Review

In Silico Strategies in Tuberculosis Drug Discovery

Stephani Joy Y. Macalino [1,2], **Junie B. Billones** [2,*], **Voltaire G. Organo** [2] and
Maria Constancia O. Carrillo [2]

1 Chemistry Department, De La Salle University, 2401 Taft Avenue, Manila 0992, Philippines;
   stephanimacalino@gmail.com
2 OVPAA-EIDR Program, "Computer-Aided Discovery of Compounds for the Treatment of Tuberculosis in
   the Philippines", Department of Physical Sciences and Mathematics, College of Arts and Sciences, University
   of the Philippines Manila, Manila 1000, Philippines; vgorgano@up.edu.ph (V.G.O.);
   mtobrerocarrillo@up.edu.ph (M.C.O.C.)
* Correspondence: jbbillones@up.edu.ph; Tel./Fax: +63-2-5772001

check for
updates

Abstract: Tuberculosis (TB) remains a serious threat to global public health, responsible for an estimated 1.5 million mortalities in 2018. While there are available therapeutics for this infection, slow-acting drugs, poor patient compliance, drug toxicity, and drug resistance require the discovery of novel TB drugs. Discovering new and more potent antibiotics that target novel TB protein targets is an attractive strategy towards controlling the global TB epidemic. *In silico* strategies can be applied at multiple stages of the drug discovery paradigm to expedite the identification of novel anti-TB therapeutics. In this paper, we discuss the current TB treatment, emergence of drug resistance, and the effective application of computational tools to the different stages of TB drug discovery when combined with traditional biochemical methods. We will also highlight the strengths and points of improvement in *in silico* TB drug discovery research, as well as possible future perspectives in this field.

Keywords: tuberculosis; druggability; docking; pharmacophore; MD simulation; QSAR; DFT

1. Introduction

In 1882, *Mycobacterium tuberculosis* (*Mtb*) was identified by Robert Koch as the causative agent of tuberculosis (TB), an infectious disease that continuous to be a relevant threat to global public health, especially in low- to middle-income countries. The pathogenesis of TB has several risk factors, including HIV infection, malnutrition, air pollution, type 2 diabetes, alcoholism, and smoking [1–5].

TB is encountered either as latent TB infection (LTBI), which is non-communicable and asymptomatic [6], or active TB, which is communicable and has symptoms such as fever, weight loss, productive cough, and hemoptysis [7]. Active infection is also classified depending on the strain: (1) drug-sensitive, (2) multidrug-resistant TB (MDR-TB), which is resistant to isoniazid and rifampicin, and (3) extensively drug-resistant TB (XDR-TB), which shows resistance to isoniazid, rifampicin, any fluoroquinolone, and aminoglycoside. Around 1.7 billion people are projected to suffer from LTBI and are at risk of progressing into active TB infection [8]. The World Health Organization (WHO) stated that active TB disease can be found in approximately 10 million people and has caused approximately 1.5 million deaths in 2018. An estimated half million individuals have rifampicin-resistant TB (RR-TB), of which 78% had MDR-TB. Furthermore, approximately 6.2% are suggested to have XDR-TB from these MDR cases [8].

2. Current Tuberculosis Management

One of the major challenges in managing TB is the estimated three million 'missing' individuals who have developed active infections but remained undetected or undiagnosed. TB can be deadly if not treated. With the help of conventional regimen, an estimated 58 million infected individuals were saved from 2000 to 2018. Global treatment outcome in 2017 shows a success rate of 85% for new TB cases and 56% for those with drug-resistant TB [8].

2.1. Latent Tuberculosis Infection

Treatment for LTBI are only provided for select groups that have a high risk of transitioning to active TB infection, including HIV-positive patients, people who were exposed to those with active TB, patients undergoing dialysis for end-stage renal disease, taking anti-tumor necrosis factor (TNF) medications, preparing for transplant surgery, or those with silicosis. Depending on whether it is beneficial or not, especially for children below 5 years of age, exposure to patients with active MDR-TB would require personalized treatment regimens and close observation. WHO recommended several different treatment regimens for LTBI, including 3 months of rifapentine and isoniazid, 3–4 months of isoniazid and rifampicin, 3–4 months of rifampicin, and 6–9 months of isoniazid [9,10]. While all these have established efficacy, poor patient compliance continues to be an issue especially with long treatment periods [9–11].

2.2. Active Drug-Sensitive Tuberculosis

In the last several decades, the treatment strategy for active drug-sensitive TB has not changed from the standard regimen of first-line drugs rifampicin, isoniazid, pyrazinamide and ethambutol (Figure 1) for the first 2 months continued by isoniazid and rifampicin for the next 4 months [12,13]. While this treatment procedure is highly efficacious and successful, its long duration primarily leads to poor patient compliance. This has long been an issue in TB management, necessitating monitoring protocols like the directly observed therapy (DOT), wherein a health professional directly supervises each dose intake [14]. Another issue brought about by the prolonged treatment is drug toxicity resulting in numerous adverse effects such as skin rash, gastrointestinal intolerance, neuropathy, arthralgia, increase in liver enzymes, hepatitis, immune thrombocytopaenia, agranulocytosis, haemolysis, renal failure, optic neuritis, and ototoxicity [15,16].
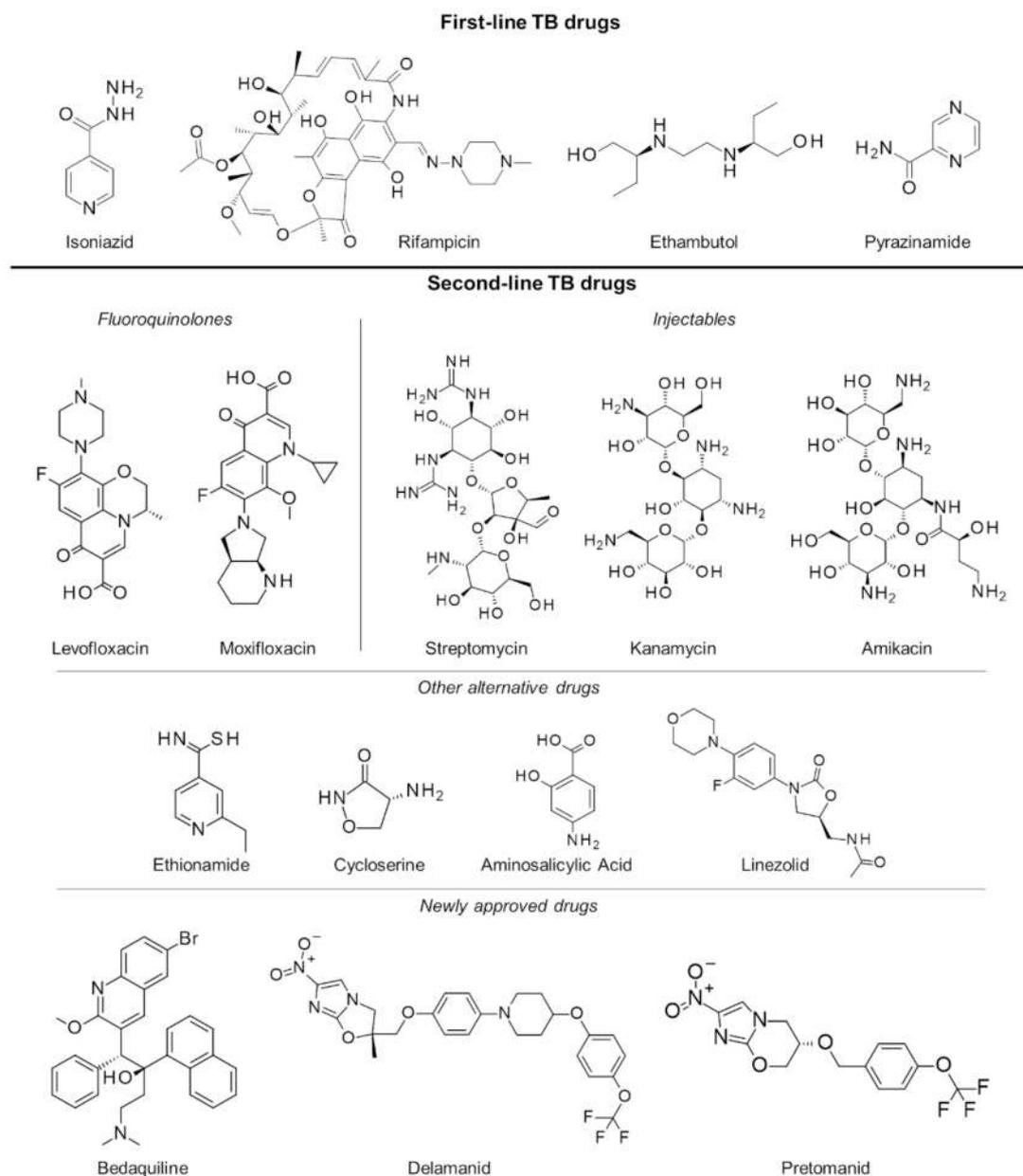
Figure 1. First- and second-line drugs approved for tuberculosis treatment.

2.3. Multiple and Extensively Drug-Resistant Tuberculosis

Failure to complete the full TB regimen leads to disease relapse and drug resistance, which is more challenging to treat. A specific regimen can be designed depending on the resistance profile of the TB strain in a patient [17,18]. These treatments are often of longer duration (18 months or more) and utilize the more expensive second-line drugs (Figure 1) which have uncertain efficacy and high toxicity, resulting in poorer compliance and undesirable outcomes. To mitigate these issues, an updated seven-drug regimen guideline for the treatment of drug-resistant TB lasting 9 to 12 months was released by the WHO last 2016 [19].

