**MASTER'S THESIS**

MASTER  OF **«** Medical Biotechnology **»**

OPTION:  **«**  Biomedical **»**

# Title

# Exploration of new anti-tuberculosis drugs by targeting dihydrofolate reductase

**Realized by:**

JAOUID Oussama

The 27th June, 2022

In front of the jury composed of:

**Pr. Azeddine Ibrahimi**, School of Medicine and Pharmacy of Rabat, **President**

**Pr. Ilham Kandoussi**, School of Medicine and Pharmacy of Rabat, **Advisor**

**Pr. Mouna Ouadghiri**, School of Medicine and Pharmacy of Rabat, **Examiner**

# Dedications

To my dear parents

I cannot find the right and sincere words to express my gratitude and affection. Your prayers and blessing have been a great help.

To my brothers Reda and Abdelmounaim for their love, support and generosity, I could not have made it through this wonderful experience without you.

To my friends Issam and Hamza, who were always there to help me find motivation to never give up on my goals.

To JAOUID family

As a testament to my gratitude, love and affection for you, thank you for your availability and unconditional support.

I dedicate this work to you with all my best wishes for happiness, health and success.

# Acknowledgments

I would like to express my sincere gratitude to **Pr. Azeddine Ibrahimi** the head of Biotechnology Laboratory for the opportunity he offered me to be part of Medbiotech family. His valuable advices and encouragements have been a great help to me during my journey at the school of medicine and pharmacy.

I am extremely grateful to my academic advisor **Pr. Ilham Kandoussi** for the continuous support during the realization of this study, for her patience, availability, and generosity. It was a great privilege to work and study under her guidance.

Special thanks to **Pr. Mouna Ouadghiri** the coordinator of the MSc Medical Biotechnology: Biomedical option, for her efforts during the organization of courses and exams, and for keeping us updated.

Thank you to all who have contributed directly or indirectly to the realization of this work.

# Abstract

Tuberculosis (TB) is one of the ten leading causes of death in the world and the main source of infection from a single infectious agent, the World Health Organization (WHO) states that one third of the world's population is now infected with TB, therefore the increase in TB treatment and so a series of drug-resistant strains have emerged, diagnosis and therapy of multidrug-resistant TB continues to be a major hurdle and is far from being fully solved. Given the urgency of the situation, the current study uses the advantages of virtual high-throughput screening approaches to identify molecules targeting *Mycobacterium tuberculosis* dihydrofolate reductase (Mt-DHFR), an enzyme critical for *Mycobacterium tuberculosis* proliferation. In the extension of the Mt-DHFR ligand pocket, there is a small hydrophobic pocket that hosts a glycerol molecule (GOL), this pocket does not exist in the human protein. Based on these data, our study explored new *Mycobacterium tuberculosis* specific inhibitors targeting Mt-DHFR by the in-silico approach. From a set of 8412 compounds, toxicity evaluation and validation of Lipinski and Veber's rule allowed to identify 11 new small molecules whose interaction with the target with and without glycerol were studied by Docking, the results were compared with 5 reference molecules chosen from the literature in addition to natural ligand of the target. Thus for the evaluation of specificity, these molecules were also tested on human DHRF, the results showed that Methotrexate is the best reference inhibitor and that the affinities of the molecules towards Mt-DHFR with glycerol are better than without glycerol, It is so 8 molecules present an affinity towards the bacterial enzyme with glycerol better than that of Methotrexate, also 9 presenting a better affinity than that of natural substrate, among which 3 have a weak affinity towards the human enzyme compared to the bacterial enzyme. In the light of the results obtained we propose 3 inhibitors of *Mycobacterium tuberculosis* targeting Mt-DHFR with better activity and interactions than the reference inhibitors and which conform to the rules of Lipinski and Veber supposed to have antitubercular potential.

**Key words:** World Health Organization (WHO), Tuberculosis, *Mycobacterium tuberculosis* dihydrofolate reductase (Mt-DHFR), Structure-based virtual screening, Molecular Docking

# Résumé

La tuberculose (TB) est l'une des dix causes de décès dans le monde et la principale source d'infection à partir d'un seul agent infectieux, l'Organisation mondiale de la santé (OMS) affirme qu'un tiers de la population mondiale est aujourd'hui infecté par la tuberculose, par conséquence l'augmentation du traitement de la tuberculose et donc une série de souches multirésistantes ont émergé, le diagnostic et la thérapie de la tuberculose multirésistante continue d'être un obstacle majeur et est loin d'être entièrement résolu .Compte tenu de l'urgence de la situation, l'étude actuelle utilise les avantages des approches de criblage virtuel à haut débit pour identifier des molécules ciblant la dihydrofolate réductase de *Mycobacterium tuberculosis* (Mt-DHFR), une enzyme critique pour la prolifération de *Mycobacterium tuberculosis*. Dans le prolongement de la poche ligand de la Mt-DHFR, il existe une petite poche hydrophobe qui accueille une molécule de glycérol (GOL), cette poche n'existe pas dans la protéine humaine. Sur la base de ces données notre étude a fait l'objet d'explorer de nouveaux inhibiteurs spécifiques de *Mycobacterium tuberculosis* ciblant la Mt-DHFR par l'approche in silico. À partir d'un ensemble de 8412 composés, l'évaluation de la toxicité et la validation de la règle de Lipinski et veber ont permis d'identifier 11 nouvelles petites molécules dont l'interaction avec la cible avec et sans glycérol ont été étudiées par Docking, les résultats ont été comparées avec 5 molécules de référence choisis à partir de la littérature en plus de ligand naturel de la cible. Ainsi pour l'évaluation de la spécificité, ces molécules ont été également testées sure la DHRF humaine, les résultats ont montré que le Methotrexate est le meilleur inhibiteur de référence et que les affinités des molécules envers la Mt-DHFR avec glycérol sont meilleures que celle sans glycérol, c'est ainsi 8 molécules présentent une affinité envers l'enzyme bactérienne avec glycérol meilleure que celle de la Methotrexate, aussi 9 présentant une meilleure affinité que celle de substrat naturel , parmi lesquelles 3 ont une affinité faible envers l'enzyme humain par rapport à l'enzyme bactérienne . A la lumière des résultats obtenue nous proposons 3 inhibiteurs du *Mycobacterium tuberculosis* ciblant la Mt-DHFR présentant une activité et des interactions meilleurs que les inhibiteurs de référence et qui conforment aux règles de Lipinski et Veber supposés avoir un potentiel antituberculeux.

**Mots clés** : Organisation mondiale de la santé (OMS), Tuberculose, *Mycobacterium tuberculosis* dihydrofolate réductase (Mt-DHFR), Criblage virtuel basé sur la structure, Docking moléculaire.

# List of tables

# List of Figures

# Table of Contents

# List of Abbreviations

| | |
|---|---|
| 2D | Two Dimensions |
| 3D | Three Dimensions |
| ADT | AutoDock Tools |
| ADME | Absorption, Distribution, Metabolism, and Excretion |
| BCG | Bacille Calmette-Guerin |
| CCL2 | Macrophage Chemokine |
| CADD : | Computer-Aided Drug Design |
| DNA | Deoxyribonucleic Acid |
| DHFR | Dihydrofolate reductase |
| EMB | Ethambutol |
| GUI | Graphical User Interface |
| HIV | Human immunodeficiency virus |
| h-DHFR | Human dihydrofolate reductase |
| IFNs | Interferons |
| IUPAC | International Union of Pure and Applied Chemistry |
| INH | Isoniazid |
| HIV | Human immunodeficiency virus |
| Kb | Kilo base |
| LTBI | Latent Tuberculosis Infection |
| LBVS | Ligand-Based Virtual Screening |
| LBDD | Ligand-Based Drug Design |
| Mtb | *Mycobacterium tuberculosis* |
| MDR-TB | Multi Drug Resistance tuberculosis |
| MTBC | *Mycobacterium tuberculosis* Complexe |
| Mt-DHFR | *Mycobacterium tuberculosis* dihydrofolate reductase |
| MTX | Methotrexate |
| PAMPs | Pathogens activated molecular patterns |
| PDB | Protein Data Bank |
| PPD | Purified Protein Derivative |

| | |
|---|---|
| PCR | Polymerase chain reaction |
| PYR | Pyrimethamine |
| QSAR | Quantitative Structure-Activity Relationship |
| RR TB | Rifampicin Resistance tuberculosis |
| RIF | Rifampicin |
| RNA | Ribonucleic Acid |
| SM | Streptomycin |
| SBDD | Structure-Based Drug Design |
| SBVS | Structure-Based Virtual Screening |
| TB | Tuberculosis |
| TLR | Toll-like receptor |
| TST | Tuberculin skin test |
| WHO | World Health Organization |
| XDR | Extensively Drug Resistanc |

# INTRODUCTION

Several unique features characterize infectious diseases - they rely on a single agent as a cause, they can be transmitted from one person to another, cause epidemics, and have a strong impact on human evolution[1]. Besides, infectious diseases can be eradicated, but also new ones may emerge, creating a dynamic stage for human-infection interplay [2]. Tuberculosis (TB), which remains a significant disease for public health worldwide and especially in emerging economies, tuberculosis is one of the world's ten causes of death worldwide and the primary infection source from a single infectious agent [3]Despite the sheer numbers, TB has been over shadowed for too long by HIV and malaria, and currently by the COVID-19. The World Health Organization (WHO) claims that 1/3 of the worldwide population is now infected with tuberculosis. About 10% of this population are predicted to develop active TB at some stage in their lifetime [4].

With the increase in TB treatment, a series of drug-resistant strains have emerged, diagnosis and therapy of multidrug-resistant TB continues to be a major hurdle and is far from being fully solved [5]. Rifampicin resistance was found in 61% of individuals with pulmonary TB worldwide in 2019, up from 51% in 2017 and 7% in 2012. A total of 206,030 individuals with MDR/RR TB were identified and recorded, a 10% rise from 2018[6]. Extended drug resistance (XDR) strains, likely stem from MDR/RR TB strains. The XDR-TB show resistant to second-line drugs fluoroquinolones (FLQ) and aminoglycosides (AMI), has been registered in 92 countries. It is estimated that 6% of multidrug-resistant TB cases are estimated to be XDR-TB [3].

Hence novel antimycobacterial drugs are urgently required to combat this growing health emergency. Alongside this, increased knowledge of gene essentiality in the pathogenic organism and larger compound databases can aid in the discovery of new drug compounds. The number of protein structures, X-ray based and modelled, is increasing and now accounts for greater than > 80% of all predicted M. tuberculosis proteins, allowing novel targets to be investigated [7].

Dihydrofolate reductase (DHFR) catalyzes the NADPH-dependent reduction of dihydrofolate to tetrahydrofolate and is critical for the synthesis of thymidylate, purines and several amino acids. Inhibition of the enzyme's activity results in the arrest of DNA synthesis and cell death. The enzyme has been widely investigated as a drug target for bacterial [7], protozoan and fungal infections, as well as for neoplastic and autoimmune diseases, inhibitors of DHFR have been proven as effective agents

for treating bacterial infections [8].

Since conventional drug research and development is an extremely long (up to 15 years), complex and costly (over a billion dollars) process. Currently, the computer-aided drug design process allows predicting possible outcomes before the process is even underway, thus saving time and money because failure at any stage of drug development would mean a huge loss to the industry. Two approaches to virtual screening are currently used: Structure-Based Drug Design (SBDD) and ligand-based Drug Design (LBDD) [9].

In the current study we used Structure-Based Drug Design (SBDD), to identify potential inhibitors of the target enzyme by screening large libraries of molecules from chemical databases. Molecular docking was used in a virtual high-throughput screening to discover potential inhibitors of *Mycobacterium tuberculosis* DHFR.

# BIBLIOGRAPHY

# I. General knowledge

## 1. History

TB or illnesses resembling TB have been described from different civilization since ancient times. The earliest such description can be found in Vedas, where TB was referred to as 'Yakshma' meaning wasting disease. Greek, Chinese and Arabic literature also describes TB like disease.[10] Mycobacterium exists on earth since last 150 million years. Typical tubercular vertebral lesions were seen in mummies from the Egyptian pre-dynastic era and Peruvian pre-Colombian era. The first weak evidence of TB in humans is from a bone lesion found in a 500 thousand year old skull in Turkey. Human TB detection using PCR sequencing in a Neolithic infant and women from 9 thousand year old settlement in the Eastern Mediterranean is the oldest strong evidence available [11].

- Galen (131e201) first suspected that TB could be contagious [11].
- It took many centuries until Girolamo Fracastorius (1483e1553) showed that some diseases could be transmitted through 'particles' by direct or indirect contact between humans [10].
- Thomas Willis (1621e1675) first described miliary TB.
- Calmette extracted a protein (tuberculin) from large cultures of the bacillus and first used for therapy known as 'tuberculinisation', which failed as treatment for TB. The Tuberculin was also used for intradermal skin test which was described by Charles Mantoux & used in the diagnosis of TB. Later this intradermal skin test was named after Charles Mantoux and is known as Mantoux test [12].
- Benjamin Marten (1690e1752) hypothesized that TB is caused by 'wonderfully minute living creatures' in his theory of 'contagious living fluid' [10].
- It was Jean Antoine Villemin (1827e1892), a French army doctor who successfully demonstrated the transmission of TB from humans to animals and from animals to animals.
- In 1834, Johann Lukas Schonlein proposed the name 'Tuberculosis' which is derived from Latin word 'tubercula' meaning 'a small lump' seen in all forms of the disease [12].
- On 24th March 1882, Robert Koch announced in the meeting of the Berlin Society of Physiology that he had discovered causative agent responsible for pulmonary TB and named it as 'tuberkel virus' in his paper published 2 weeks later [13].
- First innovative decision of staining tuberculosis bacilli and second innovative decision of culturing it on solidified cow or sheep serum gave Robert Koch the Nobel prize of medicine in 1905 [13].

- Leon Charles Albert Calmette (1863e1933) and Camille Guerin (1872e1961) developed vaccine against TB by sub-culturing Mycobacterium bovis for more than 200 times in the Guinea pig model between 1908e1921 [10], [12] .
- Arvid Wallgren, a professor from Royal Caroline medical institute, Sweden described clinical manifestations of tuberculous infection in an article titled 'The timetable of Tuberculosis' which helped in better understanding course of TB illness [13].
- The effective treatment for TB became a reality after the discovery of antitubercular drugs like Streptomycin, Paraamino salicylic acid (PAS) and isoniazid by the mid-1940s [13].
- By late 1970 it was believed that TB may no longer be a public health problem in the developed world. But the emergence of Acquired Immune Deficiency Syndrome (AIDS) in the early 1980s has ended this optimism and led to the resurgence of TB worldwide [10].

## 2. Risk factors

Among major known risk factors for TB, HIV infection is the strongest [14]; 12% of all new active TB disease cases and 25% of all TB-related deaths occur in HIV-positive individuals. The majority (75%) of HIV-associated active TB disease cases and deaths occur in Africa[15], Active TB disease was the leading cause of hospitalization among HIV-infected adults (18%) and children (10%) . TB-related in-hospital mortality was 25% among adults and 30% among children with HIV infection [16].

Other risk factors are responsible for the remaining fraction of TB cases in the general population:

- Undernutrition make an estimated 27% of TB cases worldwide[17].
- Indoor air pollution make 22% [17].
- Type 2 diabetes mellitus [18].
- Excessive alcohol use (roughly triple the risk of TB) [19].
- Smoking (which doubles the risk) [20].

