WEP Warwick Evans Publishing

International Journal of Public Health and Medical Research

ISSN: 3006-7707 (Print), ISSN: 3006-3752 (Online) | Volume 2, Number 3, Year 2024
DOI: https://doi.org/10.62051/ijphmr.v2n3.07
Journal homepage: https://wepub.org/index.php/IJPHMR/index

Research on Virtual Screening of Natural Compounds Against Protozoan Parasites Using Comparative Genomics and Molecular Docking Techniques

Menghan He [1], Shuquan Guo [2, *]

[1] Laboratory of Molecular Virology and Immunology, Technology Innovation Center, Haid Research Institute Guangdong Haid Group Co, Ltd., Panyu, Guangzhou 511400, China

[2] Institute of Hydrobiology, Jinan University Guangdong Guangzhou 510632, China

ABSTRACT

This study aims to identify conserved drug targets in protozoan parasites through genomic data mining and use computational virtual screening to discover small molecule inhibitors from the TCMSP database that target these conserved drug targets. Comparative genomics was applied to identify key drug targets in protozoan parasites, and high-throughput screening was conducted using the AutoDock Vina molecular docking program. The results were further refined using docking scores and the "ADME-Lipinski" rule, followed by an analysis of the interactions between small molecules and the drug targets. Potential small molecules with anti-parasitic activity were predicted. In this study, the DHFR-TS enzyme was identified as a conserved drug target for protozoan parasites through comparative genomics. Using DHFR-TS as the target and a known positive control ligand, six compounds with favorable drug-like properties were selected from the TCMSP database: MOL013376, MOL005018, MOL009591, MOL002047, MOL011499, and MOL012692. These compounds demonstrated superior binding affinity and interaction with DHFR-TS compared to the positive control ligand. This study combines comparative genomics and computational virtual screening to identify potential natural compounds from Traditional Chinese Medicine for protozoan parasites, providing a solid theoretical foundation for the design and experimental synthesis of novel anti-parasitic drug candidates.

KEYWORDS

Parasitic protozoa; Virtual screening; Molecular docking; DHFR-thymidylate synthase

1. INTRODUCTION

Protozoa are among the most primitive and simplest unicellular organisms in the animal kingdom, with approximately 30,000 species. Diseases caused by protozoan pathogens, such as Trypanosoma brucei (which causes African trypanosomiasis), Leishmania major (which causes leishmaniasis), Plasmodium falciparum (which causes malaria), and Toxoplasma gondii (which causes toxoplasmosis), pose significant threats to human health, resulting in tens of thousands of deaths worldwide each year [1-3]. Currently, no vaccines exist for these diseases. Due to the toxicity of available drugs and the emergence of drug resistance in pathogens, only a few treatments remain effective. Consequently, researchers are focused on discovering new drug targets and developing innovative therapies to treat diseases caused by protozoan pathogens.

The development of Western medicines for treating parasitic diseases faces challenges such as long development timelines, high costs, toxicity, and significant side effects. In contrast, Traditional Chinese Medicine (TCM), a valuable part of China's medical heritage, has been widely used both in

Content from this work may be used under the terms of CC BY-NC 4.0 licence (https://creativecommons.org/licenses/by-nc/4.0/).
Published by Warwick Evans Publishing.

the treatment of parasitic diseases and in enhancing the body's immune response [4]. Traditional Chinese herbal medicine is not only characterized by low toxicity and minimal side effects, but also by its ability to target multiple pathways. This offers a promising new approach for treating parasitic infections and preventing or reducing drug resistance. In recent years, comparative genomics and molecular virtual screening have become core technologies in the field of computer-aided drug design, widely applied in the discovery of parasite drug targets and the drug development process [5]. Comparative genomics, utilizing genomic mapping and sequencing technologies, involves the comparative analysis of the genomic structure and functional gene regions across multiple individuals (populations) of a species or between different species [6]. Molecular docking is a method of drug design based on the characteristics of the target and the interactions between the target and lead compounds. It enables the virtual screening of active ingredients in a short period of time and has become a key technology in the field of computer-aided drug research [7].

This study leverages comparative genomics and molecular docking technologies to identify conserved drug targets in protozoan parasites. The identified targets were then docked with compounds from TCMSP database, and potential lead compounds with anti-parasitic activity were virtually screened based on docking scores. This approach provides a scientific foundation for the development of treatments for parasitic diseases.

2. MATERIALS AND METHODS

2.1. Identification of Conserved Protozoan Genes

Comparative genomics was employed to identify conserved genes in protozoan parasites. In this study, the genomic data of Tetrahymena thermophila, Plasmodium falciparum, Toxoplasma gondii, Trypanosoma brucei, and Leishmania major were analyzed using the OrthoVenn2 platform (https://orthovenn2.bioinfotoolkits.net) to identify orthologous genes across these protozoan species. The parameters were set as follows: e-value $\leq 1 \times 10^{-5}$ and inflation value ≤ 1.5. The genomic information for these protozoa was retrieved from the OrthoVenn2 database [8].

2.2. Homology and Phylogenetic Analysis of Protozoan DHFR-TS

Homology analysis of the DHFR-TS amino acid sequences from five protozoan species was performed using BioEdit software and MEGA X software. The DHFR-TS gene sequences for protozoan species used for amino acid sequence analysis, homology alignment, and phylogenetic tree construction were downloaded from the NCBI database.