With the increasing threat of treatment-resistant TB infection, a number of drugs have been fast-tracked to aid with the efforts in controlling TB worldwide. At the end of 2012, the US Food and Drug Administration (FDA) conferred accelerated approval to the drug bedaquiline for the treatment of resistant TB [20]. Bedaquiline's anti-mycobacterial activity is due to its inhibition of the mycobacterial ATP synthase, a key enzyme in ATP synthesis, resulting in bacterial death. However, its use was

shown to have an increased risk of death, thereby causing concerns about its approval. During clinical trials, roughly 11.4% of patients who took bedaquiline died as compared with 2.5% of those who took placebo treatments [21]. In 2014, the use of delamanid, a nitro-dihydro-imidazooxazole derivative, in the treatment of MDR-TB in adults was given conditional approval by the European Medicines Agency (EMA) [22]. Delamanid inhibits mycolic acid biosynthesis to block the formation of mycobacterial cell wall leading to improved drug permeation and more effective treatment [23]. Just recently, pretomanid in combination with bedaquiline and linezolid has also been approved by the FDA for treatment-resistant TB [24]. Pretomanid is a prodrug activated by nitroreductase, which reduces pretomanid's imidazole ring to generate active metabolites. Specifically, a des-nitro metabolite leads to elevated levels of nitric oxide, which displays antimycobacterial activities due to its work as a poison for bacterial respiration under anaerobic conditions [25]. In aerobic conditions, it works like delaminid by targeting cell wall mycolic acid biosynthesis [26], and while there were several potential targets for this drug, its exact protein target is not yet known [27].

An increasing number of XDR-TB cases, such as in India, China, South Africa, Russia, and in eastern Europe, have proved difficult to treat even with the more intensive drug-resistant TB treatment regimen [18]. Novel therapeutics such as bedaquiline, delamanid, and pretomanid might help in curing these patients, though a suitable treatment regimen still has to be carefully designed. However, there is an additional difficulty in acquiring these drugs, especially in developing countries, resulting in a pool of patients that may remain untreated. Essentially, TB can be cured completely with the use of currently available and newly approved anti-tubercular drugs. However, difficulties in diagnosing and reporting infection, long treatment durations leading to drug toxicity and poor patient compliance, emergence of drug resistant strains, and limited acquisition of required treatment urgently necessitates the discovery and development of newer and effective drugs for TB.

3. Rise of Computer-Aided Drug Design in TB Drug Discovery

The drug discovery paradigm covers a wide range of fields, including biochemistry, chemical and structural biology, chem- and bioinformatics, computational chemistry, physical chemistry, organic synthesis, and others. The whole process entails large investments of time, money, and effort in order to produce promising candidates for the pipeline. Over the years, the drug discovery process for new antitubercular therapeutics have changed due to the increase in biological and chemical data, number of identified and validated targets, and advances in high-throughput screening technologies and software development. Moreover, the progress in data storage capacities, supercomputing powers, and parallel processing in the last several years allowed computer-aided drug design (CADD) to become an integral part of TB pharmaceutical research. This continuing expansion in computing power can soon potentially allow the exploration of the vast chemical space, thought to comprise of approximately 10^{60} organic molecules below 500 Da, in order to identify therapeutically interesting scaffolds [28]. Moreover, the boom in protein structural data, including over 150,000 macromolecular structures found in the Protein Data Bank (PDB, www.rcsb.org) [29], proved beneficial in elucidating important molecular and computational concepts for drug design studies. As with any other disease, TB has been the subject of continuous and numerous drug discovery studies, including thousands of published CADD investigations. Despite this, a paper by Ekins et al. noted gaps in the application of these methods in TB research [30], resulting in the slow output of candidates into the TB drug pipeline despite the apparent need and urgency for this disease. This suggests that more rigorous efforts are needed in TB drug discovery to maximize the advantages provided by computational tools.

Computational or *in silico* methods are knowledge-driven, rationally exploring available data to investigate protein function and design new molecular entities (NMEs) that can effectively regulate its behavior. Computational drug discovery approaches are generally divided into structure—(SBDD) and ligand-based drug design (LBDD), depending on the availability of structural data (Figure 2). However, it has been a common practice to integrate these methods in a complementary manner in order to increase the success rate of current drug discovery projects (Figure 2). SBDD requires the

target's three-dimensional (3D) structure to be able to examine and use the binding pocket for screening and design of suitable ligands, which can then be experimentally validated and optimized. In the absence of protein structural data, LBDD utilizes knowledge gained from a collection of diverse ligands with known activity to create predictive models for hit discovery and lead optimization [31]. Different types of SB and LB strategies, or a combination thereof, can be applied at different stages of TB drug discovery and development in order to alleviate the challenges involved with experimental methods. With the availability of TB genome and proteome, as well as abundant structural data, data mining and docking strategies can be employed for target identification. Virtual screening (VS) can then be applied to pick out the best potential candidates from a database containing millions of molecules for a chosen TB target. After validation of candidates, structure-activity or -property relationship (SAR/SPR) studies can be implemented to understand mechanism of action and ADMET (absorption, distribution, metabolism, excretion, and toxicity) properties in order to design compounds with better activity and pharmacokinetics. Data (both positive and negative results) taken from these investigations can be kept and used for further iteration and method optimization in the design of novel TB compounds. Both commercial and free software and webservers have been developed covering different SBDD and LBDD techniques, some of which are listed in Table 1.
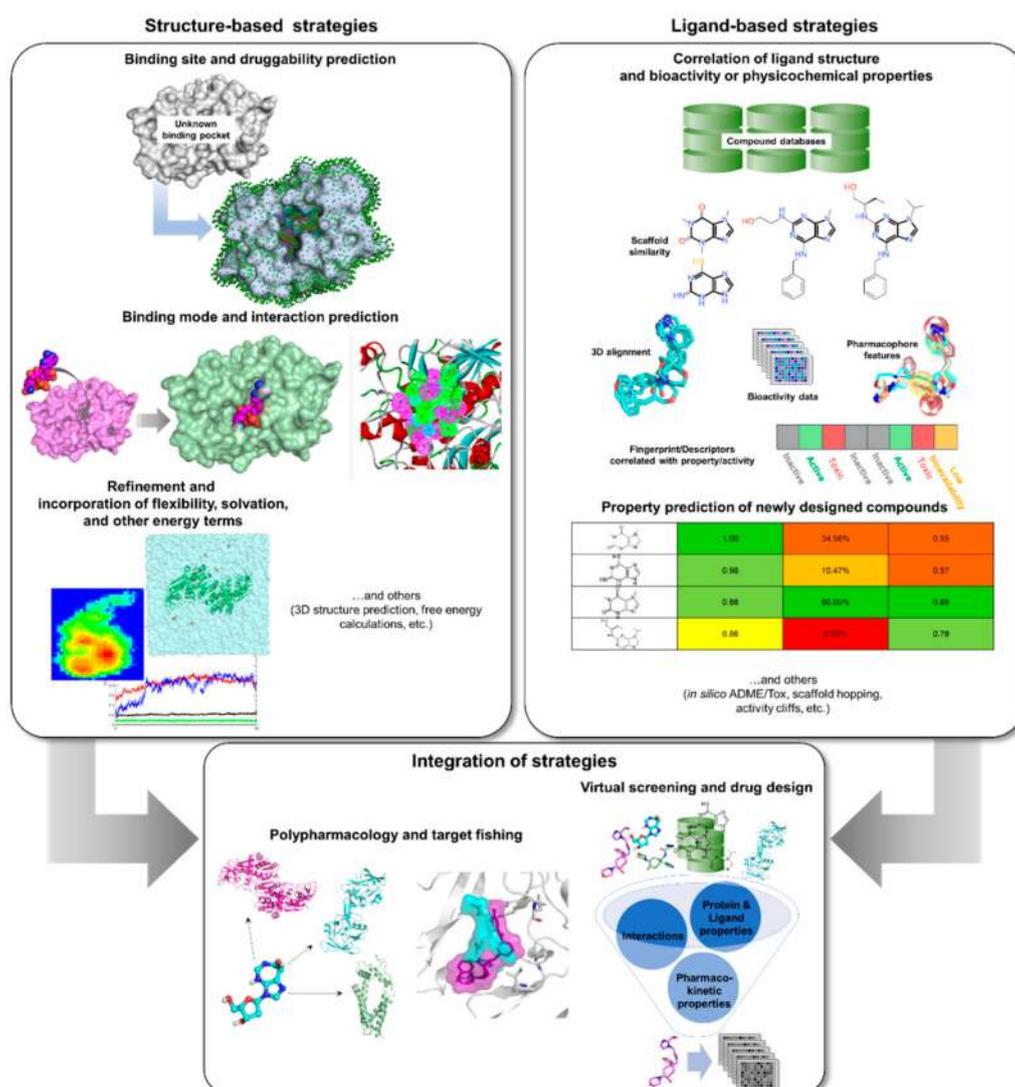

Figure 2. *In silico* tools that can be applied to TB drug design and development.

Table 1. Free and commercially available programs, webservers, and source codes for SBDD and LBDD.