## 3. Epidemiology

### 3.1. Worldwide

Tuberculosis is a contagious, bacterial, airborne disease. It is one of the 10 leading causes of death worldwide, and is the second leading cause of death by a single infectious agent, ranking behind COVID-19 and ahead of HIV/AIDS. According to the World Health Organization (WHO),

approximately 1.8 billion people are infected with TB, most of whom have the inactive form of the disease. In 2020, some 10 million people had active TB. Of these [21]:

- 5.6 million were men
- 3.3 million were women
- million were children

People living with HIV/AIDS are 15-22 times more likely to get TB due to a weakened immune system. In 2019, TB caused [21]:

- 1.23 million deaths among HIV-negative people
- 214,000 deaths among HIV-positive people

Tuberculosis causes a high number of deaths, despite being a preventable and curable disease. This is largely due to the fact that more than 4 million people worldwide who develop TB each year are not captured by public health systems [21].

**Table I : Estimated epidemiological burden of TB in 2019 globally. Rates per 100 000 population [21].**

| | POPULATION | TOTAL TB INCIDENCE | | HIV-POSITIVE TB INCIDENCE | | HIV-NEGATIVE TB MORTALITY | | HIV-POSITIVE TB MORTALITY[b] | |
|---|---|---|---|---|---|---|---|---|---|
| | | BEST ESTIMATE | UNCERTAINTY INTERVAL | BEST ESTIMATE | UNCERTAINTY INTERVAL | BEST ESTIMATE | UNCERTAINTY INTERVAL | BEST ESTIMATE | UNCERTAINTY INTERVAL |
| High TB burden countries | 4 880 000 | 8 610 | 7 600–9 680 | 668 | 585–757 | 1 040 | 966–1 120 | 165 | 134–198 |
| Africa | 1 090 000 | 2 470 | 2 190–2 750 | 595 | 515–680 | 378 | 313–448 | 169 | 139–203 |
| The Americas | 1 010 000 | 290 | 269–311 | 29 | 27–32 | 17 | 17–18 | 5.9 | 5.2–6.6 |
| Eastern Mediterranean | 717 000 | 819 | 646–1 010 | 7.9 | 5.9–10 | 76 | 65–87 | 2.7 | 2.0–3.6 |
| Europe | 930 000 | 246 | 215–281 | 30 | 23–38 | 20 | 20–21 | 4.2 | 3.1–5.5 |
| South-East Asia | 2 000 000 | 4 340 | 3 460–5 320 | 117 | 90–147 | 632 | 593–671 | 20 | 15–26 |
| Western Pacific | 1 930 000 | 1 800 | 1 480–2 150 | 36 | 28–46 | 84 | 78–91 | 6.3 | 5.2–7.5 |
| Global | 7 690 000 | 9 960 | 8 940–11 000 | 815 | 729–906 | 1 210 | 1 130–1 290 | 208 | 177–242 |

### 3.2. In Morocco

In Morocco, a total of 29,327 cases have been notified and put on treatment in 2021, within the framework of the National Tuberculosis Control Program (PNLAT), and it is clear that the young population, aged between 15 and 45, remains the most exposed [22].

Thanks to these efforts, Morocco has made significant progress resulting in a decrease of the estimated incidence of the disease by 34% and mortality by 68% during the last thirty years, with therapeutic success rates maintained at more than 85%. (This has resulted in the annual cure of more than 26,000 patients) [22].

## II. Pathological characteristics

### 1. Mode of Transmission

A simple cascade for tuberculosis transmission is proposed in which , 1 a source case of tuberculosis, generates infectious particles, that survive in the air and are inhaled by a susceptible individual who may become infected and who then has the potential to develop tuberculosis [23].

Interventions that target bacterial, host, or behavioral catalysts of transmission will interrupt tuberculosis transmission and accelerate the decline in tuberculosis incidence and mortality [24].

Understanding   who is transmitting, where transmission is occurring, and who is susceptible to infection and to disease progression, will allow us to decide what will it  take to stop tuberculosis transmission [24].
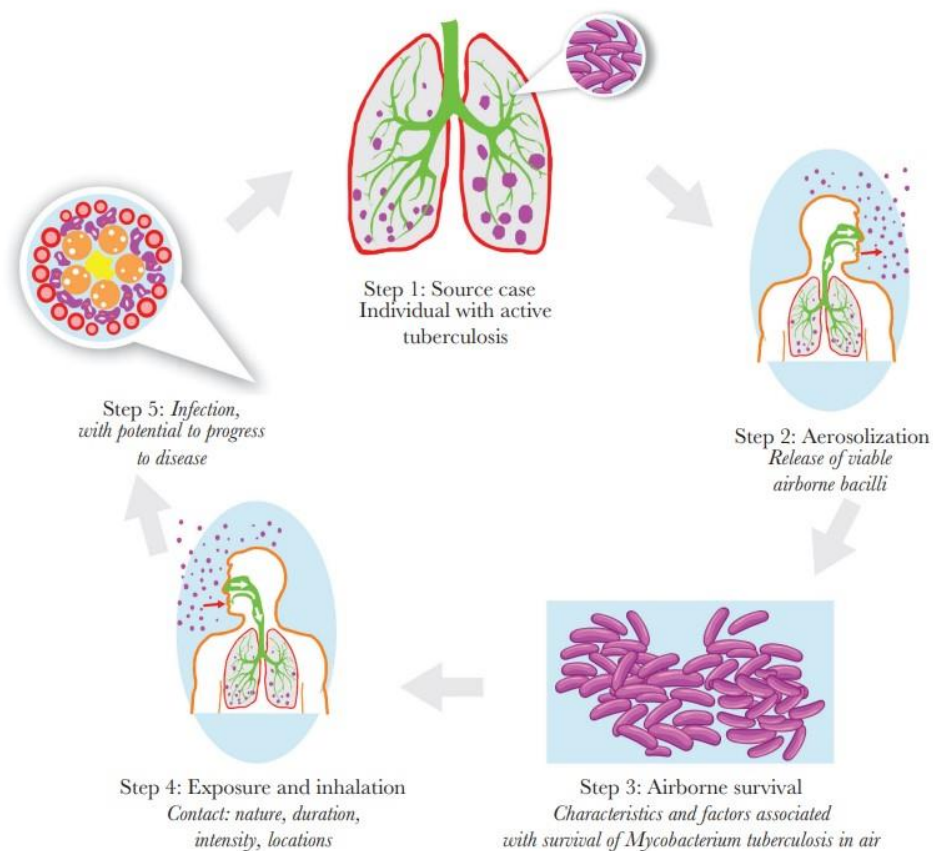


**Figure 1 : Cascade of tuberculosis transmission [24].**

## 2. Pathogenesis & Immunity

After being transferred to a new host by aerosol, M. tuberculosis is thought to be phagocytosed first by alveolar macrophages (AM) and then by interstitial macrophages [25]. In the case of other pathogens, these macrophages are microbicidal and recruited through Toll-like receptor (TLR)-mediated signaling activated by so-called pathogen-activated molecular patterns (PAMPs) present on bacterial surfaces. However, the host's innate immune system does not recognize them. Tb has developed strategies to avoid microbicidal macrophages by expressing a surface lipid, phthiocerol dimycerate (PDIM), which masks the PAMPs. Once PAMPs are masked [26].
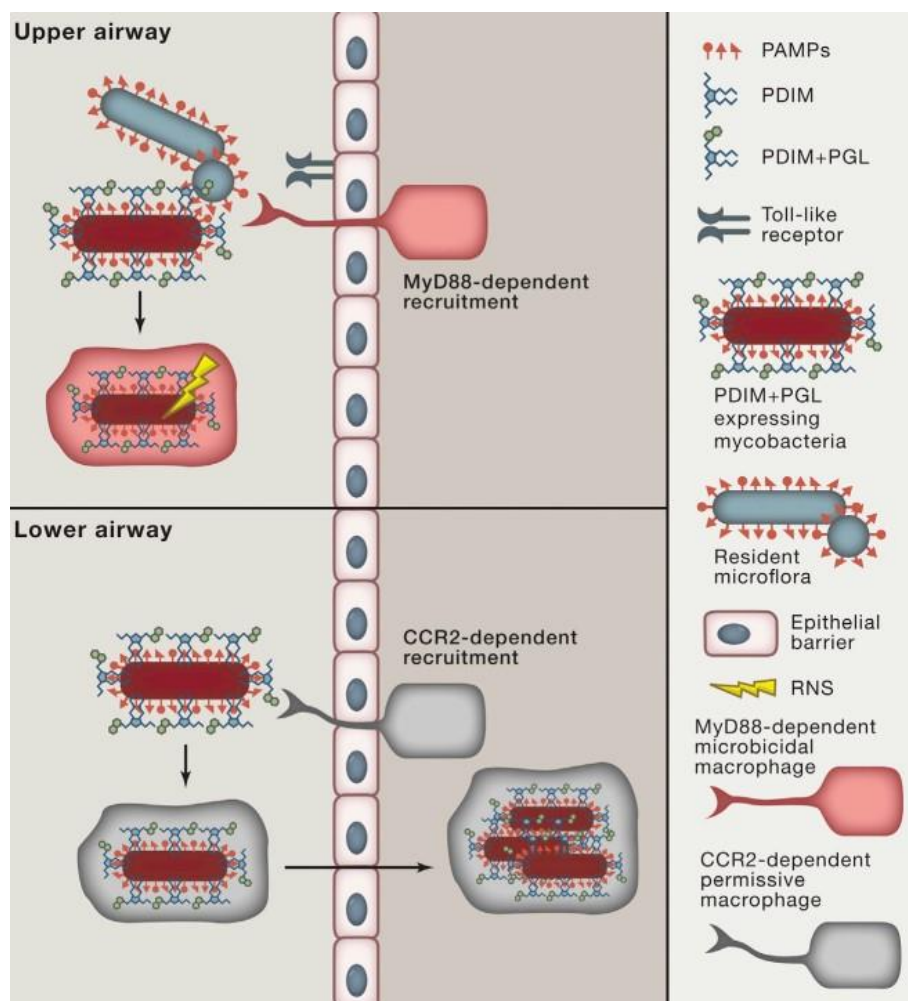


**Figure 2 : Mechanism of priminary infection of Mtb [27].**

Meanwhile MBT use a related surface lipid, phenolic glycolipid (PGL), to induce the macrophage chemokine CCL2 to recruit and infect growth-permitting macrophages [27].

The upper airways are full of TLR-stimulating commensal bacteria which make the MBT strategy is ineffective[27], Thus, tb infection must be initiated using small droplets that settle directly into the lower lung spaces since they contain little or no commensal bacteria [26].

Once internalized, ESX-1 activate the secretion system to blocks the fusion between the phagosome and lysosome. *Mycobacterium tuberculosis* releases bacterial products into the macrophage cytosol to start cytosolic surveillance which induces type I interferon response, resulting in bacteria's survival and rapid growth [27]. After phagocytosis, infected cells migrate to local draining lymph nodes. T cells recognize antigens on *Mycobacterium tuberculosis* and differentiate into specific T cells. This differentiation results in the release of lymphokines and macrophages' activation, which inhibit the growth of phagocytosed bacteria. Macrophages initiate a signaling cascade leading to the recruitment of additional monocytes and lymphocytes. This aggregate of immune cells surrounds the infection site and forms an organized cellular structure known as a granuloma [27].

After the infection, M. Tuberculosis often enters a prolonged state of latent, asymptomatic infection. This phase can last for decades within granulomas, if not for the rest of the host's life. There is a subgroup of hosts in which the latent infection re-activates, causing active disease[28].

If the bacterial load is too high, granuloma may fail to contain the infection, and bacteria will spread to other organs, including the brain. The bacterium can enter the bloodstream or respiratory tract to be released, and this is said to be the first phase of active TB disease [27].
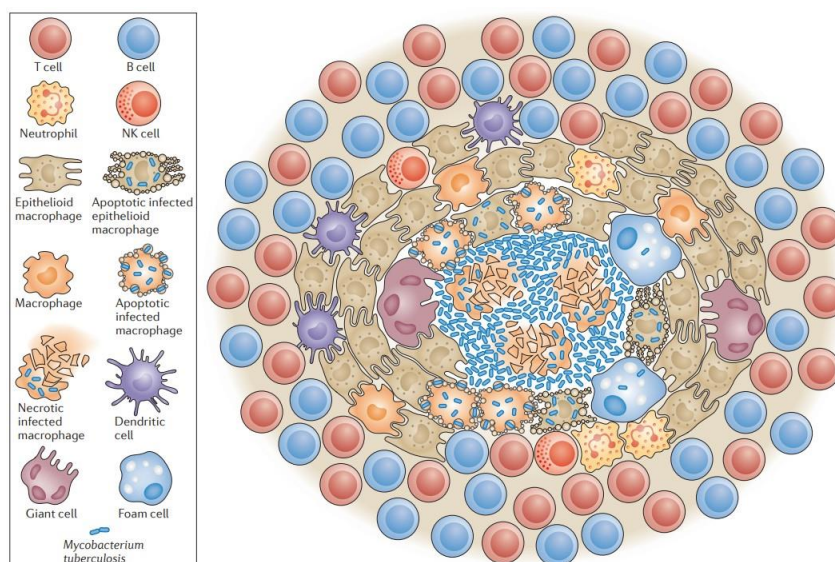


**Figure 3 : Structure and cellular constituents of the tuberculous granuloma [27].**

### 3. Clinical features of tuberculosis

Active disease may manifest with symptoms that are only minimal initially but then develop during the course of several months. Typical symptoms of active tuberculosis include a productive cough, hemoptysis, weight loss, fatigue, malaise, fever, and night sweats[29], However, some patients with active, culture-positive disease may be asymptomatic and are best described as having subclinical TB [30].

In general, clinical manifestations of illness are the first clue to the diagnosis of tuberculosis. However, they are often non-specific and misleading, and therefore the diagnosis is not always easy. This is especially the case for tuberculosis in children and in elderly subjects [31].