2.3. Preparation of Small Molecule Compound Database

TCMSP (Traditional Chinese Medicine Systems Pharmacology Database and Analysis Platform) is a comprehensive database for systems pharmacology, network pharmacology, and systems biology related to Traditional Chinese Medicine [9]. In this study, 12,990 small molecules from the TCMSP database were used as ligands for virtual screening. The three-dimensional (3D) structures of these compounds were downloaded from the database to create a molecular docking database for virtual screening.

2.4. Preparation of DHFR-TS Protein for Protozoan Parasites

Crystallographic or NMR structures of DHFR-TS proteins from protozoan parasites were retrieved from the RCSB Protein Data Bank (http://www.rcsb.org/). The crystal structure of the DHFR-TS protein from Toxoplasma gondii (PDB ID: 4KY4) was chosen as the receptor for docking studies. The 3D structure of DHFR-TS protein was downloaded, and preprocessing was carried out using

AutoDockTools software, which involved removing redundant atoms from chains other than chain A and adding hydrogen atoms.

2.5. Virtual Screening of Active Compounds from Traditional Chinese Medicine

All virtual screening operations were conducted on the Yinfu Cloud Computing Platform (https://cloud.yinfotek.com/). The small molecules were processed using the platform's tools (for preparing compound structures), which generated their 3D structures, added hydrogen atoms, and performed energy optimization using the MMF94 force field. The processed 3D structure of the DHFR-TS protein was uploaded, and the binding pockets were defined according to the active binding sites (ILE402, LEU516, PHE520, MET608, ASP513, ALA609, ASN409) reported in the literature [10]. The center of the binding box was set at (-16.262, -22.303, -49.858) with a size of (26, 26, 26). Semi-flexible docking was performed using AutoDock Vina [11] , with one best conformation output per compound. Compounds with affinity > -9 kcal/mol were filtered out, and the top 100 ranked compounds were selected for further screening.

2.6. ADME and Lipinski Rule-Based Screening

"ADME" refers to the absorption, distribution, metabolism, and excretion of a drug, representing the processes by which the body handles the drug. ADME-based screening is crucial in drug development [12]. In clinical drug development, compounds with oral bioavailability (OB) ≥ 20% and drug-likeness (DL) ≥ 0.18 are considered for screening. The Lipinski Rule of Five, an empirical guideline, is used to assess whether a compound is likely to be an effective oral drug, based on its molecular properties. The rule specifies: molecular weight (MW) < 500 Da, hydrogen bond donors (Hdon) ≤ 5, lipophilicity (clogP) < 5, hydrogen bond acceptors (Hacc) ≤ 10, and rotatable bonds (RBN) ≤ 10[13]. However, researchers have identified limitations in the application of this rule, leading to the development of simplified versions such as the "Four Rules" and "Three Rules." In this study, compounds ranked in the top 100 were further screened using ADME parameters and the simplified Lipinski criteria, excluding the constraints on hydrogen bond acceptors (Hacc) and rotatable bond numbers (RBN), to identify small molecules with favorable pharmacokinetic properties and oral bioavailability.

2.7. Analysis of Screening Results

The small molecules selected based on molecular docking, ADME, and Lipinski principles were re-docked using AutoDock Vina to generate 3D ligand-receptor complexes. Binding sites for the small molecules were analyzed using Discovery Studio Visualizer and amino acid sequence homology alignment. If the DHFR-TS proteins from all five protozoan species exhibited binding at the same amino acid residues, the compound was considered a potential lead inhibitor for protozoan parasites.

3. RESULTS AND ANALYSIS

3.1. Comparative Genomic Analysis

The OrthoVenn2 orthologous gene cluster analysis revealed a total of 9,081 gene clusters, of which 8,655 were orthologous gene clusters (containing genes from at least two species) and 426 were single-copy gene clusters. Among these, Leishmania major had 5,222 orthologous genes and 2,302 single-copy genes; Plasmodium falciparum had 2,312 orthologous genes and 2,461 single-copy genes; Tetrahymena thermophila had 3,621 orthologous genes and 10,568 single-copy genes; Trypanosoma brucei had 5,224 orthologous genes and 2,484 single-copy genes; and Toxoplasma gondii had 2,761 orthologous genes and 6,504 single-copy genes (Figure 1b). The core genome, represented by the shared gene clusters across the five protozoan species, contained 626 gene clusters, accounting for

7.66% of the total gene clusters. Among these, Plasmodium falciparum had the fewest unique gene clusters (94), while the other protozoan species had the following unique gene clusters: Tetrahymena thermophila (1,867), Toxoplasma gondii (202), Trypanosoma brucei (162), and Leishmania major (107) (Figure 1a). Further analysis of the 626 shared gene clusters revealed that 1,285 orthologous gene clusters exhibited a high degree of conservation across species and could serve as potential common drug targets (Figure 1c).