Function	Software/ Webserver Name	Availability	Website
Comparative modeling	SWISS-MODEL [32]	Free webserver	https://swissmodel.expasy.org/
Structural geometry confirmation	MODELLER [33]	Free standalone program for academic license or commercially available through BIOVIA	https://salilab.org/modeller/ https://www.3dsbiovia.com/
	Robetta [34]	Free webserver	http://new.robetta.org/
	Prime [35]	Commercially available through Schrödinger	https://www.schrodinger.com/prime
	I-TASSER [36–41]	Free webserver or standalone program for academic license	https://zhanglab.ccmb.med.umich.edu/I-TASSER/
	HHPred [42–44]	Free webserver	https://toolkit.tuebingen.mpg.de/tools/hhpred
Structural geometry confirmation	PROCHECK [45]	Free webserver and source code	https://www.ebi.ac.uk/thornton-srv/software/PROCHECK/
	ProSA [46]	Free webserver	https://prosa.services.came.sbg.ac.at/prosa.php
	VERIFY3D [47]	Free webserver	https://servicesn.mbi.ucla.edu/Verify3D/
	ERRAT [48]	Free webserver	https://servicesn.mbi.ucla.edu/ERRAT/
Druggability and binding site prediction Druggability and binding site prediction	PockDrug [49]	Free webserver	http://pockdrug.rpbs.univ-paris-diderot.fr/cgi-bin/index.py?page=home
	DoGSiteScorer [50]	Free webserver	https://proteins.plus/
	fpocket [51,52]	Free/open source platform	https://github.com/Discngine/fpocket
	CASTp [53–55]	Free webserver	http://sts.bioe.uic.edu/castp/calculation.html
	PocketQuery [56]	Free webserver	http://pocketquery.csb.pitt.edu/
	PASS [57]	Free/open source platform	http://www.ccl.net/cca/software/UNIX/pass/overview.html
	SiteMap [58]	Commercially available through Schrödinger	https://www.schrodinger.com/sitemap
	ConCavity [59]	Free webserver	https://compbio.cs.princeton.edu/concavity/
	PrankWeb [60]	Free webserver	http://prankweb.cz/
	ProFunc [61]	Free webserver	http://www.ebi.ac.uk/thornton-srv/databases/ProFunc/
Docking, pharmacophore, and virtual screening Docking, pharmacophore, and virtual screening	AutoDock [62] and AutoDock Vina [63]	Free standalone program	http://autodock.scripps.edu/
	DOCK [64]	Free/open source platform	http://dock.compbio.ucsf.edu/
	GOLD [65]	Commercially available through CCDC	https://www.ccdc.cam.ac.uk/solutions/csd-discovery/components/gold/
	Glide [66]	Commercially available through Schrödinger	https://www.schrodinger.com/glide/
	Induced Fit [67]	Commercially available through Schrödinger	https://www.schrodinger.com/induced-fit
	FlexX [68]	Commercially available through BioSolveIT	https://www.biosolveit.de/flexx/index.html
	RosettaLigand [69]	Free/open source platform for academic license	https://www.rosettacommons.org/software
	CDOCKER [70]	Commercially available through BIOVIA	https://www.3dsbiovia.com/
	SwissDock [71,72]	Free webserver	http://www.swissdock.ch/docking
	Pharmer [73]	Free/open source platform	http://smoothdock.ccbb.pitt.edu/pharmer/
	CATALYST [74]	Commercially available through BIOVIA	https://www.3dsbiovia.com/products/collaborative-science/biovia-discovery-studio/pharmacophore-and-ligand-based-design.html
	PharmGist [75]	Free webserver	http://bioinfo3d.cs.tau.ac.il/pharma/php.php
	LigandScout [76]	Commercially available through Inte:Ligand	http://www.inteligand.com/ligandscout/
	SwissSimilarity [77]	Free webserver	http://www.swisssimilarity.ch/

Table 1. *Cont.*

Function	Software/ Webserver Name	Availability	Website
	LEA3D [78]	Free webserver	https://chemoinfo.ipmc.cnrs.fr/LEA3D/index.html
	PyRx [79]	Free (no support) or commercially available	https://pyrx.sourceforge.io/
	Phase [80]	Commercially available through Schrödinger	https://www.schrodinger.com/phase
Molecular Dynamics	AMBER [81,82]	Commercially available	https://ambermd.org/
	CHARMM [83]	Free or commercially available through CHARMM or BIOVIA	http://charmm.chemistry.harvard.edu/ https://www.3dsbiovia.com/products/collaborative-science/biovia-discovery-studio/simulations.html
	CHARMMing [84]	Free webserver	https://www.charmming.org/charmming
	GROMACS [85,86]	Free/open source platform	http://www.gromacs.org/
	NAMD [87]	Free/open source platform	https://www.ks.uiuc.edu/Research/namd/
	Desmond [88]	Commercially available through Schrödinger	https://www.schrodinger.com/desmond
	SwissParam [89]	Free webserver	http://www.swissparam.ch/
	CHARMM-GUI [90]	Free webserver	http://www.charmm-gui.org/
	ParamChem CGenFF [91–93]	Free webserver	https://cgenff.umaryland.edu/
	VMD [94]	Free/open source platform	https://www.ks.uiuc.edu/Research/vmd/
Molecular Descriptors, Fingerprints, and Quantitative Structure-Activity Relationship	Dragon [95]	Commercially available through Talete	http://www.talete.mi.it/products/dragon_description.htm
	E-Dragon [96]	Free webserver	http://146.107.217.178/lab/edragon/start.html
	Canvas [97]	Commercially available through Schrödinger	https://www.schrodinger.com/canvas
	RDKit [98]	Free/open source platform	https://www.rdkit.org/docs/source/rdkit.ML.Descriptors.MoleculeDescriptors.html
	PyDescriptor [99]	Free/open source platform	https://ochem.eu/home/show.do
	Mordred [100]	Free/open source platform	https://github.com/mordred-descriptor/mordred
	Open3DQSAR [101]	Free/open source platform	http://open3dqsar.sourceforge.net/?Home
	ChemSAR [102]	Free webserver	http://chemsar.scbdd.com/
	SeeSAR [103]	Commercially available through BioSolveIT	https://www.biosolveit.de/SeeSAR/
Pharmacokinetic properties	QikProp [104]	Commercially available through Schrödinger	https://www.schrodinger.com/qikprop
	ADMET Predictor [105]	Commercially available through SimulationsPlus, Inc.	https://www.simulations-plus.com/software/overview/
	ACD Percepta [106]	Commercially available through ACD/Labs	https://www.acdlabs.com/products/percepta/index.php
	FAF-Drugs4 [107]	Free webserver	http://fafdrugs4.mti.univ-paris-diderot.fr/
	PatchSearch [108]	Free webserver	http://mobyle.rpbs.univ-paris-diderot.fr/cgi-bin/portal.py#forms::PatchSearch
	TOPKAT [109] and ADMET [110]	Commercially available through BIOVIA	https://www.3dsbiovia.com/products/collaborative-science/biovia-discovery-studio/qsar-admet-and-predictive-toxicology.html
	PASS Online [111]	Free webserver or commercially available standalone program	http://pharmaexpert.ru/Passonline/index.php
	SwissADME [112]	Free webserver	http://www.swissadme.ch/
	MetaSite [113]	Commercially available through Molecular Discovery	https://www.moldiscovery.com/software/metasite/
	ToxPredict [114]	Free webserver	https://apps.ideaconsult.net/ToxPredict#
	VirtualToxLab [115–118]	Free standalone software	http://www.biograf.ch/index.php?id=home
	admetSAR [119–121]	Free webserver	http://lmmd.ecust.edu.cn/admetsar1/home/
	MetaTox [122,123]	Free webserver	http://way2drug.com/mg2/

More available tools and detailed descriptions for the programs and servers can be found at https://www.click2drug.org/.

3.1. Databases

The era of big data has greatly affected the current drug discovery paradigm through innovations in data storage, management, and mining. Moreover, drastic cost reductions in sequencing technologies allowed the study of multi-omics (e.g., genomics, transcriptomics, proteomics, and metabolomics) for several species including *M. tuberculosis* [124–129]. In order to take advantage of the benefits provided by SB and LB techniques, these biological and/or chemical data must be acquired for analysis via numerous publicly accessible databases on the internet [130–132]. Given that TB is an old disease, vast amounts of data points have already been gathered and are waiting to be used in the fight against this infection.

One of the most extensive and widely-used protein information resource is UniProt (https://www.uniprot.org/) [133], which consists of annotations from several other databases for protein function, omics, and structural data. More specific to TB, the TB Database (http://tbdb.bu.edu/tbdb_sysbio/MultiHome.html) [134,135] contains information on mycobacterium genomes, genes, gene expression correlation, gene epitopes, and experimental and computational models of TB molecular pathways. Alternatively, genomic and proteomic data for various pathogenic mycobacteria can also be found in the Mycobrowser (https://mycobrowser.epfl.ch/) [136], which is linked to UniProt for mycobacterium protein information. On the other hand, patient clinical data is provided by the TB Portals (https://tbportals.niaid.nih.gov/) [137], which is an open-access platform containing socioeconomic, geographic, clinical, laboratory, radiological, and genomic data from patients infected with drug-resistant TB, from the National Institute of Allergy and Infectious Diseases (NIAID) in collaboration with data scientists and clinicians and scientists from countries suffering from heavy TB burden.