### 4. Taxonomy and description of the genus

*Mycobacterium tuberculosis* belongs to [11]:

- ORDER- Actinomycetales
- CLASS- Actinomycetes
- FAMILY- Mycobacteriaceae
- GENUS- *Mycobacterium* Genera

**Table II : Features of genus Mycobacterium [11].**

| Features | |
|---|---|
| **Mycobacteria** | aerobic, non-spore forming, non-motile |
| **Shape** | slightly curved or straight rods |
| **Size** | 0.2e0.6 mm by 1e10 mm |
| **Colony morphology** | varies from species to species, ranging |
| **Cell wall** | from rough to smooth and from non-pigmented to pigmented (carotenoid pigment) N-acetyl muramic acid High content of Mycolic acid (70e90 carbon atoms)-renders acid fastness |
| **DNA** | High G þ C content (61e71 mol %) |
| **Generation time** | Slow- ranging from 20 hours to 36 hours for *Mycobacterium Tuberculosis* |

## 5. Characteristics of *Mycobacterium Tuberculosis* genome

The genome of a reference M. tuberculosis strain used to this day, strain H37Rv, was sequenced and annotated in 1998 [43]. It is a genome consisting of a single circular chromosome of 4,411,532 base pairs, comprising 3,995 reading frames and forming about 4,000 genes. The DNA of M. tuberculosis has a high guanine and cytosine content (65.6%) except in a few specific regions such as the genes that code for transmembrane proteins. A large part of the genes (6% of the genome) seems to code for enzymes involved in the synthesis and degradation of lipids. More than 50% of the encoded proteins have known functions currently. The homology between the DNAs of the different MTBC subspecies is very high (>99.9%)[44], [45], which makes it a single bacterial species in the strict sense. which makes it strictly speaking a single bacterial species. This restriction of nucleotide diversity can be explained either by an unusual replication fidelity, or by a very efficient error repair system, or by a very recent evolutionary origin, or by a very low mutation rate considering the lifestyle of this bacillus (intracellular), none of these reasons being exclusive. The genome is also characterized by the use of the Guanine-Tyrosine-Guanine (GTG) codon as an initiation codon in more than 35% of genes [43].
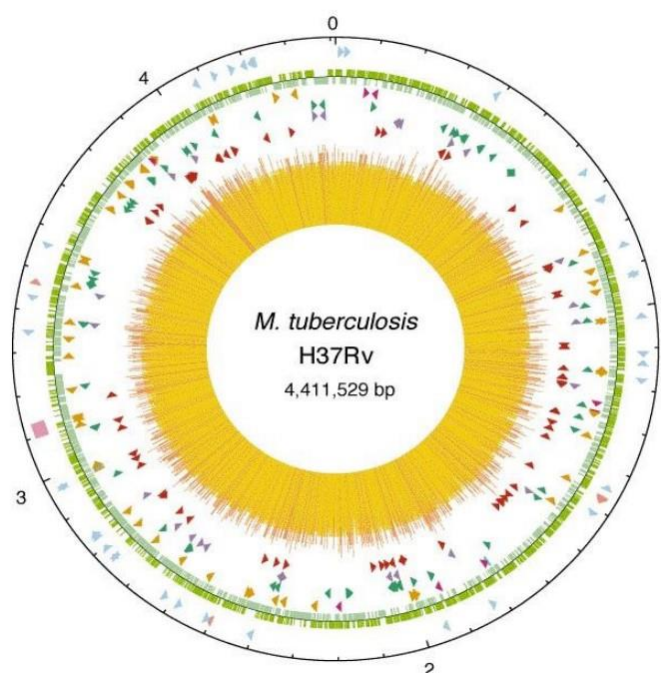


**Figure 4 : Representation of the M. tuberculosis H37Rv genome [43].**

The outer circle shows the scale in megabases, 0 represents the origin of replication. The first circle on the outside shows the positions of the stable RNA coding genes (tRNAs in blue, others in pink) and the Direct Repeat region (pink square); the second circle on the inside shows the coding sequences per strand (clockwise, dark green; counterclockwise, light green); the third circle shows the repeated sequences (insertion sequences in orange; REP13E12 family in dark pink; prophages in blue); the fourth circle shows the positions of EPP family members (green); the fifth circle shows EP family members (purple, excluding PGRS); and the sixth circle shows the positions of PGRS sequences (dark red). The histogram in the center represents the G+C content, with <65% G+C in yellow, and >65% G+C in red. The figure was generated with DNASTAR® software.

## III.  Diagnostics and treatment

### 1.  Diagnosis

Several methods for TB diagnosis and mycobacterial identification with different levels of sensitivity and specificity are currently in use. Diagnostic methods are based on direct microscopic examination of biological specimens, detection of bacilli in these specimens by culture or by molecular biology tests based on DNA or immunological tests. The identification methods are based on the morphological aspect of the colonies from the culture or on biochemical tests or molecular tests. Some molecular tests serve as both a diagnostic test for TB and a test for mycobacteria. Some of the more commonly used tests and reference tests are listed below [31]:

### 1.1.  Direct microscopic examination:

If pulmonary TB is suspected, based on a patient's clinical signs, bacteriologic tests are performed on the patient's sputum. Bacteriological tests may also be performed on gastric tubes (swallowed bronchial secretions) or bronchial aspirates. For tuberculosis patients with a suspected extra-pulmonary location, biopsies are taken according to the location of the disease [32].

A smear, or spread of the biological sample on a thin plate, is taken and then stained to reveal the specific acid-fastness of the mycobacteria. Two staining methods are used [33]:

- Ziehl-Neelsen staining
- Auramine (fluorochrome) staining

Direct examination can detect the presence of acid-fast bacilli (AFB) in a biological sample only when there are 0.5 to 1.1 bacteria per microliter of biological sample [33].

**Table III: Methods of obtaining sputum sample [33].**

| Method | Description | Advantage | Disadvantage |
|--------|-------------|-----------|--------------|
| Sputum sample | Patient coughs up sputum into sterile container | Easy to perform | <ul><li>Education and supervision of the patient is required</li><li>Patient may not be able to cough up sputum</li></ul> |
| Nebulisation/ Sputum induction | A tube is inserted into the stomach through the patients mouth or nose to obtain swallowed sputum | Used to obtain sputum in children who do not cough up sputum | <ul><li>Must be done early morning before eating</li><li>Patient may need to be hospitalized</li></ul> |
| Bronchoscopy | A scope is passed through the mouth or nose to the diseased part of the lung to obtain sputum or lung tissue | Used to obtain sputum when the patient cannot cough and gastric aspirate cannot be done. | <ul><li>Requires special equipment</li><li>Must be done in a hospitalby a specialist</li></ul> |

### 1.2. Interferon gamma release assays

In these tests, the patient's T cells are incubated ex vivo with antigens that are found in MTB but not found in BCG nor in most NTMs. Levels of interferon gamma are then measured in the supernatant or cells producing interferon gamma counted in an Enzyme-Linked ImmunoSpot (ELISPOT) assay [34].

IGRAs are a more specific test for MTB infection[35]. They can also be performed in a single visit, which is a significant advantage. However, IGRAs are expensive and require phlebotomy. There is substantially less experience informing their use in epidemiologic studies. Another major problem is that the mode of the frequency distribution for positive reactions is very close to the cut point recommended by the manufacturer. This means that small differences in the choice of cut point can have a substantial impact on both inference and prevalence estimates [34].

### 1.3. Radiography

Chest radiography is used to establish an initial assessment of thoracic lesions of varying shape, location and size. The appearance of the lesions is unrelated to the intensity of the disease. The specificity of radiography for pulmonary tuberculosis varies greatly according to different studies (27% to 81%). There are four types of lesions: the nodule, the tuberculoma, infiltrate and cavern [36].

- **The infiltrate** is manifested as early lesions of the infection. It manifests as a faint, heterogeneous and extensive part of the radiographic imaging.
- **The nodule** is a granuloma of variable size, isolated or grouped.
- **The tuberculoma** is an isolated pseudotumor nodule.
- **The cavern** is an empty area due to a loss of substance within a thick-walled infiltrate

In its pulmonary form, TB is manifested by the presence of infiltrates and nodules mainly located at the top of the lungs and sometimes associated with caverns [36].

### 1.4. Tuberculin skin test

The tuberculin skin test (TST) or Mantoux test is a diagnostic test for TB based on the observation of a delayed hypersensitivity skin reaction after intradermal injection of tuberculin (mycobacterial antigens composed of a mixture of proteins extracted from cultured MTBC and purified called PPD / "Purified Protein Derivative") [37].

A negative result does not rule out the diagnosis of TB disease as various conditions, including HIV, may suppress the reaction. This reaction indicates the existence of cell-mediated immunity to mycobacteria in the patient, induced either by prior BCG vaccination or by previous contact with Bacille de Koch or certain atypical mycobacteria. The TST is used for the diagnosis of both active and latent TB. diagnosis of latent TB [38].

The skin test result depends on the size of the raised, hard area or swelling. It also depends on the person's risk of being infected with TB bacteria and the progression to TB disease if infected [38].

- **Positive skin test**: This means the person's body was infected with TB bacteria. Additional tests are needed to determine if the person has latent TB infection or TB disease. A health care worker will then provide treatment as needed.
- **Negative skin test**: This means the person's body did not react to the test, and that latent TB infection or TB disease is not likely.

### 1.5. Molecular testing

Today, many molecular techniques suitable for the diagnosis of TB using mycobacterial DNA are used. These tests are based on gene amplification by the classical PCR (Polymerase Chain Reaction) on specific sequences of M. tuberculosis strains[39], by real-time PCR as in the case of Xpert MTB/RIF [40], or also by LAMP-PCR (Loop-mediated isothermal amplification) [41] allowing the detection of specific nucleic sequences of tubercle bacilli; this PCR may or may not be followed by hybridization of the amplification products on specific probes. However, these techniques currently have insufficient sensitivity when applied directly to biological samples, on the other hand, they have excellent sensitivity and specificity when used on culture extracts. culture extracts [42].

### 1.6. Culture

Culture is the gold standard for TB diagnosis. Only a positive culture for M. tuberculosis is definitive proof of a TB diagnosis. It is more sensitive than microscopy. Its sensitivity varies from 80% to 85% while the sensitivity varies from 50% to 80% for microscopy. It is performed on a solid, enriched culture medium (Löwenstein-Jensen medium or Colestos medium) where results are obtained within 4 to 8 weeks; or on a liquid medium (Middlebrook 7H11, 7H10, 7H9, MGIT/Mycobacteria Growth Indicator Tube or Dubos) where results are obtained more rapidly (approximately 15 days) [32].

## 2. Principles of anti-tuberculosis treatment

A tubercular lesion has 2 different bacillary populations:

- A very rich population, present in caverns, whose rapid multiplication is responsible for the development of colonies immediately resistant to each of the antibiotics, which prohibits monotherapy and justifies an initial phase of intensive treatment based on the simultaneous administration of several antibiotics [32].

- A slower multiplying population, present in solid caseous sites and in macrophages, less accessible to antibiotics and which can be the cause of relapses. The eradication of these bacilli requires a consolidation phase prolonged over several months [32].

## 3. Anti-tuberculosis drugs

- Four anti-tuberculosis drugs are used in first line [46]:

- **Isoniazid** and **Rifampicin**, known as major anti-tuberculosis drugs because they have the following properties: they are bactericidal, their good diffusion allows them to reach to reach intra- and extracellular bacilli. the simultaneous administration of these 2 antibiotics allows a

rapid reduction in the number of of extracellular BK and therefore a rapid negativation of expectorations.

- **Pyrazinamide**, which is effective on intracellular bacilli, can shorten the the duration of the treatment

- **Ethambutol**, a bacteriostatic

In addition to these 4 essential drugs, streptomycin can, in some cases, replace ethambutol.

## 4. Treatment regimens

The treatment regimen varies depending on whether or not the patient has been previously treated but in all cases, it includes two distinct phases [47].

- An initial or intensive phase that lasts two months and serves to rapidly destroy the *M.tuberculosis* bacilli, to prevent the appearance of resistant bacilli and to eliminate contagiousness

- A maintenance phase: of variable duration depending on the clinical situation. This phase serves to sterilize the lesions and thus prevent relapses. In practice, there are three clinical situations requiring different therapeutic schemes [48]:
  - new cases
  - cases of re-treatment
  - multidrug-resistant tuberculosis

### 4.1. Treatment regimens for new cases of pulmonary tuberculosis

- An initial intensive phase combining 4 anti-tuberculosis drugs: isoniazid (H), rifampicin (R), pyrazinamide (Z) and ethambutol (E), for 2 months (2HRZE)

- a consolidation phase combining the 2 major anti-tuberculosis drugs for 4 months: 4HR

The initial phase should be extended by one month if the direct examination is still positive at 2 months and if the antibiogram result is not yet available [49].

### 4.2. Treatment regimens for previously treated patients

It is also recommended, whenever possible, to perform a culture and an antibiogram at the beginning of the treatment in order to detect early any resistance, especially multi-resistance, which is 5 times more frequent in cases of retreatment than in new cases (15% vs. 3%), antibiotic treatment must be strictly adapted to the sensitivity of the bacilli [21].

### 4.2.1. In case of low risk of multi-resistance

The risk is considered low when the patient is relapsing or re-treating after discontinuation. In this situation, the recommended treatment consists of 5 drugs with the addition of streptomycin to HRZE for 2 months, then 4 drugs for 1 month and 3 drugs for 5 months (2SHRZE/1HRZE/5HRE). The treatment will be secondarily adapted to the antibiogram data [50].

### 4.2.2. In case of high risk of multi-resistance

These are cases of failure of a previous treatment and cases of exposure to a patient with multidrug-resistant tuberculosis. In these situations, it is recommended to prescribe a multidrug-resistant tuberculosis treatment regimen while waiting for the results of the antibiogram[48], [51].

## IV.    Dihydrofolate reductase

The enzyme, dihydrofolate reductase (DHFR), catalyzes NADPH-dependent reduction of dihydrofolate to tetrahydrofolate, which is a precursor of cofactors necessary for the synthesis of thymidylate, purine nucleotides, methionine, serine, and glycine that are required for DNA, RNA, and protein synthesis Due to its critical role in regulating cellular THF, DHFR is a well-preserved regulatory enzyme in both eukaryotic and prokaryotic cells [52].
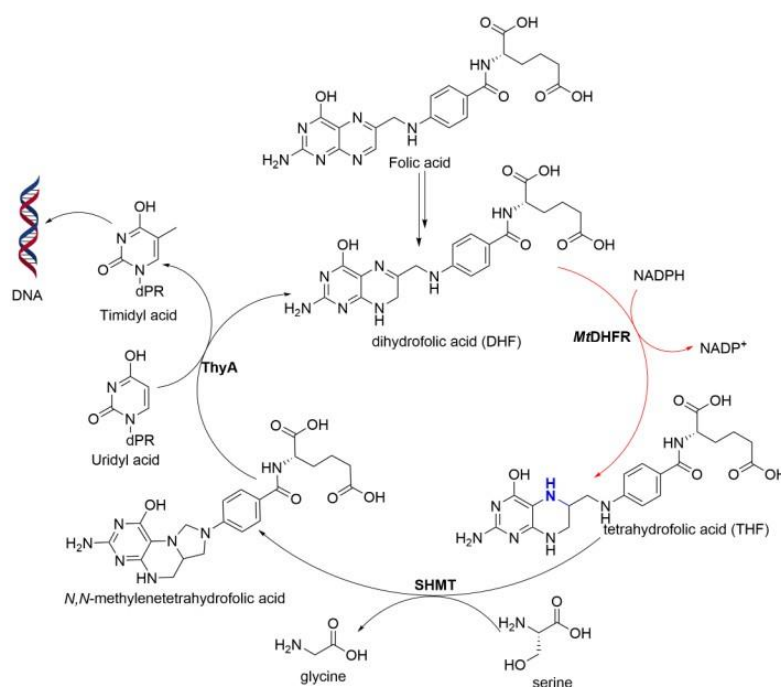


**Figure 5: Simplified scheme of the metabolic pathway of folic acid [51].**

Enzymes in bold: MtDHFR: *Mycobacterium tuberculosis* Dihydrofolate Reductase; SHMT: serine hydroxymethyltransferase; ThyA, thymidylate synthase.

### 1.  Mt-DHFR's structure

In structural terms, Mt-DHFR [PDB ID: 1DF7] contains 159 amino acid residues, and its secondary structure is formed by a central β-sheet flanked by four α-helices. This β-sheet is formed by seven parallel strands and a C-terminal antiparallel strand Mt-DHFR has 26% similarity to the amino acid sequence of h-DHFR [PDB ID: 1OHJ], and crystallographic studies of the two enzymes also demonstrated a similarity between overall protein folds [53].

The structural comparison of mtDHFR and hDHFR reveals striking differences in the NADPH, MTX, and PT523-binding sites, respectively. In particular, in the ternary complex of MTX with mtDHFR, a glycerol molecule (glycerol A) is found close to MTX in a pocket formed by Trp22, Asp27, and Gln28 while in hDHFR, the glycerol-binding site is packed with three hydrophobic residue side chains, Leu22, Pro26, and Phe31. These differences can be exploited for the design of a selective inhibitor for mtDHFR [54].
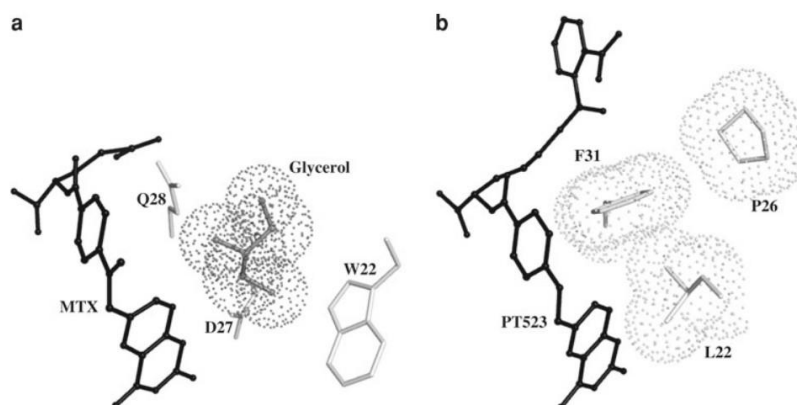


**figure 6: Difference between binding sites of MTX and PT523 in mtDHFR and hDHFR, respectively [54].**

MTX and PT523 are shown as ball-and-stick rendering colored in black. Glycerol is shown as ball-and-stick rendering in grey surrounded by grey dots while side-chains of interacting residues are shown as stick rendering in grey.

