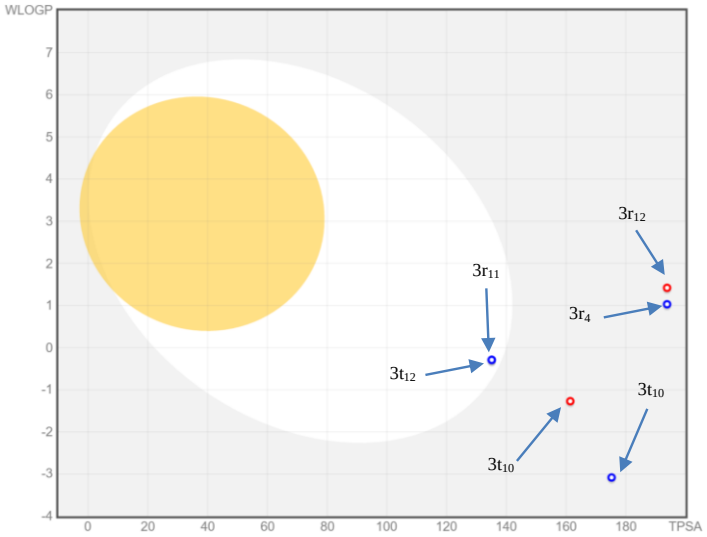


Figure 8. Boiled egg graph of the top 10 compounds with the highest binding scores. The BOILED-Egg model illustrates the predicted gastrointestinal absorption and brain penetration of molecules. The yellow ellipse depicts the compounds able to permeate through the BBB to access the CNS. The white ellipse depicts compounds able to be passively absorbed by the gastrointestinal tract. Molecules not predicted to be well absorbed nor BBB permanent are in the grey zone. Blue and red dots represent molecules that are predicted to be a substrate of the P-glycoprotein or non-substrate of the P-glycoprotein, respectively. [33]

Table 3 : ADME and drug-likeness predictions of the 10 best binders

Compound	Molecular Weight (g/mol)	Lipophilicity (MLog P)	Hydrogen Bond Acceptors	Hydrogen Bond Donors	log Kp (cm/s)	GI absorption	Abbott Bioactivity score	Lipinski violations (Follows)
$3r_{12}$	564.54	0.44	8	3	-8.21	Low	0.11	1 (Yes)
$3r_9$	557.49	-1.6	10	3	-9.39	Low	0.11	2 (No)
$1s_9$	543.46	-0.69	10	3	-9.14	Low	0.11	2 (No)
$3r_{10}$	586.55	-0.1	10	3	-9.08	Low	0.11	2 (No)
$3t_{10}$	528.47	-4.14	10	1	-11.11	Low	0.11	2 (No)
$3r_{11}$	535.52	-3.21	8	1	-9.94	High	0.55	1 (Yes)
$3r_4$	593.38	0.38	8	3	-8.78	Low	0.11	1 (Yes)
$3t_{10}$	557.53	-3.75	10	1	-10.8	Low	0.17	2 (No)
$3r_7$	544.51	-0.48	9	3	-9.00	Low	0.11	2 (No)
$3t_{12}$	535.52	-3.21	8	1	-9.94	High	0.55	1 (Yes)

The Abbott Bioavailability Score: The probability of oral bioavailability of the compound of at least 10% in rats. [35]

4. Conclusions and future outlook

To date, there have only been two reported bioactive compounds with the hipposudoric acid scaffold and Hussain et al. have reported a series of bioactive compounds with the 2,7-dichlorofluorenes scaffold. [13-16] The structural resemblance of the hipposudoric acids to the 2,7-dichlorofluorenes scaffold provide a promising avenue for the discovery of novel hipposudoric acid analogues. The principal aim of this study was to investigate the binding interaction of hipposudoric acids within the DHFR binding site. It was found that the compounds had an affinity towards the DHFR binding site, with several of the virtual compounds showing binding affinities that exceeded the affinity of the docked endogenous ligand (dihydrofolic acid).

Analysis of the best binders revealed similarities in the binding motifs relative to the co-crystallised ligand, with the highest scoring compound ($3r_{12}$) possessing a similar pharmacophore. According to this discovery, it is conceivable that hipposudoric acids may exhibit their reported antimicrobial activity by operating in a similar manner as this known DHFR inhibitor.

This study has generated a diverse array of novel hipposudoric acids and has identified potential hit compounds. Our future studies will aim to synthesise the high scoring compounds in order to explore their pharmacological activity and to establish structure-activity relationships and generate lead compounds.

Conflicts of interest

The authors declare no conflict of interest.