Advances in structural and computational biology techniques led to the surge in structural data, resulting in thousands of three-dimensional protein structures generated from X-ray crystallography, nuclear magnetic resonance (NMR), cryo-electron microscopy (EM), homology modeling, and molecular dynamics (MD) simulations. Data from these experiments are customarily deposited to structure databases such as PDB [29], PDBsum [138], etc. Associated with this, the size of the virtual chemical space [28] and improvements in combinatorial chemistry [139] also permitted the availability of chemical libraries (Table 2).

Table 2. Publicly available compound libraries.

Database	Size (Approximate)	Website
GDB-17 [140]	166 billion	http://gdb.unibe.ch/
Enamine REAL [141]	700 million	https://enamine.net/
PubChem [131]	97 million	https://pubchem.ncbi.nlm.nih.gov/
ChemSpider [142]	77 million	http://www.chemspider.com/
ZINC [143]	230 million	http://zinc.docking.org/
ChEMBL [144]	1.9 million	https://www.ebi.ac.uk/chembl/
NCI [145]	460,000	https://cactus.nci.nih.gov/download/roadmap/

Protein subcellular localization databases are also available for study, such as eSLDB (eukaryotic Subcellular Localization database) for general eukaryotes [146], LOCATE for human and other mammals [147], and PSORTdb for bacteria and archaea [148]. Lead optimization and drug repurposing researches can also benefit from protein binding databases like ReLiBase, which consists of interaction information for receptor-ligand complexes from PDB [149], BindingDB, which describes interactions and affinity information between protein target and drug-like molecules [150,151], and Database of Interacting Proteins (DIP) [152], Biological General Repository for Interaction Datasets (BioGRID) [153] and Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) [154], which contains data on protein-protein interactions.

3.2. Structure-Based Tools

3.2.1. Comparative Modeling, Binding Site Prediction, and Druggability

Employment of SBDD tools (Figure 2 and Table 1) require not only the availability of 3D structural data but also information on its druggability and potential binding sites. In the absence of structural data obtained from experiments, such as X-ray crystallography and NMR, a computational model can also be generated either through homology modeling or protein threading techniques (Table 1), which are well-established methods in protein comparative modeling. Homology modeling entails the use of a structural template with suitably similar sequence as the target protein [155]. The most critical stage of any homology modeling procedure is the initial sequence alignment. While there are numerous bioinformatics tools available for this, such as NCBI Blast [156–158], COBALT [159], Clustal Omega [160], KAlign [161], etc., manual inspection and modification of the alignment is crucial, especially if a researcher's knowledge about specific protein folds and domains need to be further incorporated. Then, the secondary structures (i.e., alpha helices, beta strands, loops, etc.) are copied from the template based on the final sequence alignment in order to approximate the target structure. The final model is then refined through minimization or MD and its stereochemical quality is checked using tools like those listed in Table 1 until the structure has improved and is suitable for further computational studies.

Druggability is the capacity of a protein target to be modulated by a ligand. It is important to characterize this property as it helps avoid intractable proteins and allows the identification and prioritization of significant targets. Some predictive approaches that include this property are listed in Table 1. A druggability database, the Druggable Cavity Directory (DCD), is also publicly available to allow researchers to submit protein pocket and druggability information, which is later verified and made available to other researchers [162]. After ascertaining that a given target is indeed tractable, binding pocket information should be acquired either from protein structures complexed with natural substrates or known inhibitors, or from mutational data distinguishing key interaction residues. Ideally, a binding pocket is a concave area in the receptor that is characterized by chemical features with which a ligand can desirably interact to attain the required receptor behavior (e.g., inhibition or activation) [163]. However, if binding information is unknown, several *in silico* methods and webservers (Table 1) are available to identify potential receptor binding sites from a given structure and have been described in detail elsewhere [31]. Otherwise, a number of studies have also used 'blind' docking [164], wherein the whole protein is set as the binding site, allowing ligands to freely bind anywhere in the structure in the hopes of finding a suitable pocket.

It is also prudent to remember that other potential binding sites may be present on the target surface, i.e., allosteric sites. Conventional drug discovery efforts often target the primary (orthosteric) binding site to block substrate binding. But, as in the case of uncompetitive and noncompetitive inhibitors, ligands can also allosterically modulate activity within the protein structure. Such is the case in the study done by Shi and colleagues, where they identified a second druggable binding site in *Mtb* UDP-galactopyranose mutase (UGM) [165]. NMR and kinetics studies classified MS-208, a known *Mtb*UGM inhibitor, as a noncompetitive/mixed inhibitor and therefore binds in another site in the enzyme to allosterically affect substrate binding. Blind docking in AutoDock Vina [63] was performed to identify possible allosteric sites for MS-208. Two regions, A- and S-site, were initially identified and docked complexes featuring binding to either sites were further subjected to simulation studies using Amber [81]. The A-site-bound structure exhibited the most stable complex formation with excellent interaction energy, as well as the most number of contacts, suggesting that this is the second druggable binding site in *Mtb*UGM [165].

3.2.2. Pharmacophore Modeling and Molecular Docking

Virtual screening is one of the most popular *in silico* drug discovery approaches as it allows researchers to quickly extract data from unexplored chemical space in a cost-effective manner. It has

become customary to complement high-throughput screening with VS for the prioritization and identification of novel ligands with the most potential as starting points for drug discovery efforts [166]. Two of the most common VS methods are pharmacophore modeling [167] and docking [168].

Pharmacophore models can be generated from a receptor alone or a receptor-ligand complex. Previously, due to lack of protein structural data, ligand-based pharmacophore has been more customarily used (see Section 3.3.2). A pharmacophore is a group of geometrically-mapped chemical features, such as *H*-bond donors and acceptors, hydrophobicity, and ionizable groups, required for optimal interactions to elicit a response between a receptor target and its partner molecule [169,170]. In line with 3D-mapping in a pharmacophore, exclusion volumes can also be included in the model to incorporate binding site shape [171]. After generating a 3D model, large chemical databases can be efficiently searched for candidates that match pharmacophore elements. However, the pharmacophore database screen only provides a fit score that cannot be translated as affinity. The fit score weighs the alignment quality between the ligand substructures and center of model features. Weights and penalties can also be employed for features deemed significant to activity [171]. These models can also be used for scaffold hopping, allowing for the discovery of novel chemotypes based on fit of interaction and geometric characteristics [171,172].

Another well-established *in silico* method is docking, which can be used to facilitate the investigation of how ligands can fit and complement receptor binding pocket features in order to modulate its activity [168]. Numerous docking methods (Table 1) have been developed and comparative studies and detailed reviews about these have been published elsewhere [173–177]. In the early days of computational drug discovery, docking was developed to be able to predict the bioactive conformation within a set of docking results. However, protein-ligand interactions need to be evaluated using a scoring function to find the best pose using estimated affinity, distinguish actives from inactives, and prioritize candidates for further testing and optimization. This was soon discovered to be the most challenging part of docking due to approximations applied to other crucial factors, such as protein flexibility, solvent involvement, and system entropy [168]. Despite advances in scoring functions through increased understanding of protein-ligand interactions, it is difficult to handle all these aspects while still maintaining method efficiency. Moreover, scoring functions often depend on the protein families and ligand sets from which it was generated and validated [31,178]. And while there is currently no ideal scoring function that can be utilized across all druggable targets, implementation of method validation before starting any VS project establishes whether a chosen docking method and scoring function can be applicable or not [171].

Both pharmacophore modeling and docking have been applied in combination with other *in silico* tools for the identification of novel antimycobacterial agents. Pharmacophore screening followed by docking can be employed as complementary screening tools, resulting in faster processing and more optimized results. A recent study published by our group exemplifies both pharmacophore-based and docking VS by targeting *Mtb* 7,8-diaminopelargonic acid aminotransferase (BioA), an important enzyme in its lipid biosynthesis pathway with no corresponding human ortholog [179]. A receptor-based pharmacophore was generated in Discovery Studio [180] using the BioA structure, characterizing 25 functionalities (nine hydrophobic, nine H-bond donors, and seven *H*-bond acceptors), and was employed to screen 4.5 million compounds from the Enamine REAL database. Compounds with good pharmacophore fit, as well as the co-crystallized inhibitor, were subsequently docked to the BioA protein via CDOCKER [70] and ligands with better binding energy values than the known inhibitor were chosen for the TOPKAT protocol [109] to filter out potentially toxic compounds. This step-by-step screening led to the identification of 45 virtual hits, 17 of which were available for purchasing and validation. Whole-cell assay was performed to eliminate compounds that cannot penetrate the distinctive thick, waxy lipid layer of the mycobacterium, identifying compound **7** ((*Z*)-*N*-(2-isopropoxyphenyl)-2-oxo-2-((3-(trifluoromethyl)-cyclohexyl)amino)acetimidic acid) as a potential BioA inhibitor with a minimum inhibitory concentration of ~25 μM [179].