## 2. Mt-DHFR's domains

In Mt-DHFR's functions as a monomeric enzyme, and contains an adenosine binding domain and loop domain, The secondary structure of DHFR is comprised of an eight stranded β-sheet, with strands labeled βA – βH, four α-helices (B, C, E, F), and four short mobile loops.[55, p. 1] The adenosine binding domain consists of β strands B, C , D , and E , as well as helices C , E and F . The loop domain contains three loops: the L1 loop , L4 loop , and L5 loop [56].
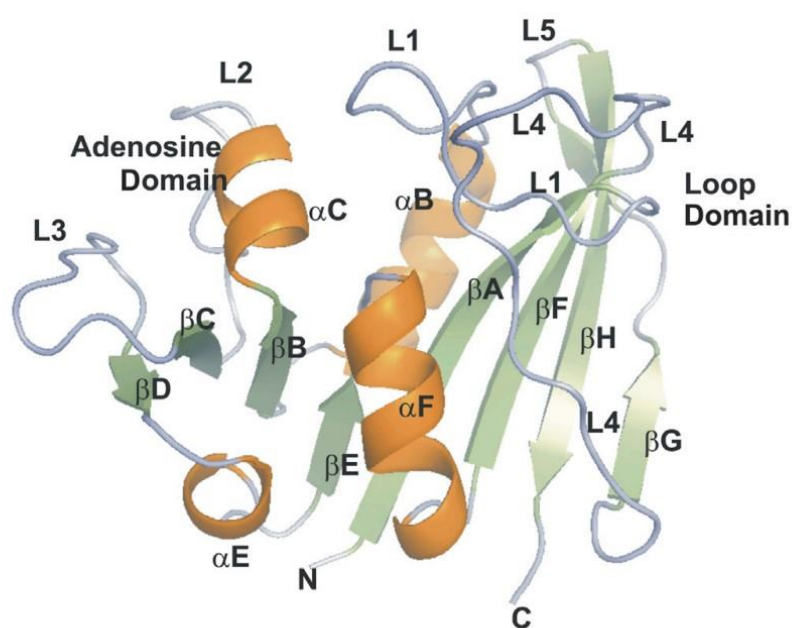


**figure 7: Structure of Mt-DHFR with the Elements of Secondary Structure Labeled and the organization into two domains [56].**

The adenosine binding domain binds the cofactor NADPH, and the loop domain contains the substrate, DHF, and these two domains facilitate hydride transfer from C-4 of NADPH to C-6 of DHF and protonation of N-5 of DHF [57].
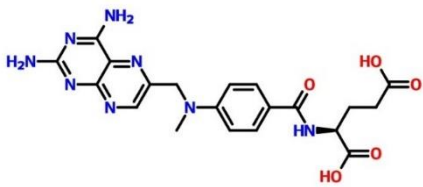
## 3. Inhibitors

Mt-DHFR was studied in order to increase the ability to treat tuberculosis patients is the more relevant, since DHFR inhibitors have been shown to be of use in Cote d'Ivoire in the treatment of patients infected both by HIV and M. tuberculosis [53].

### 3.1. Methotrexate

Methotrexate (MTX) is a DHFR inhibitor that binds to both human DHFR (h-DHFR) and mt-DHFR without any significant selectivity. This inhibition is the basis for the use of methotrexate (MTX) as a cancer chemotherapy drug and it has been widely used in the past for the treatment of a variety of malignancies, including breast, head and neck, leukaemia, lymphoma, lung, osteosarcoma, bladder and trophoblastic neoplasms [58].
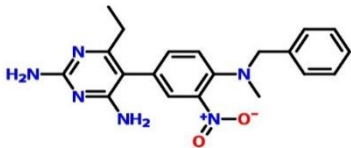
**Table IV: Methotrexate properties (PubChem CID 126941).**

| Structure | |
|---|---|
| |  |
| Molecular Formula | $C_{20}H_{22}N_8O_5$ |
| Weight | 454.4 g/mol |
| IUPAC Name | (2S)-2-[[4-[(2,4-diaminopteridin-6-yl)methyl-methylamino]benzoyl]amino]pentanedioic acid |

### 3.2. Methylbenzoprim

Methylbenzoprim is one of the most potent known inhibitors of mammalian (rat liver) DHFR (Ki 0.009 nM), possibly because it is competitive with NADPH as well as dihydrofolate. Among a large series of related structures submitted to the NCI Developmental Therapeutics Program for antitumor evaluation [59].
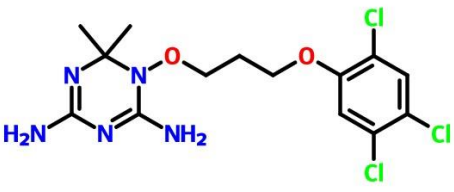
**Table V: Methylbenzoprim properties (PubChem CID 72438).**

| Structure | |
|---|---|
| |  |
| Molecular Formula | $C_{20}H_{22}N_6O_2$ |
| Weight | 378.4 g/mol |
| IUPAC Name | 5-[4-[benzyl(methyl)amino]-3-nitrophenyl]-6-ethylpyrimidine-2,4-diamine |

### 3.3. WR99210

WR99210 is a folate pathway antagonist with potent activity against Plasmodium malaria parasites . Action of WR99210, originally designated as BRL 6231, in parasites is against the bifunctional dihydrofolate reductase thymidylate synthase enzyme (DHFR-TS), where the compound binds to the DHFR active site and blocks the production of tetrahydrofolate [60].
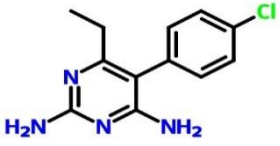
**Table VI: WR99210 properties (PubChem CID 121750).**

| Structure | |
|---|---|
| |  |
| Molecular Formula | $C_{14}H_{18}Cl_3N_5O_2$ |
| Weight | 394.7 g/mol |
| IUPAC Name | 6,6-dimethyl-1-[3-(2,4,5-trichlorophenoxy)propoxy]-1,3,5-triazine-2,4-diamine |

### 3.4. Pyrimethamine

An antimalarial drug that targeting plasmodium dihydrofolate reductase (pDHFR), has been proved to have antitumor activity, In addition to its antimalarial effects, PYR exhibits the activity of inducing apoptosis of tumor cells [61].

**Table VII : : Pyrimethamine properties (PubChem CID 4993).**

| Structure | |
|---|---|
| |  |
| Molecular Formula | $C_{12}H_{13}ClN_4$ |
| Weight | 248.71 g/mol |
| IUPAC Name | 5-(4-chlorophenyl)-6-ethylpyrimidine-2,4-diamine |

### 3.5. Trimethoprim

Is a well-known dihydrofolate reductase inhibitor and one of the standard antibiotics used in urinary tract infections[62].

**Table VIII : Trimethoprim properties (PubChem CID 5578).**

| Structure | |
|---|---|
| |  |
| Molecular Formula | $C_{14}H_{18}N_4O_3$ |
| Weight | 290.32 g/mol |
| IUPAC Name | 5-[(3,4,5-trimethoxyphenyl)methyl]pyrimidine-2,4-diamine |

## V. Computer-Aided Drug Design (CADD)

The drug discovery and development process is essentially a patient-centered science, where researchers strive to improve existing drugs or invent a completely new chemical molecule, which should ideally be more potent than any existing drug in a similar class. If not, then at least it should be safer than existing ones. This process is a very time-consuming and expensive activity that requires the expertise of many eminent researchers. It takes nearly 12-14 years of exhaustive research and a huge financial investment for the discovery of a single drug. From chemical synthesis to clinical development and finally formulating it into an appropriate form. Failure at any stage would mean a huge loss for the company. Therefore, a lot of planning is required before the project is even underway. Recently, with the use of technology, the process is becoming a less risky activity, due to the ability of computers to predict possible outcomes [9].

CADD methods are mathematical tools for manipulating and quantifying the properties of potential drug candidates as implemented in a number of programs. These include a range of publicly and commercially available software [63].

The general approaches to computer-aided drug design (CADD) that exist to date are structure-based drug design (SBDD) and ligand-based drug design (LBDD) [63].
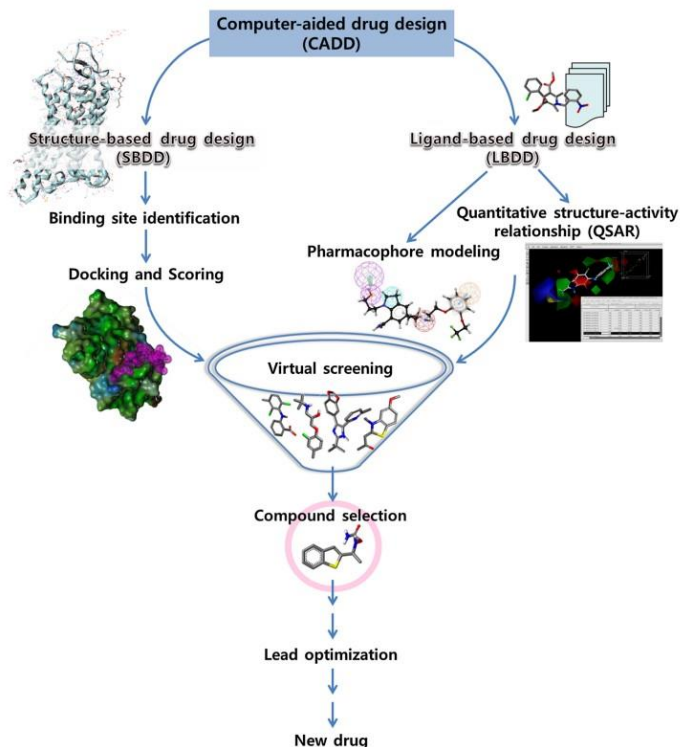


**Figure 8 : Representative workflow for computer-aided drug design [62].**

## 1. Structure-Based Drug Design (SBDD)

To design a drug to treat a disease or alleviate a symptom, a clear understanding of the disease pathway and relevant processes is crucial for the selection of a therapeutic target [9].

The so-called Structure Based Drug Design or SBDD approach, is a relatively straightforward approach, based on the knowledge and appropriate analysis of the three-dimensional structural data of the macromolecular target with which the small molecular chemical ligands interact, in order to determine the key sites and interactions that are important for their biological functions[64]. These 3D structures are obtained through the use of X-ray crystallography, nuclear magnetic resonance (NMR), cryo-electron microscopy (EM), homology modeling and molecular dynamics (MD) simulations [65].

## 1.1. Protein structure determination

To determine the three-dimensional structure of a protein, many ways are used, the most common one is by X-ray crystallography and nuclear magnetic resonance spectroscopy [66].

## 1.2. Homology modeling

It is a key technique within molecular modelling. It consists in the prediction of unknown 3D structure of the target protein, using as a model, a known 3D structure of another homologous protein. The idea behind this technique is the fact that evolutionary proteins often share similar structures. Knowing structures that have amino acid sequences similar to the target sequence of interest can help predict the target structure, its function, and even binding sites possible functionalities of the structure [67], [68].

## 1.3. Fold recognition (Threading)

Fold recognition is based on the fact that, during the evolution of proteins, a considerable sequence divergence has been observed, but at the folds, only slight changes are observed[69]. Proteins may not have similar sequences, but they can have similar folds. The sequence of a known protein structure is replaced by the sequence of the target of interest whose structure is not known. The new "threaded" structure is then evaluated using different scoring methods. This process is repeated for all the experimentally determined 3D structures in a database until the structure that has the best fit for the desired sequence is obtained [70].

## 1.4. Ab initio (de novo) modeling

Ab initio or de novo modeling is used when there is no sufficiently homologous structure to use comparative modeling. De novo modeling is not based on a model structure; it models the structure of the target by sequence function only. The ab initio modeling implemented in Rosetta is a technique which uses a "knowledge-based scoring function "to guide a fragment-based Monte Carlo search in conformational space. This method generates a protein-like structure with centroid atoms to represent side chains. In order to refine this centroid-based structure, an "all-atom refinement" function is used to relax the structure [71].

## 1.5. Binding site identification

The binding site is a small region, a pocket of bumps, where ligand molecules can best fit or bind to activate the receptor or target and produce the desirable effect. Thus, recognizing the binding site or the active site residues in the target structure is of high importance in SBDD. Because the proteins are capable of undergoing conformational changes, recognizing the accurate binding site residues is difficult [72], But still, there are just a few computational programs that can capably spot out the

binding site residues. With the targets and their binding site having been defined, the next crucial step in SBDD is hit discovery, which probably results in a library of compounds that can interact with the target [73].

## 1.6. Docking and scoring functions

Docking, a process of predicting the ligand conformation and its orientation inside the target structure plays a vital part in SBDD. The interaction or fit between the ligand and the protein structure is best represented as the "hand and glove" model [74]. Docking is often carried out in two parts. The first part includes the effective search of conformational space through a "posing" mechanism where the ligand is placed inside the receptor in different orientations to facilitate the identification of the actual binding mode of the ligand molecules. An energy-based score is provided for each pose in terms of their interactions with the receptor. The scoring functions are features that aid in investigating the interactions between the small molecule and the biological target, thereby providing context about biological activity. The second part of the docking process is the ranking of poses based on their computed scores [75].

## 2. Ligand-Based Drug Design (LBDD)

Is a rather indirect approach where the analysis depends entirely on information about other molecules that bind to the target compound of interest. These other molecules can be used to develop a pharmacophore model that delineates the minimal structural features that a molecule must possess to successfully bind to the target [76].

The compound set should encompass a wide range of concentration (at least 4 orders of magnitude) to generate a reliable ligand-based screening model. Common ligand-based design techniques are quantitative structure–activity relationships (QSARs) and pharmacophore-based methods [76].

## 2.1. Similarity searches

This approach is based on the selection of new components, which have chemical and physical similarity with known drugs. The principle is simple and effective, based on the fact that structurally similar molecules tend to have similar binding properties. However, these similarities do not give information about known ligands activity of a target [63].

## 2.2. Quantitative Structure-Activity Relationships (QSAR)

Its studies are based on the premise that changes in bioactivity are associated with structural and molecular variations in a set of compounds [76], [77]. A statistical model is generated from this correlation to develop and mathematically predict the biological property of novel compounds [76].

## 2.3. Pharmacophore modeling

It aims to identify compounds containing different scaffolds, but with a similar 3D arrangement of key interacting functional groups [78]. Binding site information can be incorporated into the pharmacophore model by exploiting the bioactive conformation of candidate compounds. Pharmacophore modeling is also often performed in the molecular alignment stage of QSAR modeling studies [76].

## 2.4. Assessment of ADME and toxicity

Lead compounds in drug discovery need to be optimized for both efficacy and safety. Its optimization consists of obtaining the desired pharmacological profiles to reach the required affinity, pharmacokinetic properties, drug safety, and ADME properties. The free binding energy reflects the potency of the drug against the target of interest, and its calculation is done by dynamic molecular simulations. The aforementioned process is done gradually by converting one atom of the ligand to another and measuring the affinity [79].