Notes and references

1. Fair RJ, Tor Y. Antibiotics and bacterial resistance in the 21st century. *Perspect Medicin Chem* 6 25-64 (2014).
2. Antimicrobial Resistance C. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *Lancet* 399(10325), 629-655 (2022).
3. Rao KN, Venkatachalam SR. Dihydrofolate reductase and cell growth activity inhibition by the beta-carboline-benzoquinolizidine plant alkaloid deoxytubulosine from Alangium lamarckii: its potential as an antimicrobial and anticancer agent. *Bioorg Med Chem* 7(6), 1105-1110 (1999).
4. Du QR, Li DD, Pi YZ *et al*. Novel 1,3,4-oxadiazole thioether derivatives targeting thymidylate synthase as dual anticancer/antimicrobial agents. *Bioorg Med Chem* 21(8), 2286-2297 (2013).
5. Askari BS, Krajinovic M. Dihydrofolate reductase gene variations in susceptibility to disease and treatment outcomes. *Current genomics* 11(8), 578-583 (2010).
6. Blakley RL, Cocco L. Dismutation of dihydrofolate by dihydrofolate reductase. *Biochemistry* 23(11), 2377-2383 (1984).
7. Alotaibi M, Reyes BD, Le T *et al*. Structure-based analysis of Bacilli and plasmid dihydrofolate reductase evolution. *J Mol Graph Model* 71 135-153 (2017).
8. Brown KA, Kraut J. Exploring the molecular mechanism of dihydrofolate reductase. *Faraday Discussions* 93(0), 217-224 (1992).
9. Hsiao Y-L, Chang P-C, Huang H-J, Kuo C-C, Chen CY-C. Treatment of acute lymphoblastic leukemia from traditional chinese medicine. *Evidence-based complementary and alternative medicine : eCAM* 2014 601064-601064 (2014).
10. Raimondi MV, Randazzo O, La Franca M *et al*. DHFR Inhibitors: Reading the Past for Discovering Novel Anticancer Agents. *Molecules (Basel, Switzerland)* 24(6), 1140 (2019).
11. Bennett BC, Wan Q, Ahmad MF, Langan P, Dealwis CG. X-ray structure of the ternary MTX.NADPH complex of the anthrax dihydrofolate reductase: a pharmacophore for dual-site inhibitor design. *Journal of structural biology* 166(2), 162-171 (2009).
12. Hussein EM, Al-Rooqi MM, Abd El-Galil SM, Ahmed SA. Design, synthesis, and biological evaluation of novel N4-substituted sulfonamides: acetamides derivatives as dihydrofolate reductase (DHFR) inhibitors. *BMC Chemistry* 13(1), 91 (2019).
13. Hussein EM, Alsantali RI, Abd El-Galil SM *et al*. Bioactive fluorenes. part I. Synthesis, pharmacological study and molecular docking of novel dihydrofolate reductase inhibitors based-2,7-dichlorofluorene. *Heliyon* 5(6), e01982 (2019).
14. Alsantali RI, Hussein EM, Obaid RJ *et al*. Bioactive Fluorenes. Part II. Unprecedented biologically active thiazole derivatives based-2,7-dichlorofluorene as competent DHFR inhibitors: Design, synthesis, and molecular docking approaches. *Arabian Journal of Chemistry* 13(5), 5451-5462 (2020).
15. Hussein EM, Alsantali RI, Morad M *et al*. Bioactive fluorenes. Part III: 2,7-dichloro-9H-fluorene-based thiazolidinone and azetidinone analogues as anticancer and antimicrobial against multidrug resistant strains agents. *BMC Chem* 14(1), 42 (2020).
16. Hussein EM, Malik MS, Alsantali RI *et al*. Bioactive fluorenes. Part IV: Design, synthesis, and a combined in vitro, in silico anticancer and antibacterial evaluation of new fluorene-heterocyclic sulfonamide conjugates. *Journal of Molecular Structure* 1246 131232 (2021).
17. Hashimoto K, Saikawa Y, Nakata M. Studies on the red sweat of the Hippopotamus amphibius %J Pure and Applied Chemistry. 79(4), 507-517 (2007).
18. Bolin JT, Filman DJ, Matthews DA, Hamlin RC, Kraut J. Crystal structures of Escherichia coli and Lactobacillus casei dihydrofolate reductase refined at 1.7 A resolution. I. General features and binding of methotrexate. *Journal of Biological Chemistry* 257(22), 13650-13662 (1982).
19. Rossum GV, Drake FL. *Python 3 Reference Manual*. CreateSpace, (2009).
20. O'boyle NM, Banck M, James CA, Morley C, Vandermeersch T, Hutchison GR. Open Babel: An open chemical toolbox. *Journal of Cheminformatics* 3(1), 33 (2011).
21. Chemical Computing Group Inc. SSW, Suite #910, Montreal, Qc, Canada, H3a 2r7, 2015. Molecular Operating Environment (MOE). (2015.10).
22. Halgren TA. Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *Journal of Computational Chemistry* 17(5-6), 490-519 (1996).
23. Galli CL, Sensi C, Fumagalli A, Parravicini C, Marinovich M, Eberini I. A Computational Approach to Evaluate the Androgenic Affinity of Iprodione, Procymidone, Vinclozolin and Their Metabolites. *PLOS ONE* 9(8), e104822 (2014).
24. Molecular Operating Environment (MOE), Scientific Vector Language (SVL) source code provided by Chemical Computing Group Inc., 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2015.
25. Molecular Operating Environment (MOE), Obtained from Content Manuel provided by Chemical Computing Group Inc., 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2015.
26. Wlodawer A, Minor W, Dauter Z, Jaskolski M. Protein crystallography for non-crystallographers, or how to get the best (but not more) from published macromolecular structures. *The FEBS journal* 275(1), 1-21 (2008).

27. Biovia DS, San Diego: Dassault Systèmes. Discovery Studio Visualizer. (Release 2021).
28. Halgren TA. Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *J Comput Chem.* Vol. 17((5-6)), (1996).
29. SciFinderⁿ CAS [Internet]. Scifinder-n.cas.org. 2022 [accessed 20 March 2022]. Available from: https://scifinder-n.cas.org/search/reaction/623caea97c008e5975750d45/1 (Structure search yielded no hits).
30. In the paper by Hussein et al. the 2,7-dichlorofluorene derivative is named compound 6j.
31. Hollingsworth SA, Dror RO. Molecular Dynamics Simulation for All. *Neuron* 99(6), 1129-1143 (2018).
32. Fujitsu Limited AaHWFP. Scigress [Computer Program]. (Release 2019, Version 2.9.1).
33. Daina A, Michielin O, Zoete V. SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Scientific Reports* 7(1), 42717 (2017).
34. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev* 46(1-3), 3-26 (2001).
35. Martin YC. A bioavailability score. *Journal of medicinal chemistry* 48(9), 3164-3170 (2005).