Throughout the years, various improvements have been harnessed to enhance ranking performance, such as rescoring or consensus scoring. Given the different strengths and limitations of each scoring function, rescoring with the help of a separate scoring function not used in a docking study provides users with a different perspective for selection of final hits. For instance, a faster scoring function can be employed for pose prediction while another one is used for affinity prediction and ranking [171]. Currently, consensus scoring is more commonly used for docking studies and has already been examined by several groups in the last couple of decades as an improved protocol for finding potential hits [181–185]. This strategy aims to characterize the intricacy of target molecular recognition based on various energy functions which can be covered by several scoring schemes, resulting in decrease in false positives [184]. However, there is also a risk of rejecting true positives, which have favorable scores in only 1 function used. Thus, it is also imperative to validate a consensus scoring workflow against specific targets [171]. An exemplary case features salicylate synthase MbtI, a critical enzyme in the biosynthesis of siderophore mycobactins, which is used by *Mtb* to chelate iron required for growth and survival in the host. Absence of siderophores prevents bacterial growth in the persistent state after engulfment by macrophages [186]. Previously identified MbtI inhibitors include those based on the MbtI reaction intermediate isochorismate [187], benzimidazole-2-thione [188], and chromane scaffolds [189]. Chiarelli et al. [190] discovered furan-based MbtI inhibitors through structure-based pharmacophore and consensus docking. The pharmacophore model was generated from important binding features in the MbtI-inhibitor complex, including interaction with the conserved Y385, lipophilic interactions, and ionizable interaction with the Mg^{2+} ion. Screening of 1.5 M compounds from Enamine [141] led to over 2,000 pharmacophore hits which were subjected to consensus docking. Docking methods including AutoDock [62], AutoDock Vina [63], DOCK [64], FRED [191], GOLD [65] (comprising four scoring functions), and PLANT [192] were first evaluated using several MbtI-inhibitor complexes to identify which methods are most reliable. GOLD and PLANT were employed for the consensus procedure of the pharmacophore hits, and those with similar binding modes and consistent scores across all scoring functions were further examined if the docked conformation still matched the 3D arrangement of the pharmacophore model. From these, five virtual hits progressed to bioassays, wherein two compounds showed potent MbtI inhibitory activity. MD simulation was additionally applied to study enzyme-ligand interactions and provide information for further optimization. The furan scaffold from the more potent hit was used as a starting point for lead optimization, resulting in a candidate with promising activity against MbtI and suitable antimycobacterial activity [190].

Inclusion of limited protein flexibility, such as in the binding pocket, while still maintaining efficiency has been considered in methods like induced fit docking (IFD) and ensemble docking. IFD incorporates the principle that ligand binding induces changes in residue side chain conformations within the specified pocket, thereby inciting tighter binding with the receptor [193,194]. However, backbone movement should also be considered as it increases the accuracy of side chain positioning and orientations [31]. An example of this is shown in Figure 3, in which *Mtb* InhA exhibits backbone and side chain conformational differences between its apo (PDB ID: 4DRE) [195], fatty acyl substrate-bound (PDB ID: 1BVR) [196] and isoniazid (INH)-bound (PDB ID: 4TRO) structures [197]. These movements can change the binding pocket shape and volume, potentially affecting VS results. The utilization of several experimentally—(X-ray crystallography, NMR, or cryo-electron microscopy) or computationally-derived (MD trajectory) protein structures to integrate both backbone and side chain movements has also been employed in ensemble docking [198]. Structural ensembles provide better reproducibility of experimental conditions as rigid protein structures, such as those obtained from X-ray crystallography experiments, only provide a snapshot of a dynamic ensemble of conformations.
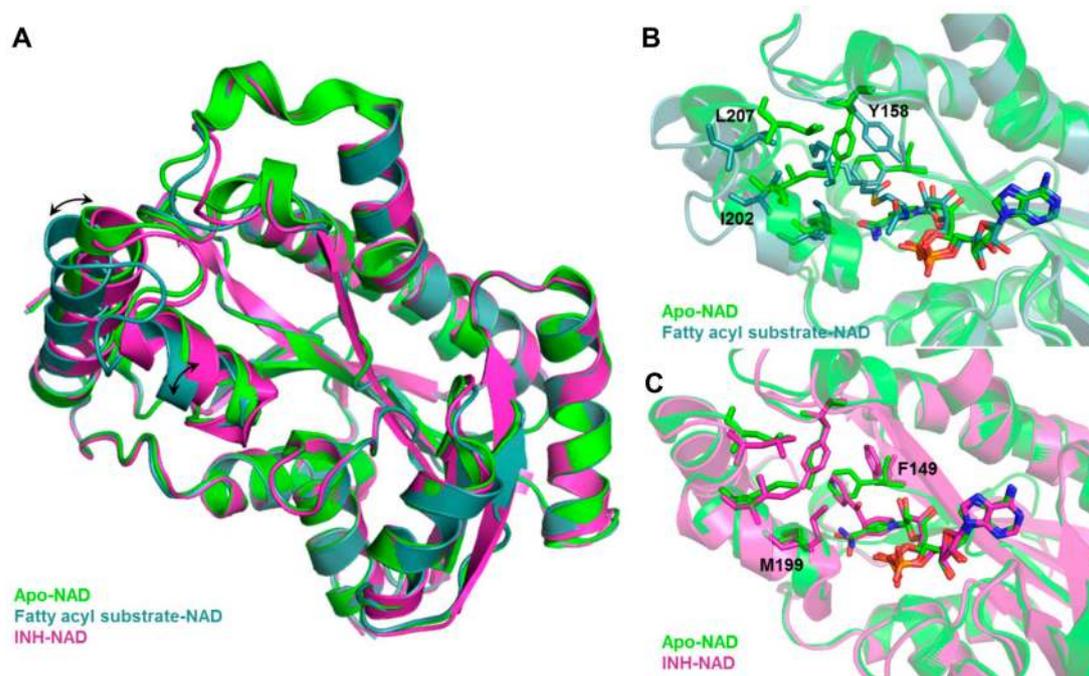
Figure 3. Backbone and sidechain flexibility shown by Mtb InhA, apo vs. fatty acyl-bound vs. INH-bound. (**A**) Structural overlay of apo, fatty acyl-bound, INH-bound Mtb InhA shows backbone movement upon substrate (fatty acyl) binding. Binding site comparison of (**B**) fatty acyl-bound and (**C**) INH-bound vs. apo Mtb InhA structure shows distinct changes in residue side chain positions and conformations. Black arrows indicate movement of alpha helices, side chains that showed large conformational change upon fatty acyl or INH binding are labelled in black.

Various docking approaches were applied by Brindha et al. [199] for drug repurposing against *Mtb* murE, which is an attractive target due to its significance in the peptidoglycan biosynthesis of tuberculosis bacteria and lack of eukaryotic homolog. VS of compounds from DrugBank [200] was first implemented through the parallel use of Glide Standard Precision (SP) [66] and AutoDock Vina [63]. To improve prioritization of compounds through the incorporation of binding site flexibility, common hits from both methods were further subjected to IFD [67]. Final rankings were done using Glide eXtra Precision (XP) scoring and AutoDock Vina binding energy prediction, resulting in 17 common top hits identified as repurposed antitubercular drug candidates [199]. In another example, ensemble docking using three enzyme structures was performed to better elucidate ligand binding interactions, especially due to the binding site flexibility of the *Mtb* Type II dehydroquinase (MtDHQase), an essential virulence factor in TB [201]. Conformation of key residues were determined by analyzing superimposed MtDHQase structures and rotamer distribution of each residue from the penultimate rotamer library [202]. The benzene sulfonamide containing compound with the best activity, a Schaeffer's acid amide, was docked using GOLD [65] and scored using ChemPLP [203]. Varying the side chain flexibility during the docking procedure led to the identification of residues that are required to transition the binding site into an open conformation, which is the preferred conformation of the inhibitor to display its activity. Along with the interaction and flexibility data, Schaeffer's acid amide can be optimized into a potent antitubercular therapeutic compound [201].

3.2.3. Molecular Dynamics

The availability of 3D structural information has greatly helped in structure-based drug design by presenting atomic-level insights into molecular interactions. Nonetheless, these provide only partial interpretations of biomolecular structures, as well as related aspects of molecular recognition and binding. In physiological conditions, proteins frequently undergo conformational changes upon

binding with a partner, such as a small molecule, peptide, or another protein, to perform a specific function. At times, these transformations only involve side chain conformations and small to medium movements in the backbone. However, there are cases in which significant deviations are seen in the overall protein fold and/or subunit arrangement [204–207].

Molecular dynamics simulation, a method that was first developed in the 70s [208], can be employed to analyze these protein dynamics and study the binding energy landscape. The availability of MD platforms (Table 1) allowed for the routine assimilation of simulation studies for systems containing ~50,000–100,000 atoms. The investigation of even larger systems is made possible using graphics processing units (GPUs), which are high-performance processors that can support heavy computational load, and high-performance computing (HPC) technologies featuring messaging passing interface (MPI), a system which employs multiple cores in parallel to distribute computational load and reduce the time required for simulation [206,209]. Popular MD packages have been adapted for these tools, and while MD simulation projects commonly use a combination of both, the speedy development of more advanced GPUs increasingly allows for the use of personal workstations [206].