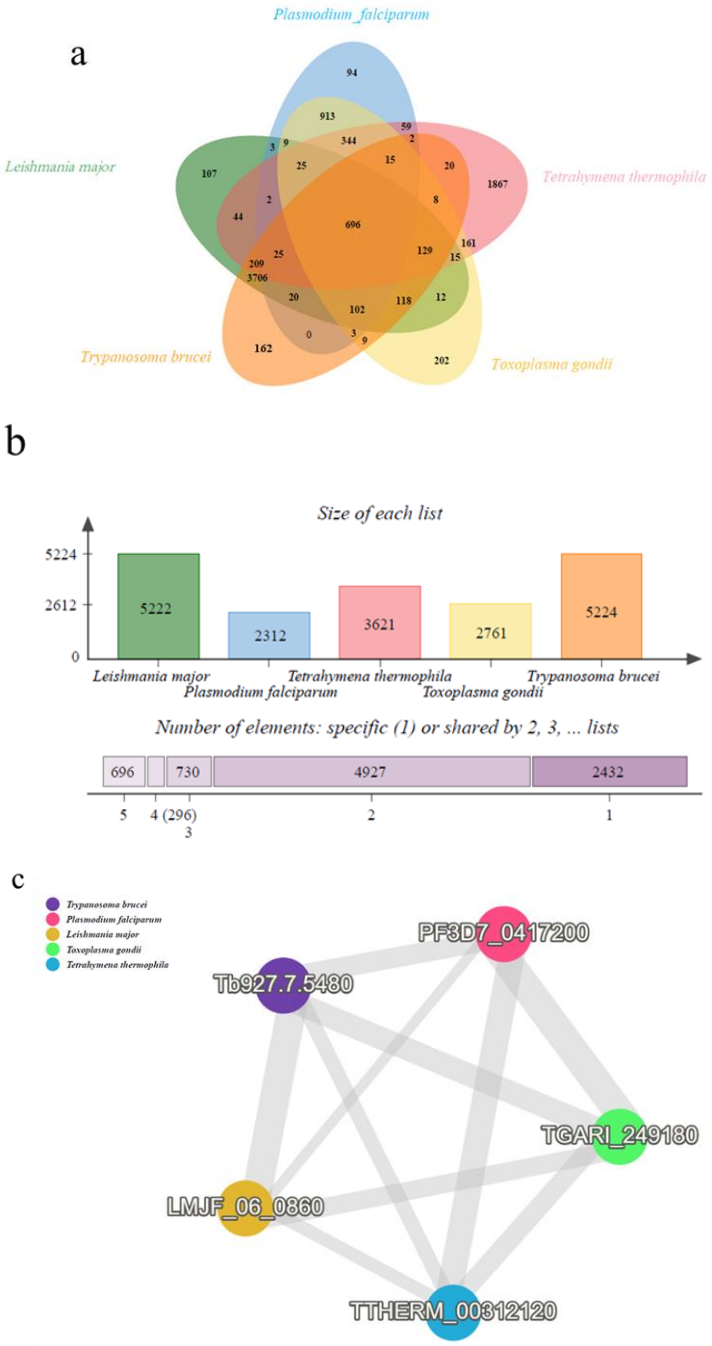


Figure 1. Cluster analysis of 5 species of protozoa direct homology

Note: a: Analysis of core genes, unique genes, and concurrent genes of 5 species of protozoa; b: Number analysis of 5 species of protozoa gene families; c: Similarity network diagram of homologous genes.

3.2. Homology Analysis of DHFR-TS Gene Amino Acid Sequences in Five Protozoan Species

Homology analysis using the NCBI database confirmed that the gene family in question corresponds to Dihydrofolate Reductase-Thymidylate Synthase (DHFR-TS). Amino acid sequence alignment revealed that the TS domain of this protein is highly conserved across the five protozoan species (Figure 2). BLAST analysis (NCBI, https://blast.ncbi.nlm.nih.gov/) indicated that the TS amino acid sequence of Toxoplasma gondii (326–610 aa) shares sequence similarities of 61.40%, 58.04%, 59.23%, and 62.46% with Plasmodium falciparum (327–607 aa), Trypanosoma brucei (243–527 aa), Leishmania major (236–520 aa), and Tetrahymena thermophila (201–484 aa), respectively. Furthermore, the active-site residue binding locations within the TS domain were found to be highly conserved across all five protozoan species (Figure 2). These results suggest that the thymidylate synthase region of DHFR-TS could serve as a shared drug target for these protozoan parasites.