# MATERIALS AND METHODS

## I. Materials

In order to propose potential mtDHFR inhibitors, structure-based virtual screens were performed. Many inhibitors were newly proposed against our target.

### 1. Databases

#### 1.1. Protein Data Bank (PDB)

Protein Data Bank (PDB) is a worldwide database containing 3D structure data of biological macromolecules, such as proteins and nucleic acids. The structures are usually obtained by X-ray crystallography, NMR spectroscopy, or cryo-electron microscopy, and submitted by biologists and biochemists from all over the world, they are freely available [66].

#### 1.2. PubChem

The PubChem database was used to retrieve the 3D structures of the compounds used in this study. It contains mainly small molecules, but also larger molecules such as nucleotides, carbohydrates, lipids, peptides and chemically modified macromolecules. PubChem provides information on chemical structures, identifiers, chemical and physical properties, biological activities, health, safety, toxicity data and much more [80].

#### 1.3. Mcule

Mcule is an online drug discovery platform that provides molecular modeling tools for hit identification, hit expansion, toxicity checker and lead optimization [81].

#### 1.4. Selleckchem

Selleck Chemicals supplies a big library of inhibitors used in the study of cell signaling pathways, it is a platform that provides information on chemical structures, identifiers, chemical and physical properties, biological activities, health, safety, toxicity data and much more[82].

#### 1.5. Binding Database

BindingDB is a public, web-accessible database of measured binding affinities, focusing chiefly on the interactions of protein considered to be drug-targets with small, drug-like molecules. BindingDB contains 41,296 Entries, each with a DOI, containing 2,519,702 binding data for 8,810 protein targets and 1,080,101 small molecules [83].

## 2. Software

### 2.1. ADT

AutoDockTools, or ADT, is the free GUI for AutoDock developed by the same laboratory that develops AutoDock. It is used to set up, run and analyze AutoDock dockings and isocontour AutoGrid affinity maps, as well as compute molecular surfaces, display secondary structure ribbons, compute hydrogen-bonds, and do many [84].

### 2.2. Autodock Vina

Auto Dock Vina is a new generation of docking software from Molecular Graphics Lab. It significantly improves the average accuracy of docking mode predictions, while being up to two orders of magnitude faster than Auto Dock4 [84].

### 2.3. PyMOL

PyMOL is an open source molecular visualization system, capable of producing high quality 3D images of small molecules and biological macromolecules, such as proteins [85].

### 2.4. Open Babel

Open Babel is a computer software, a chemistry expert system mainly used to convert chemical file formats, it is designed to support molecular modeling, chemistry and many related fields [86].

### 2.5. Toxicity checker-Mcule

It is an online drug discovery platform to estimate the hazard of chemical compounds. It offers a unique solution to pharmaceutical and biotechnology companies by providing the highest quality database of purchasable compounds and molecular modeling tools [87].

## VI.    Methods

### 1.  Target selection and preparation

The crystallographic structure of our protein of interest (*Mycobacterium tuberculosis* dihydrofolate reductase) was obtained in ".pdb" format from the Protein Data Bank (PDB) database (https://www.rcsb.org/structure/1DF7) under PDB: 1DF7. The three-dimensional structure of DHFR was determined by X-ray crystallography with a resolution of 1.70 Å;

The different 3D conformations are downloadable under pdb extension readable by docking software. Information on the primary structure, the heteroatoms (ligand, metals, modified residues... etc.), the

secondary structure and the X, Y, Z atomic coordinates which determine the exact position of each atom in a conformation are available.

The 3D structure of DHFR was visualized with PyMol which allows to rotate and resize in real time. This visualization showed us that DHFR is a monomer formed of one single chain consisting of 159 amino acid residues, complexed with NADPH and methotrexate linked by hydrogen bonds to the protein chain of the enzyme forming its active site that we removed with this software.

In addition, there is glycerol molecule (GOL) and several water molecules. These have been eliminated thanks to the graphical interface Auto Dock Tools (ADT) which also allowed us to assign partial atomic charges to our macromolecule (Kollman United Atom charges). Then the file was saved in " pdbqt " format.
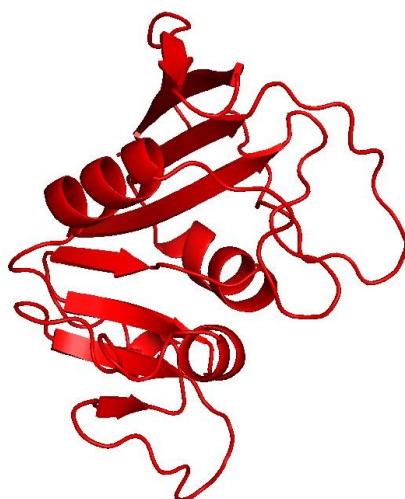


**Figure 9 : structure of mt-DHFR after preparation as seen in PyMOL**

## 2. Construction of ligand databases

Virtual screening projects aimed at identifying novel, selective and potent inhibitors of a target protein typically use large-scale compound libraries containing several thousand, or even millions, of small molecules to initiate the screening process. Depending on the purpose and type of study.

We searched for ligands that have structural complementarity with the active site of the target by searching the protein file in the format 'DHFR.pdb.qt' in the selleckchem platform (https://selleckchem.com). A library of 8213 molecules was collected, then we selected 11 molecules which present an affinity of at least -7 Kcal/mol and which respect Lipinski's **rule of five (RO5)** also

31

known as **Pfizer's rule of five**. It is a rule of thumb to evaluate drug likeness or determine if a chemical compound with a certain pharmacological or biological activity has chemical properties and physical properties that would make it a likely orally active drug in humans. The rule describes molecular properties important for a drug's pharmacokinetics in the human body, including their absorption, distribution, metabolism, and excretion (ADME). However the rule does not predict if a compound is pharmacologically active [88], [89].

Lipinski's rules state that, in general, an orally active drug has no more than one violation of the following criteria [88]:

- o more than 5 hydrogen bond donors (the total number of nitrogen-hydrogen and oxygen-hydrogen bonds)

- No more than 10 hydrogen bond acceptors (all nitrogen or oxygen atoms)

- A molecular mass less than 500 daltons

- An octanol-water partition coefficient (logP) that does not exceed 5

Based on the aforementioned rule, a number of compounds have been selected for molecular docking with the mt-DHFR.

PubMed literature concerning mt-DHFR inhibitors was selected using the query "Mt-dhfr and Inhibitors". After reading, we identified five approved drugs active against mt-HDFR, namely Bromo_WR99210, Methylbenzoprim, Methotrexate, Pyrimethamine, Trimethoprim, and also the dihydrofolic acid as a positive control. From each of the aforementioned molecules, we searched for their 3D structures using PubChem database [80].

## 3. Toxicity assessment

The prediction of the toxicity of compounds is an important part of the process of drug development and design. The toxicity assessment of the molecules used in our study was performed using Mcule Toxicity Checker. Based on this evaluation, 11 molecules from the 8412 molecules obtained from selleckchem were selected as potential DHFR inhibitors [87].
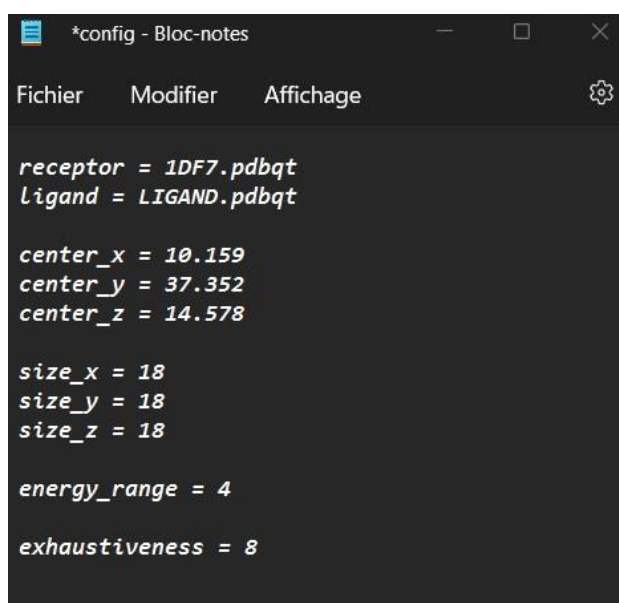
## 4. Ligand preparation

The small molecules (ligands) were downloaded from PubChem and selleckchem Databases in a sdf file format, and then converted into a pdb file format using Open Babel software. Formatted ligand files for AutoDock Vina must be in pdbqt format and contain atom types supported by AutoDock plus extra records that specify rotatable bonds.

As positive controls we used the dihydrofolic acid the natural ligand for the protein.

## 5. Configuration file preparation

Docking of the previously prepared ligands with the target was performed by the Auto Dock Vina software. The latter requires an input configuration file that contains all the parameters used to set up the docking, including the protein name and ligand[90]. Vina has a parameter called 'exhaustiveness' that controls the duration of its search. Another parameter is the energy range which represents the energy difference between the highest and lowest score, by default it is set to 4.



**Figure 10 : The configuration file config.txt**

From "grid.txt" file, we have written the center_x, y, and z coordinates, and also the size_x, y, and z of the grid box. Save this file as "config.txt".

## 6. Molecular docking

After placing the following files in the same folder:

- 1DF7.pdbqt

- Ligand.pdbt

- Config.txt

- All the MGL_Tools, Autodock Tools, and Autodock Vina setup files.

We open the command prompt and enter the folder where all the docking files are placed. And then we type the following command:

```
"C : \ Program Files (x86) \ The Scripps Research Institute \ Vina \ vina.exe" --receptor
1DF7.pdbqt --ligand ligand. Pdbqt --config config.txt --log log.txt --out output. Pdbqt
```

NB: The word **ligand** in the command is modified with the exact name of each ligand set for every docking run.

## VII.  Global workflow used to identify inhibitors of mt-DHFR

The global workflow representing major steps, databases and software used in the current study to identify potential inhibitors mt-DHFR are summarized.



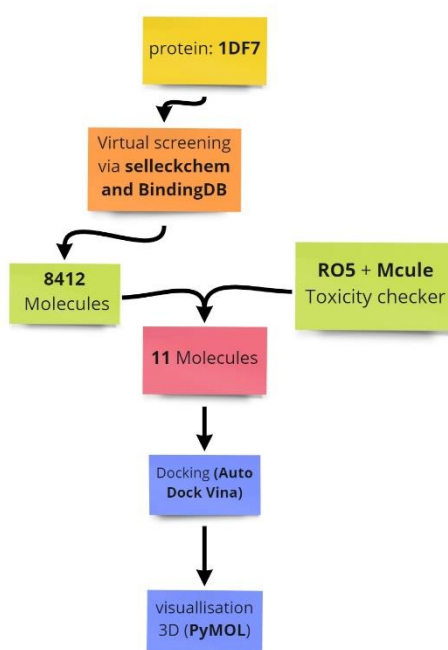**Figure 11 Major steps carried out to
identify potential inhibitors of DHFR**

# RESULTS AND DISCUSSION

# I. Results

## 1. Generation dataset

5 commonly used inhibitor of DHFR were also tested against the proteins to compare their affinity to the proteins and against the 11 resulted of the SBVS, as positive control, we used dihydrofolic acid as it is the physiological substrate of the DHFR (**Table IX**).

**Table IX: All selected inhibitors, dihydrofolic acid and their IUPAC identifier.**

| Molecule | IUPAC |
|---|---|
| **BDBM18226** | 5-[3-[(2,4-diamino-5-methylpyrido[2,3-d]pyrimidin-6-yl)methyl]-4-methoxyphenoxy]pentanoic acid |
| **BDBM50145795** | 3-[11-[(2,4-diaminopteridin-6-yl)methyl]benzo[b][1]benzazepin-3-yl]oxypropanoic acid |
| **BDBM18225** | 4-[3-[(2,4-diamino-5-methylpyrido[2,3-d]pyrimidin-6-yl)methyl]-4-methoxyphenoxy]butanoic acid |
| **BDBM18227** | 6-[3-[(2,4-diamino-5-methylpyrido[2,3-d]pyrimidin-6-yl)methyl]-4-methoxyphenoxy]hexanoic acid |
| **BDBM18228** | 4-[2-[(2,4-diamino-5-methylpyrido[2,3-d]pyrimidin-6-yl)methyl]-4-methoxyphenoxy]butanoic acid |
| **BDBM18234** | 4-[11-[(2,4-diaminopteridin-6-yl)methyl]benzo[b][1]benzazepin-3-yl]but-3-ynoic acid |
| **BDBM18235** | 5-[11-[(2,4-diaminopteridin-6-yl)methyl]benzo[b][1]benzazepin-3-yl]pent-4-ynoic acid |
| **BDBM18236** | 6-[11-[(2,4-diaminopteridin-6-yl)methyl]benzo[b][1]benzazepin-3-yl]hex-5-ynoic acid |
| **BDBM50145798** | 4-[11-[(2,4-diaminopteridin-6-yl)methyl]benzo[b][1]benzazepin-3-yl]oxybutanoic acid |
| **BDBM50145799** | 5-[11-[(2,4-diaminopteridin-6-yl)methyl]benzo[b][1]benzazepin-3-yl]oxypentanoic acid |
| **BDBM50514994** | 1-N-[2-(2-amino-4-oxo-3,7-dihydropyrrolo[2,3-d]pyrimidin-6-yl)ethyl]-4-N-(pyridin-2-ylmethyl)benzene-1,4-dicarboxamide |
| **Bromo_WR99210** | 1-[3-(4-bromophenoxy)propoxy]-6,6-dimethyl-1,3,5-triazine-2,4-diamine |
| **Methylbenzoprim** | 5-[4-[benzyl(methyl)amino]-3-nitrophenyl]-6-ethylpyrimidine-2,4-diamine |
| **Methotrexate** | (2$S$)-2-[[4-[(2,4-diaminopteridin-6-yl)methyl-methylamino]benzoyl]amino]pentanedioic acid |
| **Pyrimethamine** | 5-(4-chlorophenyl)-6-ethylpyrimidine-2,4-diamine |
| **Trimethoprim** | 5-[(3,4,5-trimethoxyphenyl)methyl]pyrimidine-2,4-diamine |
| **Dihydrofolic acid** | (2$S$)-2-[[4-[(2-amino-4-oxo-7,8-dihydro-3$H$-pteridin-6-yl)methylamino]benzoyl]amino]pentanedioic acid |

## 2. Molecular docking results

Structure-based virtual screening approach was used to identify small molecules with high affinity to Mt-DHFR. For this purpose, we used selleckchem platform to obtain 8000+ compounds. From them we retrieved those with the best profile, based on Lipinski's rules and the toxicity assessment. Then, we performed docking using Autodock Vina to confirm and to validate the obtained results. Based on the affinity, 11 molecules with the best binding affinity to Mt-DHFR ranging from -9,6 kcal/mol to -6.9 kcal/mol were selected. The ligand 1 and 2 (BDBM18226 and BDBM50145795) were the most important ones since they showed a greater affinity than the rest of the inhibitors with best affinity (-9,6 kcal/mol) for Mt-DHFR and (-9,2 and -9 kcal/mol respectively) for the Mt-DHFR and finally a (-7,3 kcal/mol) for h-DHFR, the affinity of the molecules, the number and the distance of the hydrogen bonds are with each protein are shown in the **(Table X)**, **(Table XI)** and **(Table XII)**.