To start an MD simulation, a 3D protein structure of the required system (e.g., apo protein, protein-ligand or protein-protein complex) must be obtained experimentally or through homology modeling, and represented based on the duration and details of study [206,210]. Another critical aspect of system preparation in MD is the solvent model, which can be explicit or implicit. Explicit solvent is more frequently used due to its simplicity and its proficiency in recovering native solvent effects to protein structure [211]. However, large system size resulting from this model makes conformational sampling challenging. To speed up conformational sampling, an implicit solvent model can be generated by adding approximations to the system, but this may affect the free energy landscapes [212]. Once a solvent model has been chosen, the next step is to select an appropriate force field, which is used to define the forces acting on every atom in the system and to calculate the potential energy within the molecular structure. While there are numerous force fields that have been and are still being developed and improved, some of the most popular force fields applied in simulation systems are currently CHARMM [213], Amber [214], GROMOS [215], and OPLS-AA [216,217]. Different force fields use different parameterization to characterize atomistic molecular simulations, distinguishing their applicability in atomistic molecular simulations of diverse target structures and systems, but are often equivalent [206]. To ensure efficiency while keeping the calculations accurate, simple molecular representations in force fields include springs depicting bond length and angles, periodic functions depicting rotations and Lennard-Jones potentials, and Coulomb's law characterizing van der Waals and electrostatic interactions within the system [206]. Newton's law of motion is then employed for the computation of accelerations and velocities during atom movement. Minimization and equilibration are typically performed ahead of a production run to adjust the system to the applied force field, relax steric clashes, and to stabilize system temperature and pressure. Once the prepared system is correctly minimized and equilibrated, the production run can be performed for a suitable amount of time (ps, ns, µs, etc.) depending on research needs (Figure 4). A timestep of 1 or 2 fs is frequently used for atomistic MD simulations [206].

Figure 4. Typical molecular dynamics simulation workflow.

After obtaining simulation trajectories, this information can be used for additional analyses, including but not limited to: (1) verifying stability through root-mean-square deviation (RMSD), root-mean-square fluctuation (RMSF), or radius of gyration (RoG) [218,219], which can identify critical components for protein flexibility, (2) investigating protein structural or energy networks through network analysis [220], which can pinpoint residues that are pivotal in allosteric or long-range communications within the protein, (3) studying protein energy landscape by mapping trajectories [221],

providing information about protein folding and function, as well as most stable populations within a given trajectory. Findings from each analysis provide crucial information that would otherwise be imperceptible with other techniques, thereby increasing our understanding of a given system.

With current computational capabilities, simulation times are nearly of biological relevance, allowing researchers to observe biological events, such as allosteric regulation, transient protein changes and binding, and enzyme catalysis. A number of MD studies have already been done for the structural and functional elucidation of validated TB targets and design new antitubercular agents [222–224]. One such study investigated the differences of inhibitor binding against wild-type and mutant structures of InhA, a very well-known TB target [223]. Mutations for this protein led to lower affinity for its co-factor, NADH, resulting in isoniazid resistance. MD simulations of the wild-type and mutant structures of InhA bound to NADH were performed using Amber [81] to understand the underlying aspects affected by structural mutations. Schroeder and colleagues found that mutations in the glycine-rich loop (I21V and I16T) disturbed the NADH binding conformation, specifically that of its pyrophosphate moiety, and decreased its direct and indirect *H*-bond contacts within the binding pocket. Isoniazid requires the formation of a covalent adduct with NADH within the InhA binding site. Changes in binding interactions and conformation of NADH can negatively affect this, hence, contributing to isoniazid resistance [223]. A very recent MD study for TB drug discovery using Amber [81] has been published by Cruz et al. [224], wherein the binding mechanisms of Tam1 and its analogs against polyketide synthase 13 (Pks13), an enzyme that carries out the final step in mycolic acid biosynthesis [225], were investigated to obtain insights that can aid in the design of new antitubercular agents [224]. Fluctuation analysis revealed distinct flexibility in the protein lid domain of *Mtb* Pks13, which was decreased upon ligand binding, suggesting that residues from this domain are critical for ligand interaction. Binding free energy calculations from trajectory data agreed with experimental data, identifying Tam16 as the most potent of the Tam analogs due to conformational stability offered by *H*-bond interactions at both ends of the ligand structure that was not observed for other compounds. Energy decomposition analysis further specifically identified residues that greatly contributed to inhibitor binding, which can then be targeted for optimization of Tam16 and the design of other analogs [224].

3.3. Ligand-Based Tools

3.3.1. Similarity-Based and Quantitative Structure-Activity/Property Relationship Methods

Even before the upsurge of available target structural data, rational inhibitor design has been employed with the help of substrate or product structures, and thus termed as ligand-based drug discovery and design (Figure 2). The simplest and most inexpensive LBDD approach is the similarity-based method, wherein compound candidates were designed and optimized through the principle of chemical similarity, i.e., similar (untested) ligand structures are posited to have similar activities as known inhibitors or modulators [226,227]. A typical workflow requires one or more reference structures with existing bioactivity information against a specific target. This is then used as a template to select new potential candidates from a chemical database to prioritize for assay testing.

Different molecular descriptors and parameterizations can be employed to characterize compounds and efficiently determine similarities. Descriptors can be generated as one-, two- or three-dimensional (1D, 2D, or 3D), wherein 1D descriptors comprise of global ligand properties (e.g., molecular weight, logP, number of *H*-bonds, etc.), 2D descriptors include topological and connectivity properties (e.g., aromaticity, degree of branching, etc.), and 3D descriptors involve geometrical properties (e.g., shape, volume, surface area, etc.) [228,229]. Additionally, fingerprints can also be used to depict template and database compounds by rendering structural features, such as those based on substructure (i.e., scaffold or functional groups), topology or path (i.e., fragmentation following a linear path of bonds), circular or radial (i.e., surrounding features of an atom up to a certain radius), and pharmacophoric (i.e., distance-based features, incorporating molecular shape, and interactions required for biological

activity) elements [230]. Well-known platforms that generate molecular descriptors and fingerprints are listed in Table 1.

To compare structural similarity after obtaining simplified molecular features, similarity coefficient and weighing scheme are required to measure and highlight the importance of certain aspects of a compound's structure in relation with its activity. Given that there are multiple tools available for both components, careful selection of analysis tools is crucial to have successful VS campaign. Tanimoto coefficient, which uses the ratio of shared features in both fingerprints to the total number of features between each fingerprint sets, is used as a standard for similarity evaluation of any two vectors. This coefficient returns values between 1 and 0 to depict chemical similarity [231]. Other known similarity measures include Manhattan distance, Euclidean distance, Dice index, Cosine coefficient, and others [232]. In terms of weighing schemes, some features may be 'silenced' or set as optional depending on its importance for a specific activity. There is not one method that can be considered the best for the full range of known targets and chemicals, as each method have their own data set applicability [233,234]. In this case, data fusion can be employed to obtain a consensus of outputs from different methods [235,236]. Both 2D and 3D similarity methods have been successful in identifying hits for various targets and have been established to have comparable or even better enrichment than docking [237]. While similarity-based approaches are known for their efficiency, there is a risk of obtaining low diversity hits as most similarity methods are highly dependent on the input structures used to calculate descriptors [238]. Moreover, there is a potential occurrence of 'activity cliffs,' in which small modifications in a ligand structure lead to significant difference in activity [239–241].

The similarity concept is also applied in studies involving quantitative structure-activity/property relationship (QSAR/QSPR), a method which is used to investigate the correlation between structural and physicochemical properties of ligands with known biological activities. QSAR modeling depends on the premise that ligand 2D and 3D properties can provide information to establish a statistical model of the desired biological activity, which can then be employed for activity prediction of ligand candidates [242,243]. The statistical model is generated based on an appropriate data set, consisting of compound structures with known bioactivity against a specific target, which must be checked and pre-processed. This data set should contain an adequate number of samples (i.e., a minimum of 20 experimentally-validated compounds) and, if from separate studies, identified using same assay protocols such that equivalent activities are obtained. Included in the data set preparation, especially for higher dimensions of QSAR modeling, is the conformational selection and alignment which allows the identification of scaffolds and functional groups that are critical to activity and therefore has more weight in the statistical model [242]. In this case, it is important to remember that the lowest energy conformation is not always equivalent to the bioactive conformation [244,245] and that ligands chosen for the training set should interact with the same binding site [242]. Typical alignment methods include the analysis of molecular fields, structural shapes, or pharmacophores. Pharmacophore generation for ligand alignment is more favorable as it aligns compounds based on feature similarity rather than chemical substructure [246]. As with similarity-based methods, QSAR makes use of molecular descriptors with dimensionality depending on the information available. Descriptors applied to the model should be carefully chosen to avoid autocorrelation and over-fitting. Before proceeding with and *in silico* prediction study, the model must be validated with internal and/or external data sets to establish its predictivity and applicability against a desired target [31,247].

Different QSAR methods have been developed and incorporated in various open-source platforms and commercial software (Table 1). The earliest QSAR-based algorithms include Comparative Molecular Field Analyses (CoMFA) [248] and Comparative Molecular Similarity Indices Analysis (CoMSIA) [249], both of which are still widely used today for various drug discovery endeavors [250–253]. 3D-QSAR CoMFA was used by Singh and Supuran [252] for the discovery of novel *Mtb* carbonic anhydrase inhibitors. A number of sulfonamides that target *Mtb* carbonic anhydrase 2 to regulate bacterial growth were used as the data set for QSAR modeling. The best developed model had excellent predictivity and good fit with an r^2 value of 0.93 and cross-validated coefficient q^2 value of 0.88. From the CoMFA

results, it was also determined that several steric and electrostatic features play critical roles in the inhibition of *Mtb* carbonic anhydrase 2. Using this information, nine compounds were designed and later observed to have better predicted inhibitory activities compared to the test set used. However, experimental validation is still required to determine the feasibility of these findings [252].