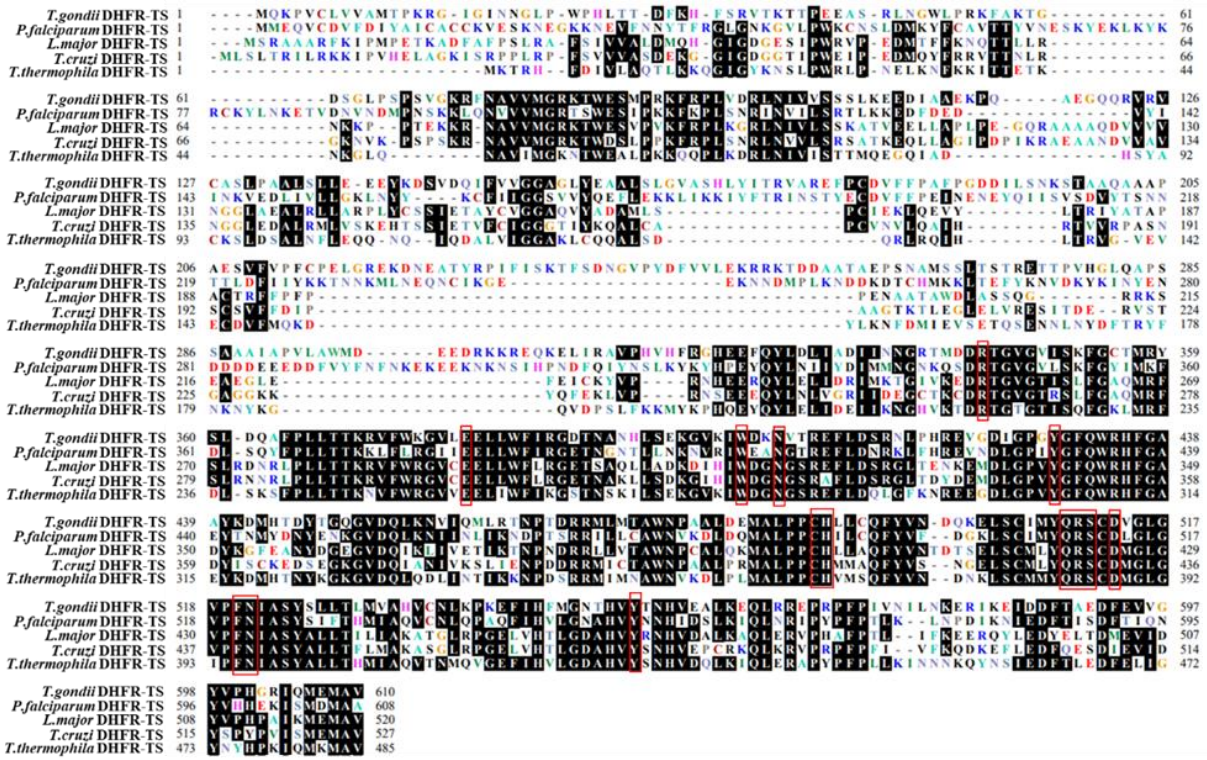


Figure 2. Amino acid sequence multiple alignments of the DHFR-TS from 5 species of protozoa

The red box represents active site residues of the TS structural domain.

3.3. Molecular Docking, ADME, and Lipinski Rule Screening Results

The three-dimensional crystal structure of Toxoplasma gondii DHFR-TS (PDB: 4KY4), with the TS domain as the docking pocket, was selected from the PDB database. The inhibitor binding sites (ILE402, LEU516, PHE520, MET608, ASP513, ALA609, ASN409) reported by [10] were used to define the docking pocket coordinates as (-16.262, -22.303, -49.858). Using AutoDock Vina for molecular docking and screening based on ADME and Lipinski's rule, six compounds with the highest potential activity were identified. The docking scores and structures of these compounds are shown in Table 1 and Figure 3. All six compounds exhibited docking scores higher than that of the reference ligand 1UE (-9.7 kcal/mol). Among them, jujubogenin (saponin aglycone from Ziziphus jujuba) had the highest binding affinity (-12 kcal/mol), followed by xambioona, korsine, melianone, Daturameteline A, and kuwanon D, all of which demonstrated good binding affinities (≥-11.3 kcal/mol).

Table 1. Six active compounds obtained by virtual screening

Compound number	Compound name	Source of natural compounds	Binding energy (kcal /mol)
MOL013376	Jujubogenin	Hoveniae Dulcis Semen	-12
MOL005018	Xambioona	Licorice	-11.9
MOL009591	Korsine	Fritiliariae Irrhosae Bulbus	-11.9
MOL002047	Elianone	Toosendan Fructus	-11.4
MOL011499	Daturameteline A	Daturae Flos	-11.4
MOL012692	Kuwanon D	Mori Cortex	-11.3
1UE	2-amino-5-(phenylsulfanyl)-3, 9-dihydro-4H-pyrimido [4, 5-b]indol-4-one		- 9.7