**Table X: Docking results of Mt-DHFR and selected molecules.**

| Ligand number | Molecule | Affinity (kcal/mol) | H bonds | Distance |
|:---:|:---:|:---:|:---:|:---|
| 1 | **BDBM18226** | **-9,2** | **7** | ARG-32 2.0Å<br>ARG-60 2.3-2.5-1.8Å<br>SER-49 2.5-2.1Å<br>ILE-14 2.9Å |
| 2 | **BDBM50145795** | **-9** | **5** | ARG-60 2.0-2.3Å<br>ARG-32 2.2Å<br>ILE-94 2.8Å<br>THR-46 2.8Å |
| 3 | **BDBM18225** | **-8,7** | **9** | ILE-5 2.4Å<br>TYR-100 2.4Å<br>ILE-94 2.4Å<br>ASP-27 2.1-2.9Å<br>ARG-60 2.0-2.1-2.5Å<br>ARG-32 2.2Å |
| 4 | **BDBM18227** | **-9,3** | **8** | ILE-14 2.5Å<br>GLY-18 2.3Å<br>SER-49 2.1-2.5Å<br>ARG-60 1.9-2.5-2.0Å<br>ARG-32 2.0Å |
| 5 | **BDBM18228** | **-8,5** | **7** | SER-49 2.8Å<br>ASP-27 2.1-2.8 Å<br>GLN-28 2.3Å<br>ILE-5 2.7Å<br>TYR-100 2.3Å<br>ILE-94 2.3Å |
| 6 | **BDBM18234** | **-9,2** | **5** | SER-49 2.6-2.4Å<br>GLY-18 2.7Å<br>ARG-60 2.1Å<br>ARG-32 2.5Å |
| 7 | **BDBM18235** | **-8,9** | **3** | ARG- 2.5Å<br>SER- 2.4Å<br>GLY- 2.8Å |
| 8 | **BDBM18236** | **-8,6** | **5** | ALA-7 2.3Å<br>TYR-100 2.4Å<br>ILE-14 2.2Å<br>GLN-28 2.7Å<br>TRP-22 2.1Å |
| 9 | **BDBM50145798** | **-8,8** | **4** | TYR-100 2.4Å<br>ILE-14 2.3Å<br>GLN-28 2.5Å<br>ILE-20 2.0Å |
| 10 | **BDBM50145799** | **-6,9** | **5** | ARG-60 2.5-1.8Å<br>ILE-5 2.5Å<br>TYR-100 2.1Å<br>ASP-27 2.1Å |
| 11 | **BDBM50514994** | **-9** | **4** | SER-49 2.3Å<br>GLN-28 2.5Å<br>ASP-28 2.4Å<br>TYR-100 2.5Å |
| 12 | **Bromo_WR99210** | **-7,3** | **5** | ASP-27 2.1-2.8Å<br>ILE-5 1.9Å<br>TYR-100 2.6Å<br>ILE-94 2.1Å |
| 13 | **Methylbenzoprim** | **-7,8** | **4** | GLN-28 2.4Å<br>ASP-27 2.6Å<br>ALA-7 2.6Å<br>TYR-100 2.6Å |
| 14 | **Methotrexate** | **-8,7** | **6** | SER-49 2.4-2.3Å<br>LYS-53 2.6Å<br>ARG-32 1.9-2.0Å<br>ARG-60 2.4Å |
| 15 | **Pyrimethamine** | **-7** | **4** | ASP-27 2.2Å<br>ILE-5 1.9Å<br>TYR-100 2.8Å<br>ILE-94 2.5Å |
| 16 | **Trimethoprim** | **-7,1** | **4** | ILE-5 1.9Å<br>ILE-94 2.2Å<br>SER-49 2.5-2.4Å |
| 17 | **Dihydrofolic acid** | **-8,8** | **4** | SER-49 2.8Å<br>ARG-32 2.1Å<br>ARG-60 2.3-2.1Å |

**Table XI: Docking results of Mt-DHFR+GOL and selected molecules.**

| Ligand number | Molecule | Affinity (kcal/mol) | H bonds | Distance |
|---|---|---|---|---|
| 1 | BDBM18226 | -9,6 | 7 | SER-49 2.4-2.1Å<br>GLY-18 2.3Å<br>ARG-32 1.9Å<br>ARG-60 2.2-2.6-1.8Å |
| 2 | BDBM50145795 | -9,6 | 6 | ILE-20 2.8-1.9Å<br>GOL-502 2.5Å<br>GLN-28 2.8Å<br>ALA-7 2.1-2.0Å |
| 3 | BDBM18225 | -9,4 | 7 | ILE-5 2.1Å<br>ILE-94 2.4Å<br>TYR-100 2.5Å<br>ARG-32 2.0Å<br>ARG-60 2.1-2.7-2.0Å |
| 4 | BDBM18227 | -9 | 8 | TYR-100 2.0Å<br>ALA-7 2.0-2.2Å<br>ILE-20 2.7-2.1-2.3Å<br>GOL-502 2.8Å<br>SER-49 2.1Å |
| 5 | BDBM18228 | -8,9 | 7 | ILE-14 2.5<br>TYR-100 2.3<br>ILE-5 2.5<br>ILE-54 2.2<br>ASP-27 2.7<br>GLN-28 2.8<br>GOL-205 2.5 |
| 6 | BDBM18234 | -7,3 | 5 | ALA-7 2.3-2.2<br>ILE-94 2.7<br>ARG-60 2.4<br>AGR-32 2.5 |
| 7 | BDBM18235 | -8,8 | 5 | ILE-20 2.3-2.1Å<br>GLN-28 2.6Å<br>ALA-7 2.4Å<br>TYR-100 2.1Å |
| 8 | BDBM18236 | -9,2 | 5 | GLN-28 2.6Å<br>ILE-20 2.1-1.0Å<br>ALA-7 2.1Å<br>TYR-100 2.6Å |
| 9 | BDBM50145798 | -8 | 8 | ASP-27 2.6-2.3-2.2Å<br>ALA-7 2.6-2.6Å<br>TYR-100 2.2-2.4Å<br>ILE-94 1.9Å |
| 10 | BDBM50145799 | -9,6 | 6 | ILE-20 1.9-2.3Å<br>ILE-94 2.8Å<br>TYR-100 2.6Å<br>GOL-502 2.4Å<br>GLN-28 2.7Å |
| 11 | BDBM50514994 | -8,9 | 5 | ARG-32 2.5Å<br>GLN-28 2.5Å<br>ILE-5 2.5Å<br>TYR-100 2.6-2.4Å |
| 12 | Bromo_WR99210 | -7,4 | 2 | TYR-100 1.9Å<br>ASP-27 2.4Å |
| 13 | Methylbenzoprim | -8,3 | 5 | TYR-100 2.6Å<br>ALA-7 2.6Å<br>ASP-27 2.5Å<br>GLN-28 2.4Å<br>GOL-502 2.6Å |
| 14 | Methotrexate | -8,9 | 6 | TYR-100 2.7Å<br>SER-49 2.3-2.4Å<br>ARG-60 2.4-2.0Å<br>ARG-32 2.2Å |
| 15 | Pyrimethamine | -8 | 2 | TYR-100 2.0Å<br>THR-46 2.4Å |
| 16 | Trimethoprim | -7,5 | 5 | ASP-27 2.0-2.3Å<br>SER-49 1.8Å<br>ILE-5 1.9Å<br>TYR-100 2.5Å |
| 17 | Dihydrofolic acid | -8,8 | 4 | SER-49 2.9Å<br>ARG-32 2.2Å<br>ARG-60 1.8-2.4Å |

**Table XII: Docking results of h-DHFR and selected molecules.**

| Ligand number | molecule | affinity (kcal/mol) | H bonds |
|:---:|:---:|:---:|:---:|
| 1 | BDBM18226 | -7,3 | 4 |
| 2 | BDBM50145795 | -7,3 | 7 |
| 3 | BDBM18225 | -8,2 | 7 |
| 4 | BDBM18227 | -8 | 6 |
| 5 | BDBM18228 | -8,9 | 6 |
| 6 | BDBM18234 | -8 | 4 |
| 7 | BDBM18235 | -6,9 | 2 |
| 8 | BDBM18236 | -8 | 3 |
| 9 | BDBM50145798 | -6,3 | 6 |
| 10 | BDBM50145799 | -8,2 | 3 |
| 11 | BDBM50514994 | -9,9 | 4 |
| 12 | Bromo_WR99210 | -7,6 | 6 |
| 13 | Methylbenzoprim | -7,6 | 3 |
| 14 | Methotrexate | -8,5 | 8 |
| 15 | Pyrimethamine | -7,3 | 3 |
| 16 | Trimethoprim | -6,7 | 5 |
| 17 | Dihydrofolic acid | -9,2 | 7 |
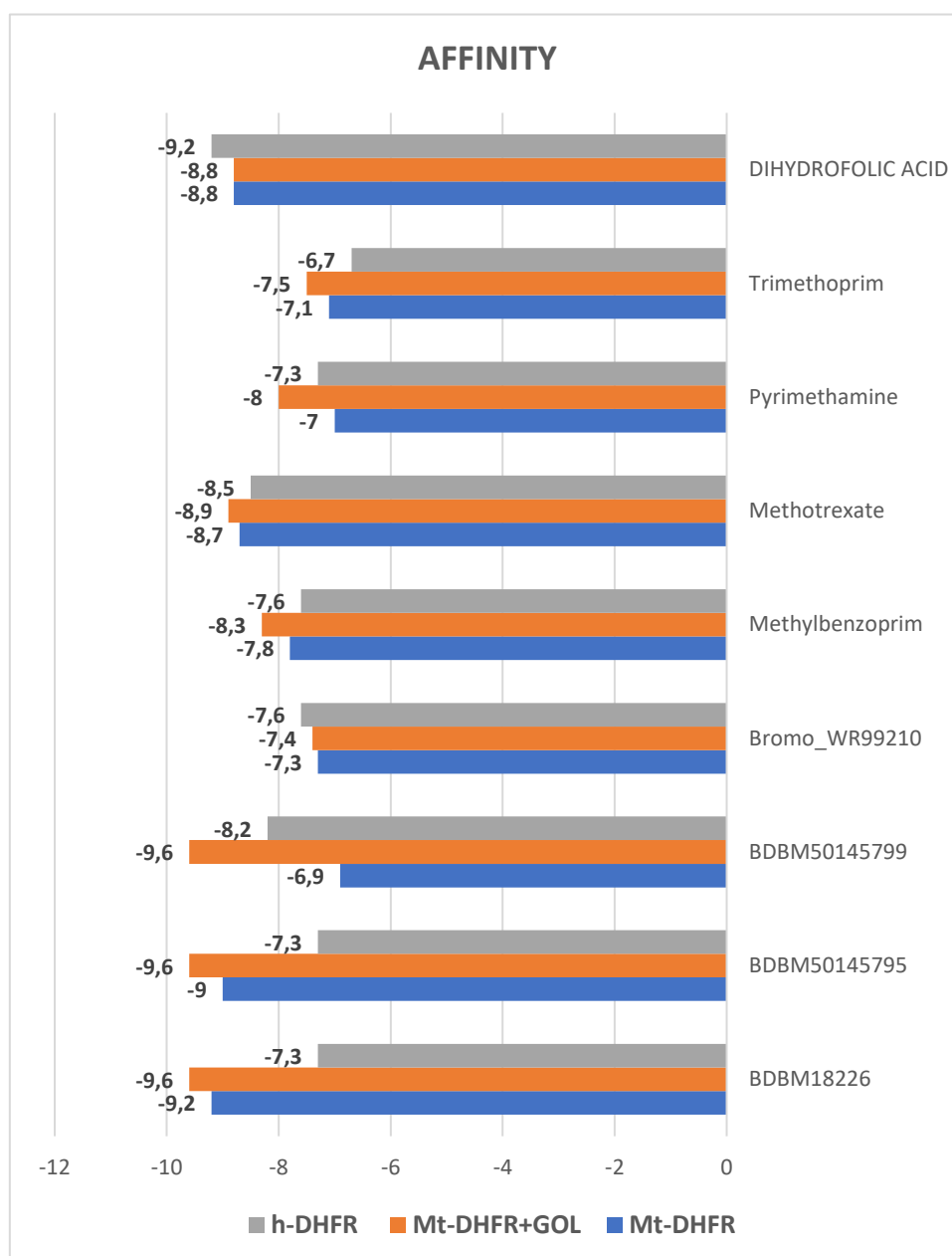
**Figure 12 : Bar graph representing the binding energies (in kcal/mol) calculated by AutoDock Vina software**

ligand BDBM18226, BDBM50145795, BDBM50145799, Bromo_WR99210, Methylbenzoprim, Methotrexate, Pyrimethamine, Trimethoprim and Dihydrofolic acid.
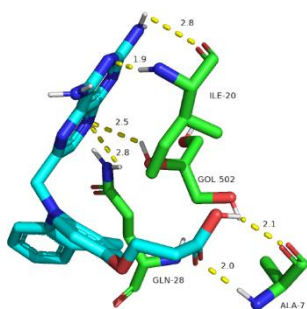
## 3. Visualization of Protein-Ligand interactions
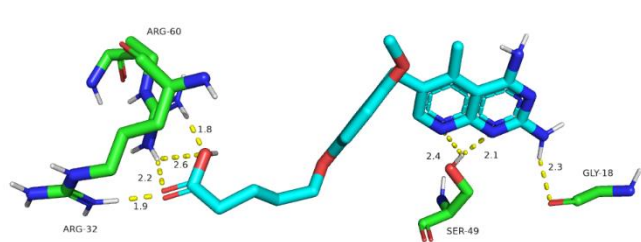


**BDBM18226 (-9,6 kcal/mol)**
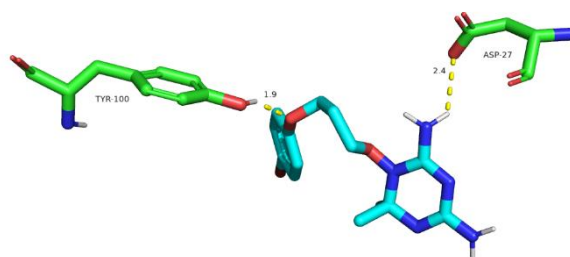
**Bromo_WR99210 (-7,4 kcal/mol)**

**BDBM50145795 (-9,6 kcal/mol)**
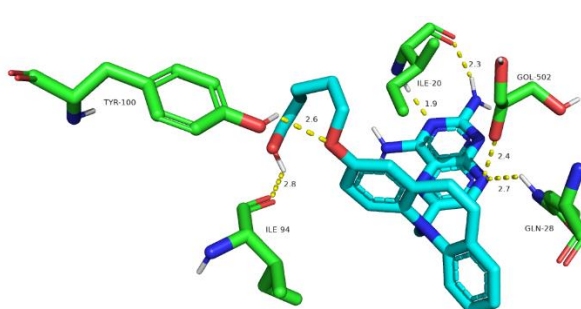
**BDBM50145799 (-9,6 kcal/mol)**
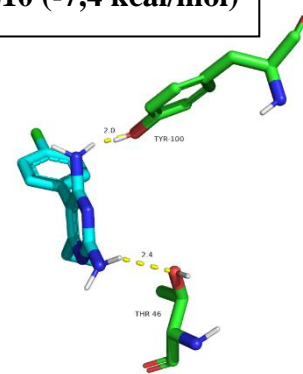
**Pyrimethamine(-8 kcal/mol)**

**Methotrexate(-8,9 kcal/mol)**

**Methylbenzoprim (-8,3 kcal/mol)**

**Trimethoprim (-7,5 kcal/mol)**

**Dihydrofolic acid (-8,8 kcal/mol)**

**Figure 13 : 3D interactions of Mt-DHFR + GOL with the selected small molecules, approved inhibitors and dihydrofolic acid obtained from PyMol.**