3.3.2. Ligand-Based Pharmacophore Modeling

Pharmacophore modelling have already demonstrated its value in ligand-based drug discovery studies. Ligand substructures required for optimal bioactivity can be aligned and characterized as a spatial arrangement of features in 3D space, which can be directly used for screening or applied to 3D-QSAR modeling [169,170,246]. Several software and webservers (Table 1) are available for the generation of ligand-based pharmacophores.

As with any other ligand-based methods, a data set of diverse ligands with known bioactivity against a specific target is required for pharmacophore generation. The training ligands used for ligand-based pharmacophore modeling must bind to the same pocket and have similar binding interactions, much like in QSAR studies. After obtaining a pharmacophore, its validity is assessed using a separate test set. It is then employed for virtual search of candidate compounds from libraries of untested molecules, wherein compounds are taken as potential hits if it 'fits' well with the pharmacophore (Figure 5). The main advantage of pharmacophore modeling is the use of molecular features rather than structural groups in depicting critical functionalities for activity, which allows for the identification of novel ligands with diverse structures (i.e., scaffold hopping) [172]. Moreover, pharmacophores can also be used for target profiling and polypharmacological studies to avoid adverse effects resulting from off-target binding [254]. This is especially useful when designing antitubercular and other antibiotic or antiviral agents as to avoid harmful interactions with human proteins.

Due to the numerous parallels between pharmacophore and 3D-QSAR modeling, these methods have been used in combination for a number of ligand-based drug discovery efforts. A study by Tawari et al. used PHASE [80] to target the *Mtb* aryl acid adenylating enzymes known as MbtA, which are involved in siderophore biosynthesis in tuberculosis [255]. A set of nucleoside bisubstrate analogs with known whole cell assay activity and bioactivity against siderophore biosynthesis in *Mtb* were used for pharmacophore and QSAR model development. *H*-bond donor, *H*-bond acceptor, and aromatic features were found to be critical for the inhibition of MbtA. The pharmacophore was also used to align molecules for the 3D-QSAR model, which exhibited suitable predictability and applicability with a Q^2 value of 0.71, RMSE of 0.65, and Pearson-R of 0.85 when assessed against a test set. The SAR studies additionally revealed the disadvantageous effects of bulky groups at the adenyl moiety *C*-6 position. Information taken from these models can be used for the rational design of new MbtA bisubstrate inhibitors as antitubercular agents [255].
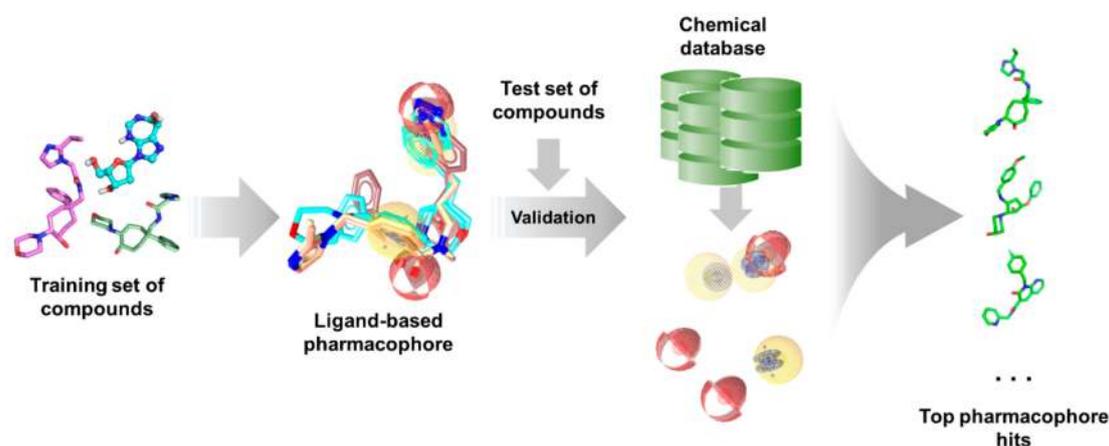

Figure 5. Typical ligand-based pharmacophore generation and screening workflow.

3.3.3. Density Functional Theory

Density functional theory (DFT) is quantum mechanical method established in the 1960s [256,257], which can be used in material science, computational chemistry, and computational physics to study the electronic properties of a many-particle (e.g., atom, molecule, condensed phase) system. DFT is based on two Hohenberg-Kohn (HK) theorems. First, the ground state properties of the many-particle system can be determined using only three spatially determined electron densities. Second, the HK theorem describes an energy functional, which can be minimized by the correct ground state electron density [258]. The use of DFT circumvents the computational expense of conventional methods like Hartree-Fock (HF) Theory since DFT relies on the premise that energies, intricate motions, and pair correlations can be derived directly from the electron probability density alone, instead of using wavefunctions. Theoretically, quantum mechanical wavefunctions consist of all the information required from a target system, and while the Schrödinger equation can be solved for a simple system, such as that of a hydrogen atom, it needs extensive computational efforts and it is impossible to solve this for a many-body system. In this case, DFT is used as an equivalent and efficient alternative to the Schrödinger equation [259], making DFT a popular tool in several computational fields [260].

In tuberculosis research, DFT has found uses in studies involving catalytic mechanisms [261,262], structure-activity relationship analysis [263], and inhibitor potency [264]. Chi and colleagues applied DFT to support their initial observations regarding a change in inhibitor binding mode in the MbtI enzyme after the addition of a substituted enolpyruvyl group to the parent compound structure previously designed from isochorismate [264]. X-ray crystal structures of MbtI complexed with its inhibitors depicted two different binding modes (Mode 1 and 2), suggesting binding site flexibility to accommodate ligand binding. The global minimum conformation of (*E*)-3-(1-carboxyprop-1-enyloxy)-2-hydroxybenzoic acid (AMT), *Z*-methyl-AMT, and *E*-methyl-AMT inhibitors in solution were calculated using Gaussian09 [265] with the B3LYP hybrid functional [266,267]. Global minimum conformation of free *Z*- and *E*-methyl-AMT were found to be similar to its bound conformation (Mode 2), indicating prearranged conformations to facilitate its binding to MbtI. Calculation of conformational entropy values for the three compounds revealed that *Z*-methyl-AMT is the least disordered, which may be due to the methyl conformational lock in its structure. Although a pure *Z*-isomer has not yet been obtained to experimentally differentiate it from the *E*-isomer, this finding rationalizes potent binding of methyl-AMT to MbtI and offers more information for the future design of novel and potent MbtI inhibitors [264].

Despite the success and popularity of DFT, it still has deficiencies due to approximations used in the development of functionals. Systems predominantly comprised of dispersion (van der Waals) forces, such as gaseous systems, or those wherein dispersion has a considerable contribution, such as biomolecular systems, are challenging to characterize using DFT [268]. However, several studies have already investigated the inclusion of van der Waals to improve this method [269–271]. Other major limitations of DFT application in computational chemistry include the characterization of charge transfer excitations, transition states, and global potential energy surfaces [272].

3.4. Integrated Tools

With the variety of available tools and structural data for drug discovery nowadays, it is more common to find studies that employ a combination of structure- and ligand-based approaches rather than exclusive application of each (Figure 2). Additionally, integration of these strategies often produces better results owing to more effective exploration of chemical and biological space. Moreover, the strengths of one method can overcome the limitations of the other, resulting in a highly complementary drug discovery process [171,238,273]. Integrated *in silico* workflows include sequential and parallel or data fusion methods [274–276], though hybrid methods have also already been developed [277,278]. Sequential methods involve the successive use of computational methods with the aim of increasing the selectivity of the VS workflow by continually reducing the number of potential hits before experimental evaluation [31]. However, it has been established that structure- and

ligand-based methods have similar enrichment and frequently yield hits with different scaffolds [276], indicating that these methods are better applied in parallel rather than sequentially [279]. Parallel application, through simultaneous employment of various computational tools, often produce a more diverse hit profile [31,280]. Nonetheless, since results from these methods are often fused to produce a final ranking, a large number of virtual hits is obtained from this approach [31].

In silico methods have already been applied to tuberculosis studies in several different ways and combinations depending on the goal of the study implemented, such as for drug discovery [179,281], understanding protein structure and function [282,283], and others [284]. One example showing the integration of computational methods is a study implemented by Li et al. involving 3D-QSAR, binding pocket prediction, docking, and MD simulation studies for FtsZ, which is a validated *Mtb* target and plays a significant role in cell division [285]. Trisubstituted benzimidazoles were found to target this protein and used for 3D-QSAR CoMFA [248] analysis in order to elucidate important structural factors related to their inhibitory activities. Homology modeling using the SWISS-MODEL server [32] was required to obtain the GDP-bound structure of *Mtb* FtsZ using *S. aureus* FtsZ as template. Afterwards, binding site prediction using ProFunc [61] was performed to identify potential binding pockets (other than the GDP binding site) for the candidate compounds. Selected trisubstituted benzimidazole analogs were docked into the *Mtb* FtsZ model using AutoDock, after which the lowest binding energy docked complex was refined using MD simulation. Using the MD-refined *Mtb* FtsZ structure, all trisubstituted benzimidazoles were docked using Surflex-Dock [286]. In the lowest energy state of the compounds, the benzimidazole scaffold and cyclohexyl group were located in a highly hydrophobic pocket within FtsZ, while the carbamate groups were oriented towards the hydrophilic area. These interactions are posited to be crucial for ligand binding stabilization and inhibition of *Mtb* FtsZ. The results of this study display how the concerted application of different *in silico* methods can lead to better understanding of protein structure, ligand design, and inhibitory activities [285].