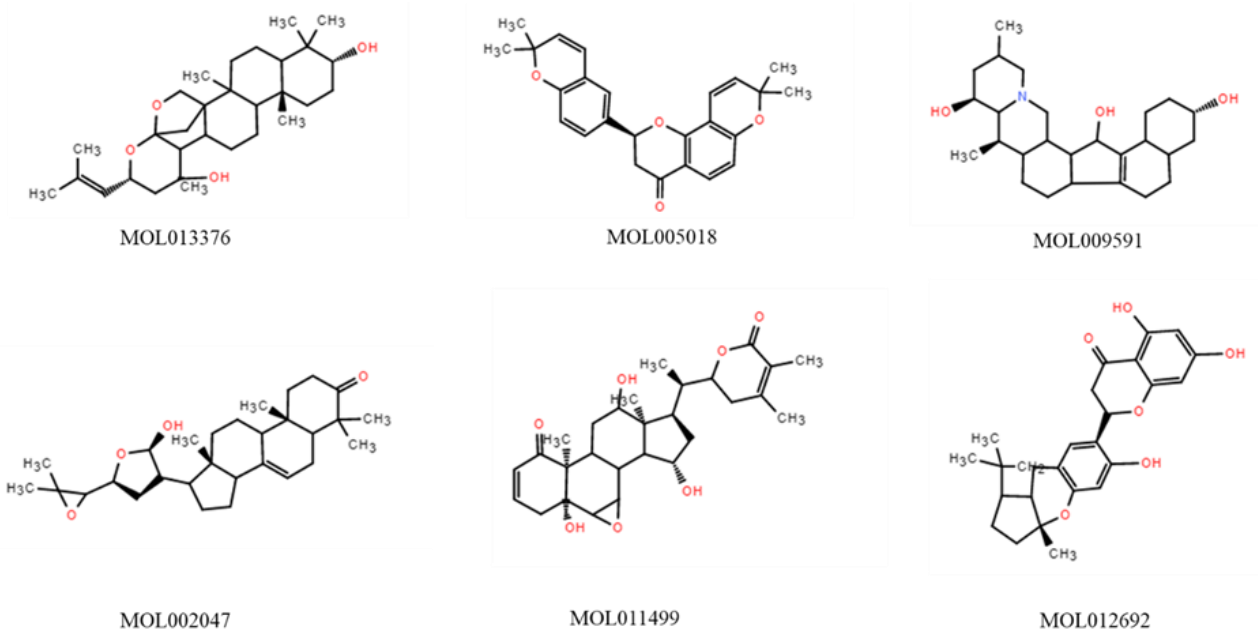
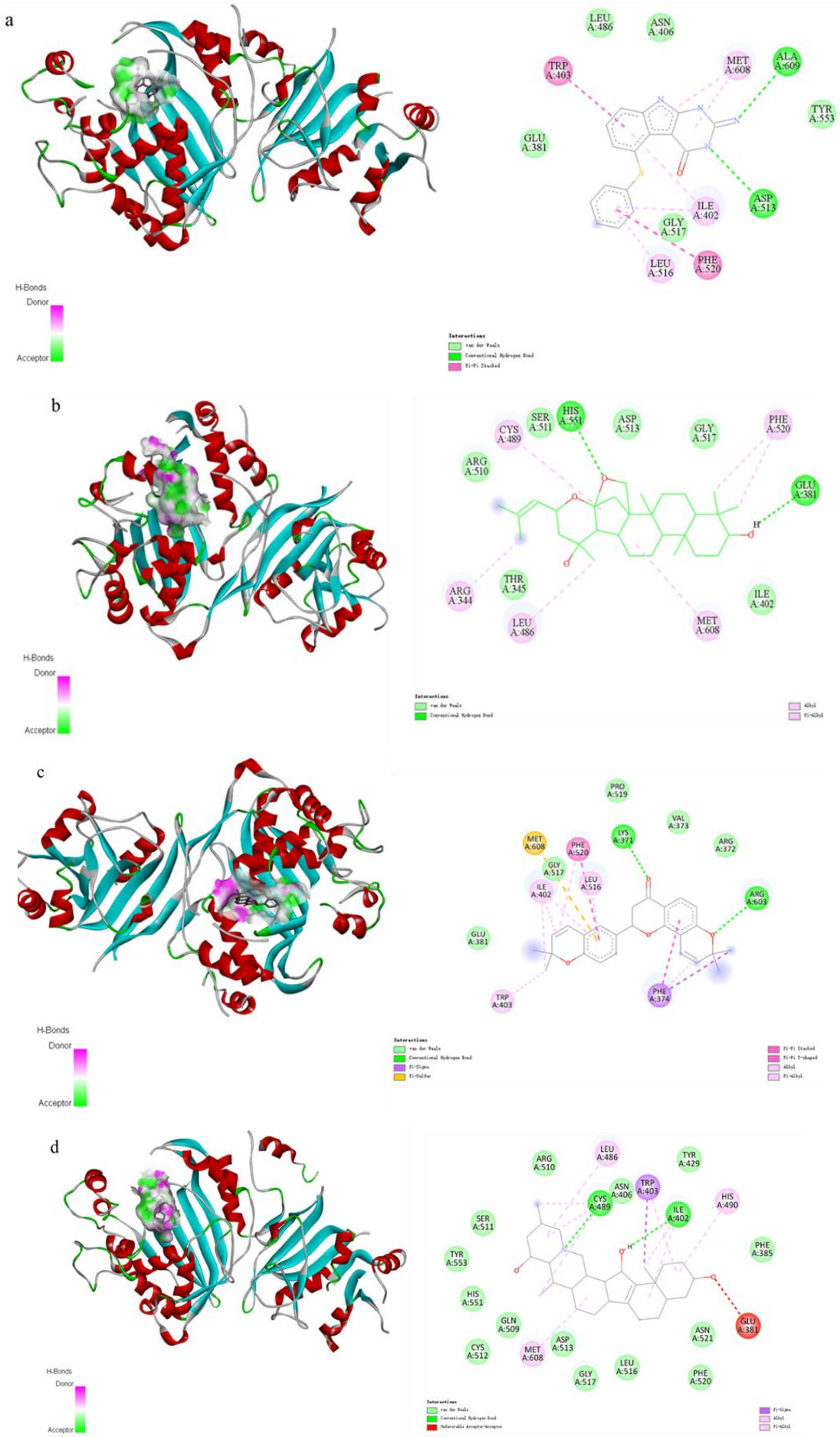


Figure 3. The Structure of active compounds

3.4. Analysis of the Binding Mode of Potential Compounds with the Target

The binding interactions between the active compounds obtained from virtual screening, as well as the original ligand, and DHFR-TS are summarized in Table 2. The binding of the original ligand with DHFR-TS primarily relies on hydrogen bonding and π interactions. In the 2D diagram, it can be observed that the N3 and 2-NH2 positions of the original ligand form hydrogen bonds with ASP513 and ALA609 residues. The benzene ring of the original ligand forms π-π stacking interactions with PHE520 and TRP403, as well as π-alkyl interactions with ILE402, LEU516, and MET608 residues. In addition, the original ligand interacts with GLU381, LEU486, and TRP498 residues via van der Waals forces (Figure 4a). The highest-scoring compound, jujubogenin (sour jujube seed saponin aglycone), binds to DHFR-TS mainly through hydrogen bonds, alkyl interactions, π-alkyl interactions, and van der Waals forces. The hydroxyl and oxygen groups in the molecule form hydrogen bonds with GLU381 and HIS551 residues, while the cyclic structure and side chains interact with CYS489, MET608, PHE520, ARG344, LEU486, ARG510, SER511, ASP513, GLY517, and ILE402 residues through alkyl interactions, π-alkyl interactions, and van der Waals forces (Figure 4b). Xambioona interacts with DHFR-TS primarily through hydrogen bonding and π interactions. The oxygen atom of the molecule forms hydrogen bonds with LYS371 and ARG603 residues, while the benzene ring of the molecule forms π-π stacking interactions with PHE374, PHE520, MET608, and LEU516