## II. Discussion

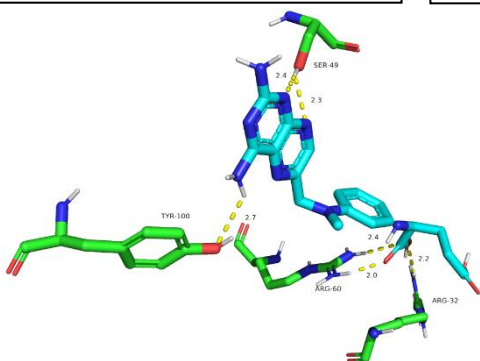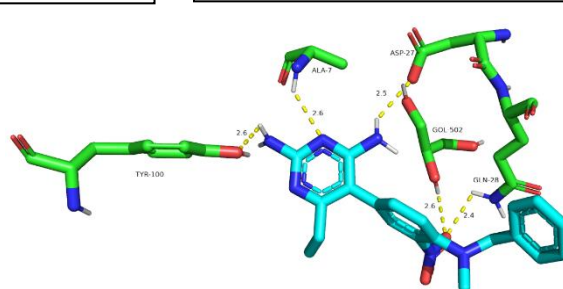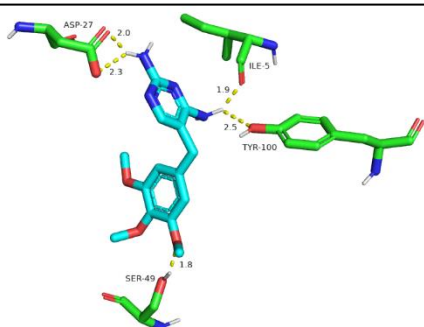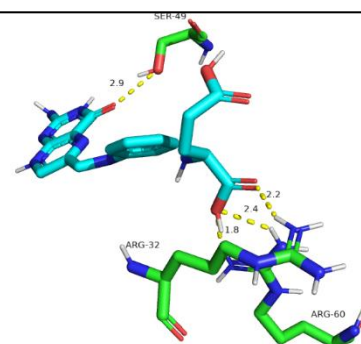The present study consist of using Structure-based virtual screening to identify novel small molecules that have an inhibitor activity against the Mt-DHFR, a data base of molecules was created using selleckchem and binding data base containing 8412 molecule, using Lipinski rule of five, veber rules and toxicity assessment, we had 11 molecules that can be used and docked against both human and bacterial enzyme to measure the potency as well as the selectivity, we've select also 5 known and FDA approved inhibitors to compare there activity and potency with our 11 molecules, as a positive control, we used the natural substrate (dihydrofolate acid) as a base line, the total of the 17 molecules were docked against the human enzyme (h-DHFR), the bacterial enzyme Mt-DHFR with and without the glycerol to see the effect of the extra glycerol pocket on the selectivity of the inhibitors and if we can use this information in the design of specific drugs, we used Autodock tools for the preparation of the ligands as well as the human h-DHFR [PDB ID: 1OHJ] and bacterial Mt-DHFR [PDB ID: 1DF7] , to perform the docking, we used Autodock Vina and the visualization of the hydrogen bonds is done by PyMol.

The 5 inhibitors of reference have shown a good affinity towards the enzyme, the most potent one is Methotrexate showed the lowest value of affinity energy (-8,9 kcal/mol), this molecule is used usually as anticancer drug, The second- best docking score goes to Methylbenzoprim with (-8,3 kcal/mol) which is frequently used in cancer chemotherapy as an DHFR inhibitor, followed by Pyrimethamine (-8 kcal/mol) a potent antimalarial agent, Trimethoprim and Bromo_WR99210 are also used as antimalarial agents as inhibitors of the DHFR that showed the highest values of affinity, -7.5 kcal/mol and -7.4 kcal/mol respectively, the natural substrate of the enzyme is the dihydrofolic acid and it showed a affinity of (-8.8 kcal/mol).

Based on the aforementioned information we take the Methotrexate as the inhibitor of reference to study the activity of the 11-inhibitor selected as it shows a better affinity than the natural substrate.

The h-DHFR docking results showed that only BDBM50514994 with (-9,9 kcal/mol) has a better affinity than the natural substrate that has a value of (-9,2 kcal/mol), followed by (-8,9 kcal/mol) for BDBM18228, (-8,2 kcal/mol) for both BDBM18225 and BDBM50145799, (-8 kcal/mol) is also the same for BDBM18236, BDBM18234 and BDBM18227 another tie for (-7,3 kcal/mol) by BDBM18226 and BDBM50145795, the BDBM18235 scored (-6,9 kcal/mol) and the last one is BDBM50145798 with a value of (-6,3 kcal/mol)

In regards to the Mt-DHFR docking results showed that (-9,3 kcal/mol) is the lowest affinity value by BDBM18227 next we have (-9,2 kcal/mol) for both BDBM18226 and BDBM18234, followed

by (-9 kcal/mol) for BDBM50145795 and BDBM50514994, BDBM18235 has a value of (-8,9 kcal/mol), BDBM50145798 scored (-8,8 kcal/mol), while BDBM18225 scored (-8,7 kcal/mol), BDBM18236 with (-8,6 kcal/mol), in the penultimate place BDBM18228 scoring (-8,5 kcal/mol) and finally BDBM50145799 with the highest affinity value of (-6,9 kcal/mol).

9 of the 11 selected molecules have a better affinity to Mt-DHFR+GOL than the natural substrate and 8 of them has better affinity then the inhibitor of reference, the lowest affinity value registered is (-9.6 kcal/mol) and it is the same between BDBM18226, BDBM50145795 and BDBM50145799, which are the best ones yet, followed by BDBM18225 (-9.4 kcal/mol), BDBM18236 (-9.2 kcal/mol), BDBM18227 (-9 kcal/mol) the BDBM50514994 and BDBM18228 share equal value of (-8.9 kcal/mol) the same one as the Methotrexate, and finally we have BDBM18235 (-8.8 kcal/mol), BDBM50145798 (-8 kcal/mol) and in last place BDBM18234 with (-7.3 kcal/mol)

This presence of glycerol showed a jump in the affinity for each of the potential inhibitors especially for BDBM18226, BDBM50145795 and BDBM50145799 from (-9,2 kcal/mol), (-9 kcal/mol) and (-6,9 kcal/mol) respectively for Mt-DHFR to a value of (-9,6 kcal/mol) with Mt-DHFR+GOL.

The best affinity found is (-9.6 kcal/mol) for the ligand BDBM18226, BDBM50145795 and BDBM50145799, with 7, 6, 6 hydrogen bonds respectively. This same affinity and different number of hydrogen bonds, is due to other types of interaction( hydrophobic and Van der Waals forces ), the Methotrexate have no selectivity for the bacterial enzyme Mt-DHFR+GOL since it's affinity is relatively close with both enzymes, (-8.9 kcal/mol) for bacterial protein and (-8.4 kcal/mol) for h-DHFR, Unlike Methotrexate the ligands BDBM18226, BDBM50145795 and BDBM50145799,have the strongest affinity toward the bacterial Mt-DHFR+GOL and the weakest one towards the human h-DHFR with (-7.3 kcal/mol) for BDBM18226, (-7.3 kcal/mol) for BDBM50145795 and finally (-8.2 kcal/mol) for BDBM50145799, which make them a specific set of new inhibitors for the Mt-DHFR.

# CONCLUSION

The present study was carried out to suggest potential inhibitors of Mt-DHFR, using virtual screening with the structure-based approach.

To do that we retrieved from Binding data base 8412 small molecules, after applying Lipinski's, Veber rules, toxicity assessment and molecular docking, 11 molecules have been identified as potential inhibitors, 3 molecules have a high specificity and affinity towards the Mt-DHFR and a low affinity to the h-DHFR.

These molecules could be proposed for the first time as potential specific inhibitors for the Mt-DHFR and could eventually replace the currently used antibiotics.

# REFERENCES

[1]  F. As and M. Dm, "The perpetual challenge of infectious diseases," *N. Engl. J. Med.*, vol. 366, no. 5, Feb. 2012, doi: 10.1056/NEJMra1108296.

[2]  M. Dm, F. Gk, and F. As, "Emerging infections: a perpetual challenge," *Lancet Infect. Dis.*, vol. 8, no. 11, Nov. 2008, doi: 10.1016/S1473-3099(08)70256-1.

[3]  World Health Organization, "World health statistics overview 2019: monitoring health for the SDGs, sustainable development goals," World Health Organization, Geneva, 2019. Accessed: Jun. 09, 2022. [Online]. Available: https://apps.who.int/iris/handle/10665/311696

[4]  C. Jr *et al.*, "Clonal expansion of both modern and ancient genotypes of Mycobacterium tuberculosis in southern Taiwan," *PloS One*, vol. 7, no. 8, 2012, doi: 10.1371/journal.pone.0043018.

[5]  K. Dheda *et al.*, "Global control of tuberculosis: from extensively drug-resistant to untreatable tuberculosis," *Lancet Respir. Med.*, vol. 2, no. 4, pp. 321–338, Apr. 2014, doi: 10.1016/S2213-2600(14)70031-1.

[6]  World Health Organization, *Global tuberculosis report 2018*. Geneva: World Health Organization, 2018. Accessed: Jun. 09, 2022. [Online]. Available: https://apps.who.int/iris/handle/10665/274453

[7]  A. Raju, S. Kulkarni, M. K. Ray, M. G. R. Rajan, and M. S. Degani, "E84G mutation in dihydrofolate reductase from drug resistant strains of Mycobacterium tuberculosis (Mumbai, India) leads to increased interaction with Trimethoprim," *Int. J. Mycobacteriology*, vol. 4, no. 2, pp. 97–103, Jun. 2015, doi: 10.1016/j.ijmyco.2015.02.001.

[8]  J. He, W. Qiao, Q. An, T. Yang, and Y. Luo, "Dihydrofolate reductase inhibitors for use as antimicrobial agents," *Eur. J. Med. Chem.*, vol. 195, p. 112268, Jun. 2020, doi: 10.1016/j.ejmech.2020.112268.

[9]  W. Yu and A. D. MacKerell, "Computer-Aided Drug Design Methods," *Methods Mol. Biol. Clifton NJ*, vol. 1520, pp. 85–106, 2017, doi: 10.1007/978-1-4939-6634-9_5.

[10] S. K. Sharma and A. Mohan, "Tuberculosis: From an incurable scourge to a curable disease - journey over a millennium," *Indian J. Med. Res.*, vol. 137, no. 3, pp. 455–493, Mar. 2013.

[11] A. Natarajan, P. M. Beena, A. V. Devnikar, and S. Mali, "A systemic review on tuberculosis," *Indian J. Tuberc.*, vol. 67, no. 3, pp. 295–311, Jul. 2020, doi: 10.1016/j.ijtb.2020.02.005.

[12] E. Cambau and M. Drancourt, "Steps towards the discovery of Mycobacterium tuberculosis by Robert Koch, 1882," *Clin. Microbiol. Infect.*, vol. 20, no. 3, pp. 196–201, Mar. 2014, doi: 10.1111/1469-0691.12555.

[13] A. Wallgren, "The time-table of tuberculosis," *Tubercle*, vol. 29, no. 11, pp. 245–251, Nov. 1948, doi: 10.1016/S0041-3879(48)80033-4.

[14] D. V. Havlir, H. Getahun, I. Sanne, and P. Nunn, "Opportunities and Challenges for HIV Care in Overlapping HIV and TB Epidemics," *JAMA*, vol. 300, no. 4, pp. 423–430, Jul. 2008, doi: 10.1001/jama.300.4.423.

[15] H. Getahun *et al.*, "Management of latent Mycobacterium tuberculosis infection: WHO guidelines for low tuberculosis burden countries," *Eur. Respir. J.*, vol. 46, no. 6, pp. 1563–1576, Dec. 2015, doi: 10.1183/13993003.01245-2015.

[16] N. Ford *et al.*, "Causes of hospital admission among people living with HIV worldwide: a systematic review and meta-analysis," *Lancet HIV*, vol. 2, no. 10, pp. e438-444, Oct. 2015, doi: 10.1016/S2352-3018(15)00137-X.

[17] K. Lönnroth *et al.*, "Tuberculosis control and elimination 2010–50: cure, care, and social development," *The Lancet*, vol. 375, no. 9728, pp. 1814–1829, May 2010, doi: 10.1016/S0140-6736(10)60483-7.

[18] C. Y. Jeon and M. B. Murray, "Diabetes mellitus increases the risk of active tuberculosis: a systematic review of 13 observational studies," *PLoS Med.*, vol. 5, no. 7, p. e152, Jul. 2008, doi: 10.1371/journal.pmed.0050152.

[19] J. Rehm *et al.*, "The association between alcohol use, alcohol use disorders and tuberculosis (TB). A systematic review," *BMC Public Health*, vol. 9, p. 450, Dec. 2009, doi: 10.1186/1471-2458-9-450.

[20] M. N. Bates, A. Khalakdina, M. Pai, L. Chang, F. Lessa, and K. R. Smith, "Risk of tuberculosis from exposure to tobacco smoke: a systematic review and meta-analysis," *Arch. Intern. Med.*, vol. 167, no. 4, pp. 335–342, Feb. 2007, doi: 10.1001/archinte.167.4.335.

[21] "Global tuberculosis report 2021." https://www.who.int/publications-detail-redirect/9789240037021 (accessed Jun. 10, 2022).

[22] "Ministry of Health, 2021." https://www.sante.gov.ma/Pages/Communiques.aspx?IDCom=408 (accessed Jun. 16, 2022).

[23] N. Swaminathan, S. R. Perloff, and J. M. Zuckerman, "Prevention of Mycobacterium tuberculosis Transmission in Health Care Settings," *Infect. Dis. Clin. North Am.*, vol. 35, no. 4, pp. 1013–1025, Dec. 2021, doi: 10.1016/j.idc.2021.07.003.

[24] D. W. Dowdy, A. S. Azman, E. A. Kendall, and B. Mathema, "Transforming the fight against tuberculosis: targeting catalysts of transmission," *Clin. Infect. Dis. Off. Publ. Infect. Dis. Soc. Am.*, vol. 59, no. 8, pp. 1123–1129, Oct. 2014, doi: 10.1093/cid/ciu506.

[25] J. A. Philips and J. D. Ernst, "Tuberculosis pathogenesis and immunity," *Annu. Rev. Pathol.*, vol. 7, pp. 353–384, 2012, doi: 10.1146/annurev-pathol-011811-132458.

[26] W. F. Wells, H. L. Ratcliffe, and C. Grumb, "On the mechanics of droplet nuclei infection; quantitative experimental air-borne tuberculosis in rabbits," *Am. J. Hyg.*, vol. 47, no. 1, pp. 11–28, Jan. 1948, doi: 10.1093/oxfordjournals.aje.a119179.

[27] C. J. Cambier, S. Falkow, and L. Ramakrishnan, "Host evasion and exploitation schemes of Mycobacterium tuberculosis," *Cell*, vol. 159, no. 7, pp. 1497–1509, Dec. 2014, doi: 10.1016/j.cell.2014.11.024.

[28] N. K. Dutta and P. C. Karakousis, "Latent tuberculosis infection: myths, models, and molecular mechanisms," *Microbiol. Mol. Biol. Rev. MMBR*, vol. 78, no. 3, pp. 343–371, Sep. 2014, doi: 10.1128/MMBR.00010-14.

[29] L. Arango, A. W. Brewin, and J. F. Murray, "The spectrum of tuberculosis as currently seen in a metropolitan hospital," *Am. Rev. Respir. Dis.*, vol. 108, no. 4, pp. 805–812, Oct. 1973, doi: 10.1164/arrd.1973.108.4.805.

[30] C. E. Barry *et al.*, "The spectrum of latent tuberculosis: rethinking the biology and intervention strategies," *Nat. Rev. Microbiol.*, vol. 7, no. 12, pp. 845–855, Dec. 2009, doi: 10.1038/nrmicro2236.

[31] C. Lange and T. Mori, "Advances in the diagnosis of tuberculosis," *Respirol. Carlton Vic*, vol. 15, no. 2, pp. 220–240, Feb. 2010, doi: 10.1111/j.1440-1843.2009.01692.x.