residues, as well as π-sulfur interactions. In addition, residues PHE374, ILE402, TRP403, GLU381, GLY517, PRO519, VAL373, and ARG372 form sigma-π interactions, π-alkyl interactions, alkyl interactions, and van der Waals forces with the molecule (Figure 4c). Korsine also binds to DHFR-TS via hydrogen bonds, alkyl interactions, π interactions, and van der Waals forces. The nitrogen and hydroxyl groups of the molecule form hydrogen bonds with CYS489 and ILE402 residues. The cyclic structure and side chains interact with TRP403, HIS490, LEU486, CYS489, MET608, and ILE402 residues through π-alkyl and alkyl interactions. The molecule also forms van der Waals forces with residues such as CYS489, TYR429, PHE385, PHE520, ASN521, LEU516, GLY517, ASP513, GLN509, HIS551, TYR553, SER511, ARG510, ASN406. Unlike other compounds, the oxygen atom of the molecule forms an unfavorable Acceptor-Acceptor interaction with the GLU381 residue (Figure 4d). Melianone binds to the target protein in a similar manner to the other compounds, relying primarily on hydrogen bonds, π interactions, and van der Waals forces. The oxygen atom of the molecule forms hydrogen bonds with LYS371, ASP513, GLN509, and GLY517 residues, while the cyclic structure and side chains interact with ILE402, MET608, LEU486, CYS489, TRP403, VAL373, and LYS371 residues via π-alkyl and alkyl interactions. The PHE520 residue forms a sigma-π interaction with the molecule (Figure 4e). Daturameteline A is a lactone compound derived from Datura. The binding interactions of this molecule with DHFR-TS include hydrogen bonds, alkyl interactions, π interactions, van der Waals forces, and unfavorable Donor-Donor interactions. The oxygen and hydroxyl groups of the molecule form hydrogen bonds with TYR553, ASP513, CYS489, and HIS551 residues, and the cyclic structure and side chains interact with ILE402, PHE520, and LEU516 residues through alkyl and π-alkyl interactions. Additionally, MET608, GLY517, ASN521, HIS490, GLN509, SER511, ARG510, LEU486, THR345, and ALA609 residues form van der Waals interactions with the molecule (Figure 4f). Kuwanon D is a flavonoid compound belonging to the dihydroflavonol class. The binding to the target protein relies mainly on hydrogen bonds, π interactions, alkyl interactions, and van der Waals forces. The oxygen and hydroxyl groups in the molecule form hydrogen bonds with HIS551 and ASP513 residues. The phenyl ring, cyclic structure, and side chains interact with ILE402, CYS489, HIS490, TRP403, TYR429, LEU486, MET608, and PHE520 residues through alkyl interactions, π-alkyl interactions, and sigma-π interactions. Van der Waals interactions are observed with residues such as SER511, THR345, ARG344, PHE411, VAL610, ASN406, ASN521, GLU381, GLY517, LEU516, and ALA609 (Figure 4g).


a
LEU A:486
ASN A:406
TRP A:403
MET A:608
ALA A:609
TYR A:553
GLU A:381
ASP A:513
ILE A:402
GLY A:517
LEU A:516
PHE A:520
H-Bonds
Donor
Acceptor
b
CYS A:489
SER A:511
HIS A:551
ASP A:513
GLY A:517
PHE A:520
ARG A:510
GLU A:381
ARG A:344
THR A:345
ILE A:402
LEU A:486
MET A:608
H-Bonds
Donor
Acceptor
c
PRO A:519
VAL A:373
MET A:608
PHE A:520
LYS A:371
ARG A:372
GLY A:517
ILE A:402
LEU A:516
ARG A:603
GLU A:381
TRP A:403
PHE A:374
H-Bonds
Donor
Acceptor
d
ARG A:510
LEU A:486
TYR A:429
ASN A:406
TRP A:403
HIS A:490
CYS A:489
ILE A:402
SER A:511
PHE A:385
TYR A:553
HIS A:551
GLN A:509
GLU A:381
ASN A:521
CYS A:512
MET A:608
ASP A:513
LEU A:516
GLY A:517
PHE A:520
H-Bonds
Donor
Acceptor

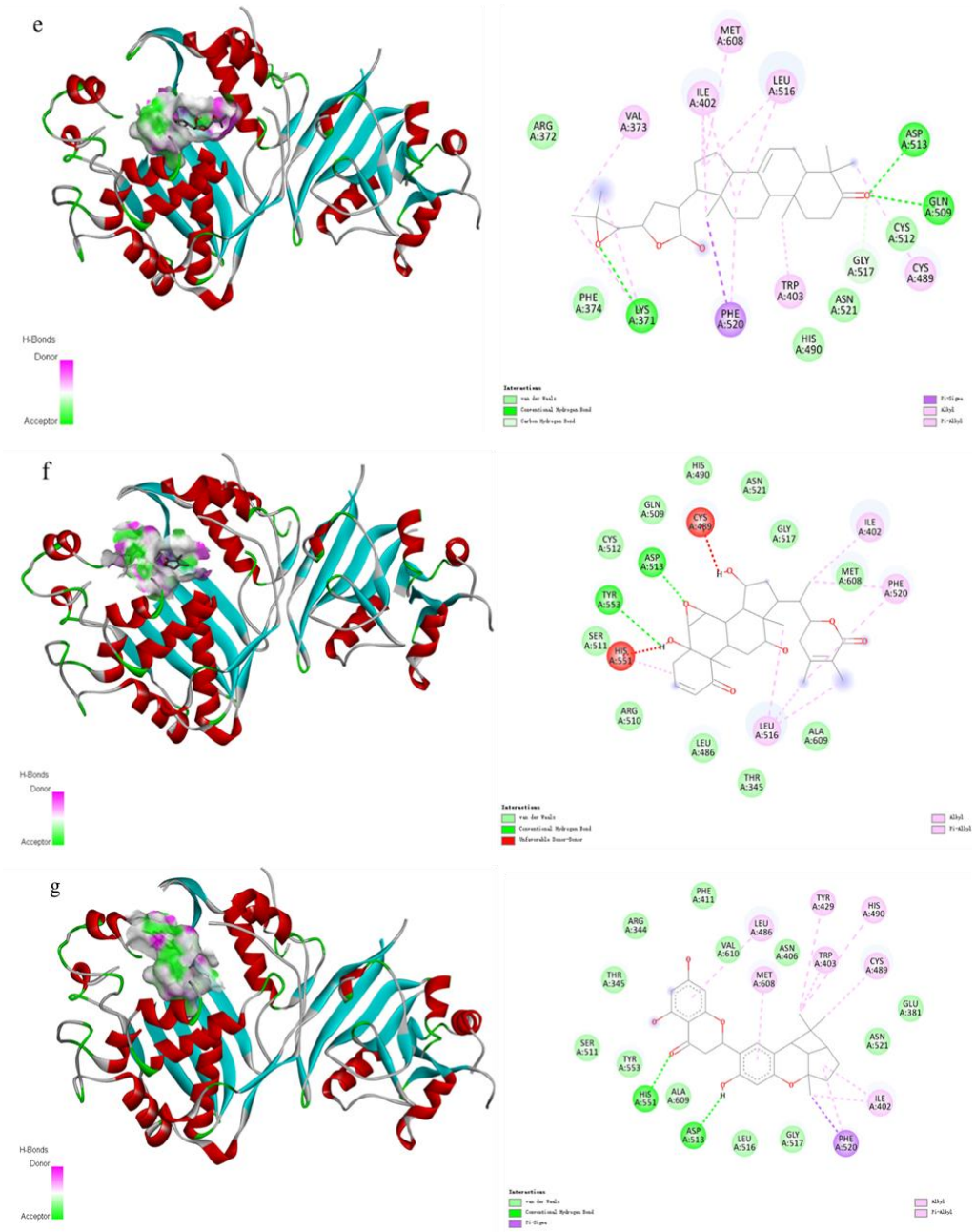


Figure 4. Molecular docking mode between DHFR-TS and potentially active compounds

(a): 1UE (b): MOL013376 (c): MOL005018 (d): MOL009591 (e): MOL002047 (f): MOL011499 (g): MOL012692.

The binding modes of the six active compounds identified through screening are similar to that of the original ligand, primarily involving hydrogen bonds, π interactions, and van der Waals forces. Additionally, the key amino acid residues involved in the interactions with the target protein are conserved. These critical residues and their corresponding binding sites are highly conserved across the DHFR-TS proteins of the five protozoan species. Specifically, the residues ARG344, GLU381, TRP403, ASN406, TYR429, CYS489, HIS490, GLN509, ARG510, SER511, ASP513, PRO519, PHE520, and HIS551 are recognized as essential active-site residues of the DHFR-TS protein [10]. Both the original ligand and the potential active compounds identified in this study interact with these key amino acid residues. Therefore, the compounds identified through this screening could serve as promising inhibitors of the DHFR-TS protein in protozoan parasites.

4. DISCUSSION

Protozoan parasites are single-celled eukaryotic organisms responsible for a range of severe diseases that pose a significant health and economic burden globally [14]. These parasites are typically transmitted through insect vectors, contaminated food, or water, disproportionately affecting tropical and remote regions with limited healthcare resources. Additionally, global climate change has expanded the range of insect vectors, leading to a broader spread of parasitic diseases [15]. The World Health Organization (WHO) has classified several parasitic diseases, such as leishmaniasis and trypanosomiasis, as neglected tropical diseases, impacting nearly one billion people worldwide. Among these, diseases like cryptosporidiosis are also categorized as "Category B biological agents" by the Centers for Disease Control and Prevention (CDC) and the National Institutes of Health (NIH) (CDC, 2023).