[32] A. Charif, "ACTUALITES DIAGNOSTIQUES ET THERAPEUTIQUES DE LA TUBERCULOSE PULMONAIRE. / CHARIF Ahmed El Mehdi," Thesis, 2020. Accessed: Jun. 10, 2022. [Online]. Available: http://ao.um5.ac.ma/xmlui/handle/123456789/18060

[33] C. Saltini, "Chemotherapy and diagnosis of tuberculosis," *Respir. Med.*, vol. 100, no. 12, pp. 2085–2097, Dec. 2006, doi: 10.1016/j.rmed.2006.09.015.

[34] J. R. Andrews *et al.*, "The dynamics of QuantiFERON-TB gold in-tube conversion and reversion in a cohort of South African adolescents," *Am. J. Respir. Crit. Care Med.*, vol. 191, no. 5, pp. 584–591, Mar. 2015, doi: 10.1164/rccm.201409-1704OC.

[35] D. Menzies, M. Pai, and G. Comstock, "Meta-analysis: new tests for the diagnosis of latent tuberculosis infection: areas of uncertainty and recommendations for research," *Ann. Intern. Med.*, vol. 146, no. 5, pp. 340–354, Mar. 2007, doi: 10.7326/0003-4819-146-5-200703060-00006.

[36] G. Aviram, "Chest radiography for tuberculosis screening: a valuable tool," *Isr. Med. Assoc. J. IMAJ*, vol. 17, no. 1, pp. 50–51, Jan. 2015.

[37] W. Wang *et al.*, "Skin test of tuberculin purified protein derivatives with a dissolving microneedle-array patch," *Drug Deliv. Transl. Res.*, vol. 9, no. 4, pp. 795–801, Aug. 2019, doi: 10.1007/s13346-019-00629-y.

[38] S. Salmanzadeh, H. Tavakkol, K. Bavieh, and S. M. Alavi, "Diagnostic Value of Serum Adenosine Deaminase (ADA) Level for Pulmonary Tuberculosis," *Jundishapur J. Microbiol.*, vol. 8, no. 3, p. e21760, Mar. 2015, doi: 10.5812/jjm.21760.

[39] R. F. D'Amato *et al.*, "Rapid diagnosis of pulmonary tuberculosis by using Roche AMPLICOR Mycobacterium tuberculosis PCR test," *J. Clin. Microbiol.*, vol. 33, no. 7, pp. 1832–1834, Jul. 1995, doi: 10.1128/jcm.33.7.1832-1834.1995.

[40] R. Blakemore *et al.*, "Evaluation of the analytical performance of the Xpert MTB/RIF assay," *J. Clin. Microbiol.*, vol. 48, no. 7, pp. 2495–2501, Jul. 2010, doi: 10.1128/JCM.00128-10.

[41] D. Cao *et al.*, "Real-time fluorescence Loop-Mediated Isothermal Amplification (LAMP) for rapid and reliable diagnosis of pulmonary tuberculosis," *J. Microbiol. Methods*, vol. 109, pp. 74–78, Feb. 2015, doi: 10.1016/j.mimet.2014.12.013.

[42] J. Y. Kim, M. J. Ferraro, and J. A. Branda, "False-negative results obtained with the Gen-Probe Amplified Mycobacterium tuberculosis direct test caused by unrecognized inhibition of the amplification reaction," *J. Clin. Microbiol.*, vol. 47, no. 9, pp. 2995–2997, Sep. 2009, doi: 10.1128/JCM.00966-09.

[43] S. T. Cole *et al.*, "Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence," *Nature*, vol. 393, no. 6685, pp. 537–544, Jun. 1998, doi: 10.1038/31159.

[44] T. Garnier *et al.*, "The complete genome sequence of Mycobacterium bovis," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 100, no. 13, pp. 7877–7882, Jun. 2003, doi: 10.1073/pnas.1130426100.

[45] N. H. Smith, R. G. Hewinson, K. Kremer, R. Brosch, and S. V. Gordon, "Myths and misconceptions: the origin and evolution of Mycobacterium tuberculosis," *Nat. Rev. Microbiol.*, vol. 7, no. 7, pp. 537–544, Jul. 2009, doi: 10.1038/nrmicro2165.

[46] S. Chetty, M. Ramesh, A. Singh-Pillay, and M. E. S. Soliman, "Recent advancements in the development of anti-tuberculosis drugs," *Bioorg. Med. Chem. Lett.*, vol. 27, no. 3, pp. 370–386, Feb. 2017, doi: 10.1016/j.bmcl.2016.11.084.

[47] S. Tiberi *et al.*, "Tuberculosis: progress and advances in development of new drugs, treatment regimens, and host-directed therapies," *Lancet Infect. Dis.*, vol. 18, no. 7, pp. e183–e198, Jul. 2018, doi: 10.1016/S1473-3099(18)30110-5.

[48] F. Tritar, H. Daghfous, S. Ben Saad, and L. Slim-Saidi, "Prise en charge de la tuberculose multirésistante," *Rev. Pneumol. Clin.*, vol. 71, no. 2–3, pp. 130–139, Apr. 2015, doi: 10.1016/j.pneumo.2014.05.001.

[49] "Diagnostic et traitement de la tuberculose. Manuel pratique Recommandations destinées au corps médical. Editeur responsable : JP Van Vooren — FARES asbl.," *vdocuments.net*. https://vdocuments.net/diagnostic-et-traitement-de-la-2018-01-12-diagnostic-et-traitement-de-la-tuberculose.html (accessed Jun. 10, 2022).

[50] E. Masson, "Recommendations of the Society of Pneumology of the French Language on the management of tuberculosis in France," *EM-Consulte*. https://www.em-consulte.com/article/144470/recommandations-de-la-societe-de-pneumologie-de-la (accessed Jun. 10, 2022).

[51] K. C. Chang, C. C. Leung, W. W. Yew, S. C. Ho, and C. M. Tam, "A nested case-control study on treatment-related risk factors for early relapse of tuberculosis," *Am. J. Respir. Crit. Care Med.*, vol. 170, no. 10, pp. 1124–1130, Nov. 2004, doi: 10.1164/rccm.200407-905OC.

[52] A. S. Dean *et al.*, "25 years of surveillance of drug-resistant tuberculosis: achievements, challenges, and way forward," *Lancet Infect. Dis.*, Mar. 2022, doi: 10.1016/S1473-3099(21)00808-2.

[53] R. Li *et al.*, "Three-dimensional structure of M. tuberculosis dihydrofolate reductase reveals opportunities for the design of novel tuberculosis drugs," *J. Mol. Biol.*, vol. 295, no. 2, pp. 307–323, Jan. 2000, doi: 10.1006/jmbi.1999.3328.

[54] M. Kumar, R. Vijayakrishnan, and G. Subba Rao, "In silico structure-based design of a novel class of potent and selective small peptide inhibitor of Mycobacterium tuberculosis Dihydrofolate reductase, a potential target for anti-TB drug discovery," *Mol. Divers.*, vol. 14, no. 3, pp. 595–604, Aug. 2010, doi: 10.1007/s11030-009-9172-6.

[55] R. S. Swanwick, P. J. Shrimpton, and R. K. Allemann, "Pivotal role of Gly 121 in dihydrofolate reductase from Escherichia coli: the altered structure of a mutant enzyme may form the basis of its diminished catalytic performance," *Biochemistry*, vol. 43, no. 14, pp. 4119–4127, Apr. 2004, doi: 10.1021/bi036164k.

[56] K. A. Brown and J. Kraut, "Exploring the molecular mechanism of dihydrofolate reductase," *Faraday Discuss.*, no. 93, pp. 217–224, 1992, doi: 10.1039/fd9929300217.

[57] "Coupling interactions of distal residues enhance dihydrofolate reductase catalysis: mutational effects on hydride transfer rates - PubMed." https://pubmed.ncbi.nlm.nih.gov/12379104/ (accessed Jun. 08, 2022).

[58] W. Hong *et al.*, "The influence of glycerol on the binding of methotrexate to *Mycobacterium tuberculosis* dihydrofolate reductase: a molecular modelling study," *Mol. Simul.*, vol. 41, no. 18, pp. 1540–1545, Dec. 2015, doi: 10.1080/08927022.2015.1048513.

[59] M. F. Stevens *et al.*, "Structural studies on bioactive compounds. 28. Selective activity of triazenyl-substituted pyrimethamine derivatives against Pneumocystis carinii dihydrofolate reductase," *J. Med. Chem.*, vol. 40, no. 12, pp. 1886–1893, Jun. 1997, doi: 10.1021/jm970050n.

[60] T. P. Remcho *et al.*, "Regioisomerization of Antimalarial Drug WR99210 Explains the Inactivity of a Commercial Stock," *Antimicrob. Agents Chemother.*, vol. 65, no. 1, pp. e01385-20, Dec. 2020, doi: 10.1128/AAC.01385-20.

[61] H. Liu *et al.*, "Antimalarial Drug Pyrimethamine Plays a Dual Role in Antitumor Proliferation and Metastasis through Targeting DHFR and TP," *Mol. Cancer Ther.*, vol. 18, no. 3, pp. 541–555, Mar. 2019, doi: 10.1158/1535-7163.MCT-18-0936.

[62] "Trimethoprim and other nonclassical antifolates an excellent template for searching modifications of dihydrofolate reductase enzyme inhibitors | The Journal of Antibiotics." https://www.nature.com/articles/s41429-019-0240-6 (accessed Jun. 08, 2022).

[63] S. J. Y. Macalino, V. Gosu, S. Hong, and S. Choi, "Role of computer-aided drug design in modern drug discovery," *Arch. Pharm. Res.*, vol. 38, no. 9, pp. 1686–1701, Sep. 2015, doi: 10.1007/s12272-015-0640-5.

[64] S. Kalyaanamoorthy and Y.-P. P. Chen, "Structure-based drug design to augment hit discovery," *Drug Discov. Today*, vol. 16, no. 17–18, pp. 831–839, Sep. 2011, doi: 10.1016/j.drudis.2011.07.006.

[65] L. G. Ferreira, R. N. Dos Santos, G. Oliva, and A. D. Andricopulo, "Molecular docking and structure-based drug design strategies," *Mol. Basel Switz.*, vol. 20, no. 7, pp. 13384–13421, Jul. 2015, doi: 10.3390/molecules200713384.

[66] wwPDB consortium, "Protein Data Bank: the single global archive for 3D macromolecular structure data," *Nucleic Acids Res.*, vol. 47, no. D1, pp. D520–D528, Jan. 2019, doi: 10.1093/nar/gky949.

[67] A. M. Lesk and C. Chothia, "How different amino acid sequences determine similar protein structures: the structure and evolutionary dynamics of the globins," *J. Mol. Biol.*, vol. 136, no. 3, pp. 225–270, Jan. 1980, doi: 10.1016/0022-2836(80)90373-3.

[68] K. Illergård, D. H. Ardell, and A. Elofsson, "Structure is three to ten times more conserved than sequence--a study of structural response in protein cores," *Proteins*, vol. 77, no. 3, pp. 499–508, Nov. 2009, doi: 10.1002/prot.22458.

[69] A. Ingles-Prieto *et al.*, "Conservation of protein structure over four billion years," *Struct. Lond. Engl. 1993*, vol. 21, no. 9, pp. 1690–1697, Sep. 2013, doi: 10.1016/j.str.2013.06.020.

[70] D. T. Jones, W. R. Taylor, and J. M. Thornton, "A new approach to protein fold recognition," *Nature*, vol. 358, no. 6381, pp. 86–89, Jul. 1992, doi: 10.1038/358086a0.

[71] K. T. Simons, C. Kooperberg, E. Huang, and D. Baker, "Assembly of protein tertiary structures from fragments with similar local sequences using simulated annealing and Bayesian scoring functions," *J. Mol. Biol.*, vol. 268, no. 1, pp. 209–225, Apr. 1997, doi: 10.1006/jmbi.1997.0959.

[72] J. J.-L. Liao and R. C. Andrews, "Targeting protein multiple conformations: a structure-based strategy for kinase drug design," *Curr. Top. Med. Chem.*, vol. 7, no. 14, pp. 1394–1407, 2007, doi: 10.2174/156802607781696783.

[73] T. Wunberg *et al.*, "Improving the hit-to-lead process: data-driven assessment of drug-like and lead-like screening hits," *Drug Discov. Today*, vol. 11, no. 3–4, pp. 175–180, Feb. 2006, doi: 10.1016/S1359-6446(05)03700-1.

[74] A. C. Anderson, "The process of structure-based drug design," *Chem. Biol.*, vol. 10, no. 9, pp. 787–797, Sep. 2003, doi: 10.1016/j.chembiol.2003.09.002.

[75] *Structure-Based Drug Discovery*. 2006. doi: 10.1039/9781847552549.

[76] C. C. Melo-Filho, R. C. Braga, and C. H. Andrade, "3D-QSAR Approaches in Drug Design: Perspectives to Generate Reliable CoMFA Models," *Curr. Comput. Aided Drug Des.*, vol. 10, no. 2, pp. 148–159.

[77] A. Lavecchia and C. Di Giovanni, "Virtual screening strategies in drug discovery: a critical review," *Curr. Med. Chem.*, vol. 20, no. 23, pp. 2839–2860, 2013, doi: 10.2174/09298673113209990001.

[78] A. Vuorinen and D. Schuster, "Methods for generating and applying pharmacophore models as virtual screening filters and for bioactivity profiling," *Methods*, vol. 71, Oct. 2014, doi: 10.1016/j.ymeth.2014.10.013.

[79] S. P. Leelananda and S. Lindert, "Computational methods in drug discovery," *Beilstein J. Org. Chem.*, vol. 12, pp. 2694–2718, 2016, doi: 10.3762/bjoc.12.267.

[80] PubChem, "PubChem." https://pubchem.ncbi.nlm.nih.gov/ (accessed Jun. 11, 2022).

[81] "mcule." https://mcule.com/ (accessed Jun. 11, 2022).

[82] "Selleckchem.com - Bioactive Compounds Expert (Bioactive Compounds,Compound Libraries)." https://www.selleckchem.com/ (accessed Jun. 11, 2022).

[83] T. Liu, Y. Lin, X. Wen, R. N. Jorissen, and M. K. Gilson, "BindingDB: a web-accessible database of experimentally determined protein-ligand binding affinities," *Nucleic Acids Res.*, vol. 35, no. Database, pp. D198–D201, Jan. 2007, doi: 10.1093/nar/gkl999.

[84] "AutoDock." https://autodock.scripps.edu/ (accessed Jun. 11, 2022).

[85] "PyMOL | pymol.org." https://pymol.org/2/ (accessed Jun. 11, 2022).

[86] N. M. O'Boyle, M. Banck, C. A. James, C. Morley, T. Vandermeersch, and G. R. Hutchison, "Open Babel: An open chemical toolbox," *J. Cheminformatics*, vol. 3, p. 33, Oct. 2011, doi: 10.1186/1758-2946-3-33.

[87] "Toxicity checker." https://mcule.com/apps/toxicity-checker/ (accessed Jun. 11, 2022).

[88] "Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings - PubMed." https://pubmed.ncbi.nlm.nih.gov/11259830/ (accessed Jun. 11, 2022).

[89] C. A. Lipinski, "Lead- and drug-like compounds: the rule-of-five revolution," *Drug Discov. Today Technol.*, vol. 1, no. 4, pp. 337–341, Dec. 2004, doi: 10.1016/j.ddtec.2004.11.007.

[90] N. M. Hassan, A. A. Alhossary, Y. Mu, and C.-K. Kwoh, "Protein-Ligand Blind Docking Using QuickVina-W With Inter-Process Spatio-Temporal Integration," *Sci. Rep.*, vol. 7, no. 1, p. 15451, Nov. 2017, doi: 10.1038/s41598-017-15571-7.