Studies have shown that parasites possess remarkable abilities to evolve drug resistance and form resistant cysts, complicating treatment efforts and prompting increased global research focus on combating resistance mechanisms [16]. As a result, there is growing interest in identifying new drug targets and developing novel therapies to treat parasitic infections. The combination of comparative genomics and computational screening techniques has become a powerful approach for identifying novel drug targets and compounds, and is widely applied in parasitic disease research [17]. The International Helminth Genomes Consortium, for example, has conducted large-scale comparative genomics of 81 parasitic worm species to identify gene families associated with parasitism, and to discover potential drug targets for dozens of trematodes and nematodes that infect both humans and livestock. By identifying potential drug targets like proteases, which destroy immune defenses, or enzymes that help parasites digest host tissues during infection, researchers can then use virtual screening methods to find existing drugs that might interact with these targets. This has led to the creation of databases that link active compounds to corresponding parasitic targets [5].

For protozoan parasites like plasmodium species, artemisinin-based combination therapies (ACT) have been highly effective. However, with increasing drug use, resistance has begun to emerge in some areas, diminishing the efficacy of these treatments. As a result, there is an urgent need to identify new drug targets and develop alternative treatments to address resistance. In this context, comparative genomics and computational screening are essential tools for uncovering conserved drug targets and identifying promising compounds for new therapies. This study focuses on combining these methods to identify conserved drug targets in protozoan parasites and to virtually screen traditional Chinese medicine TCMSP databases for potential lead compounds against protozoan infections.

5. CONCLUSION

In this study, we employed comparative genomics to identify the DHFR-TS enzyme as a conserved drug target in protozoan parasites. Using Toxoplasma gondii DHFR-TS as a model, we performed virtual screening with ligands from the TCMSP database of traditional Chinese medicines, employing AutoDock Vina for docking and ADME/Lipinski filters for compound selection. By comparing the binding sites of the selected compounds with the original ligand and assessing the conservation of these sites, we identified six potential inhibitors of protozoan parasites. The binding modes of these compounds with the DHFR-TS target protein are similar to that of the original ligand, suggesting that these compounds may act as DHFR-TS inhibitors and serve as promising lead compounds for the development of novel antiparasitic drugs.

Despite identifying potential lead compounds, the study has limitations. The virtual screening was conducted using semi-flexible docking, which may not fully predict the actual conformation of the target in the biological environment. Therefore, further validation of the docking results through pharmacological and pharmacodynamic testing, as well as X-ray crystallography of ligand-receptor complexes, will be necessary to confirm the therapeutic potential of these compounds.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the financial support from Jinan University College Student's Innovation and Entrepreneurship Training Program (project No. 202310559100).

REFERENCES

[1] Barrett M P, Burchmore R J S, Stich A, et al. The trypanosomiases [J]. The Lancet, 2003, 362(9394): 1469-1480.

[2] Baneth G, Yasur-Landau D, Gilad M, et al. Canine leishmaniosis caused by Leishmania major and Leishmania tropica: comparative findings and serology [J]. Parasites & Vectors, 2017, 10: 1-9.

[3] Xiao J, Yolken R H. Strain hypothesis of Toxoplasma gondii infection on the outcome of human diseases [J]. Acta physiologica, 2015, 213(4): 828-845.

[4] Nonaka M, Murata Y, Takano R, et al. Screening of a library of traditional Chinese medicines to identify anti-malarial compounds and extracts [J]. Malaria Journal, 2018, 17(1): 1-10.

[5] Comparative genomics of the major parasitic worms [J]. Nature genetics, 2019, 51(1): 163-174.

[6] Hardison R C. Comparative genomics [J]. PLoS biology, 2003, 1(2): e58.

[7] Meng X Y, Zhang H X, Mezei M, et al. Molecular docking: a powerful approach for structure-based drug discovery [J]. Current computer-aided drug design, 2011, 7(2): 146-157.

[8] Xu L, Dong Z, Fang L, et al. OrthoVenn2: a web server for whole-genome comparison and annotation of orthologous clusters across multiple species [J]. Nucleic acids research, 2019, 47(W1): W52-W58.

[9] Ru J, Li P, Wang J, et al. TCMSP: a database of systems pharmacology for drug discovery from herbal medicines [J]. Journal of cheminformatics, 2014, 6: 1-6.

[10] Zaware N, Sharma H, Yang J, et al. Discovery of potent and selective inhibitors of Toxoplasma gondii thymidylate synthase for opportunistic infections [J]. ACS medicinal chemistry letters, 2013, 4(12): 1148-1151.

[11] Trott O, Olson A J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading [J]. Journal of computational chemistry, 2010, 31(2): 455-461.

[12] Vugmeyster Y, Harrold J, Xu X. Absorption, distribution, metabolism, and excretion (ADME) studies of biotherapeutics for autoimmune and inflammatory conditions [J]. The AAPS journal, 2012, 14: 714-727.

[13] Lipinski C. Solubility in the design of combinatorial libraries [J]. Analysis and Purification Methods in Combinatorial Chemistry, 2003, 163: 407-434.

[14] Burgess S L, Gilchrist C A, Lynn T C, et al. Parasitic protozoa and interactions with the host intestinal microbiota [J]. Infection and immunity, 2017, 85(8): 101-108.

[15] Short E E, Caminade C, Thomas B N. Climate change contribution to the emergence or re-emergence of parasitic diseases [J]. Infectious Diseases: Research and Treatment, 2017, 10: 1178633617732296.

[16] Schaap P, Schilde C. Encystation: the most prevalent and underinvestigated differentiation pathway of eukaryotes [J]. Microbiology, 2018, 164(5): 727-739.

[17] Striepen B, Kissinger J C. Genomics meets transgenics in search of the elusive Cryptosporidium drug target [J]. Trends in parasitology, 2004, 20(8): 355-358.