4. Edges and Pitfalls of *In Silico* Methods

There are roughly 2500 protein structures for tuberculosis in the PDB and perhaps thousands of ligand candidates published. All these pieces of information are available with a few keyboard strokes and a click of the mouse. Along with existing technologies, it is now possible to analyze TB enzymes and lead candidates at the atomic level in order to understand their function and how to regulate them. While computational methods have been widely used in drug discovery nowadays due to their successful applications [287–289], it is still important to remember that these tools are like any other experimental approaches—prone to limitations dependent on the system and other various parameters being studied [290–292].

VS has been known to successfully screen millions of compounds to identify potential inhibitors for a given target [287,289]. This lends efficiency to cost, time, and efforts used in drug discovery projects as only the most promising compounds are brought forward for more rigorous experimental testing and drug development. However, optimization and validation of these methods are far from perfect and are highly dependent on the protein system and compound classes used, leading to possible bias in the computational model. Thus, it is challenging to determine which method has the advantage over another; many benchmark studies have been published regarding this matter [293,294]. Other major limitations include difficulties in incorporating protein flexibility and solvent effects due to the computational burden attached to these factors [31]. Fortunately, available technologies seem to be catching up as enhance sampling methods, HPC, and MD platforms are now routinely applied in drug discovery projects and are known to calculate up to milliseconds of simulations for various protein targets [295–297]. In terms of ligand-based drug design, its main advantage is its simplicity and efficiency. Indeed, LBDD has a long history and numerous candidates have already been discovered even with the lack of protein structural information [298–300]. Nonetheless, several factors should be considered when applying ligand-based tools. Firstly, ligand alignments are based on the lowest conformation energy, which is often different from the bioactive conformation [244,245], as well as

on the assumption that ligands bind in the same site and display the same conformation. Secondly, compounds should be evaluated by the same group (preferred) or tested using the same assay with the same parameters to be considered comparable [242]. Thirdly, the basic premise of 'similar structures display similar activities' are contradicted by the existence of activity cliffs [239–241], and so care should be taken when selecting potential candidates from a pool of virtual hits. Finally, it is also a challenge to incorporate the effects of solvation and protein flexibility due to the nature of the analysis.

As mentioned in the previous section, integration of several *in silico* methods have become common practice when designing and optimizing lead candidates to overcome the shortcomings of each individual tools. Despite requiring more computational resources, assimilation of computational methods result in better accuracy and enrichment of hits. In addition, the combination of a researcher's innate knowledge with the computational efficiency of these tools is perhaps the best integration of all, as a human's touch continues to be irreplaceable in the interpretation of all the data produced by *in silico* methods.

5. Conclusions and Future Perspectives

TB remains to be a relevant public health threat worldwide, necessitating accelerated discovery and design of novel antimycobacterial agents. Computer-aided drug design has become one of anchors of drug discovery research and continues to be a formidable tool in the hunt for promising drug leads, especially for tuberculosis. Continuous advancements in computing power and available software can enhance current computational tools and their application to different stages in the drug discovery pipeline. Nonetheless, these methods are not invincible as each tool have their own restrictions, and approximations are often used during the analysis. To overcome these, it is best to assimilate several *in silico* tools to complement the strength and limitations of each method used. The application of CADD in TB research has led to the identification of several antimycobacterial compounds that have already reached clinical evaluations, promoting its value in the drug discovery paradigm. Nonetheless, more work has to be done in order to expedite the discovery of anti-TB therapeutics.

Machine learning (ML) methods are making a comeback in drug discovery studies due to the upsurge in available data and enhanced computational powers. This has resulted in a wave drug discovery studies involving artificial intelligence (AI), wherein ML and deep learning (DL) techniques are applied to efficiently and 'intelligently' solve problems. This new shift in the drug discovery landscape is observed in personalized medicine and a number of relevant illnesses like cancer. While there are already several FDA-approved uses of AI in healthcare and diagnostics [301], it has yet to produce a successful drug candidate but it might not be far off. Currently, AI studies involving TB frequently covers diagnostics and treatment outcomes. This is perhaps one of the gaps that needs to be filled to be able to fast-track the discovery of novel and efficacious anti-TB drugs and finally alleviate the heavy burden of this infection around the globe.

Author Contributions: S.J.Y.M. and J.B.B. conceptualized and prepared the manuscript draft; J.B.B., V.G.O., and M.C.O.C. provided critical comments; all authors reviewed, edited, and approved the manuscript. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Pai, M.; Behr, M.A.; Dowdy, D.; Dheda, K.; Divangahi, M.; Boehme, C.C.; Ginsberg, A.; Swaminathan, S.; Spigelman, M.; Getahun, H.; et al. Tuberculosis. *Nat. Rev. Dis. Primers* **2016**, *2*, 16076. [CrossRef] [PubMed]
2. Havlir, D.V.; Getahun, H.; Sanne, I.; Nunn, P. Opportunities and challenges for HIV care in overlapping HIV and TB epidemics. *JAMA* **2008**, *300*, 423–430. [PubMed]

3. Leung, C.C.; Yew, W.W.; Chan, C.K.; Chang, K.C.; Law, W.S.; Lee, S.N.; Tai, L.B.; Leung, E.C.; Au, R.K.; Huang, S.S.; et al. Smoking adversely affects treatment response, outcome and relapse in tuberculosis. *Eur. Respir. J.* **2015**, *45*, 738–745. [CrossRef] [PubMed]

4. Imtiaz, S.; Shield, K.D.; Roerecke, M.; Samokhvalov, A.V.; Lonnroth, K.; Rehm, J. Alcohol consumption as a risk factor for tuberculosis: Meta-analyses and burden of disease. *Eur. Respir. J.* **2017**, *50*, 1700216. [CrossRef] [PubMed]

5. Restrepo, B.I. Diabetes and Tuberculosis. In *Understanding the Host Immune Response against Mycobacterium tuberculosis Infection*; Springer International Publishing: Cham, Switzerland, 2018; pp. 1–21.

6. Getahun, H.; Matteelli, A.; Chaisson, R.E.; Raviglione, M. Latent *Mycobacterium tuberculosis* infection. *N. Engl. J. Med.* **2015**, *372*, 2127–2135. [CrossRef]

7. Miller, L.G.; Asch, S.M.; Yu, E.I.; Knowles, L.; Gelberg, L.; Davidson, P. A population-based survey of tuberculosis symptoms: How atypical are atypical presentations? *Clin. Infect. Dis.* **2000**, *30*, 293–299. [CrossRef]

8. World Health Organization (WHO). *Global Tuberculosis Report 2019*; World Health Organization (WHO): Geneva, Switzerland, 2019.

9. World Health Organization (WHO). *Guidelines on the Management of Latent Tuberculosis Infection*; World Health Organization (WHO): Geneva, Switzerland, 2015.

10. World Health Organization (WHO). *Latent Tuberculosis infection: Updated and Consolidated Guidelines for Programmatic Management*; World Health Organization (WHO): Geneva, Switzerland, 2018.

11. Getahun, H.; Matteelli, A.; Abubakar, I.; Aziz, M.A.; Baddeley, A.; Barreira, D.; Den Boon, S.; Borroto Gutierrez, S.M.; Bruchfeld, J.; Burhan, E.; et al. Management of latent *Mycobacterium tuberculosis* infection: WHO guidelines for low tuberculosis burden countries. *Eur. Respir. J.* **2015**, *46*, 1563–1576. [CrossRef]

12. World Health Organization (WHO). *Treatment of Tuberculosis: Guidelines*; World Health Organization (WHO): Geneva, Switzerland, 2010.

13. Nahid, P.; Dorman, S.E.; Alipanah, N.; Barry, P.M.; Brozek, J.L.; Cattamanchi, A.; Chaisson, L.H.; Chaisson, R.E.; Daley, C.L.; Grzemska, M.; et al. Official American Thoracic Society/Centers for Disease Control and Prevention/Infectious Diseases Society of America Clinical Practice Guidelines: Treatment of Drug-Susceptible Tuberculosis. *Clin. Infect. Dis.* **2016**, *63*, e147–e195. [CrossRef]

14. Volmink, J.; Garner, P. Directly observed therapy for treating tuberculosis. *Cochrane Database Syst. Rev.* **2007**, CD003343. [CrossRef]

15. Horsburgh, C.R., Jr.; Barry, C.E., 3rd; Lange, C. Treatment of Tuberculosis. *N. Engl. J. Med.* **2015**, *373*, 2149–2160. [CrossRef]

16. Saukkonen, J.J.; Cohn, D.L.; Jasmer, R.M.; Schenker, S.; Jereb, J.A.; Nolan, C.M.; Peloquin, C.A.; Gordin, F.M.; Nunes, D.; Strader, D.B.; et al. An official ATS statement: Hepatotoxicity of antituberculosis therapy. *Am. J. Respir. Crit. Care Med.* **2006**, *174*, 935–952. [CrossRef]

17. Dheda, K.; Barry, C.E., 3rd; Maartens, G. Tuberculosis. *Lancet* **2016**, *387*, 1211–1226. [CrossRef]

18. Dheda, K.; Gumbo, T.; Gandhi, N.R.; Murray, M.; Theron, G.; Udwadia, Z.; Migliori, G.B.; Warren, R. Global control of tuberculosis: From extensively drug-resistant to untreatable tuberculosis. *Lancet. Respir. Med.* **2014**, *2*, 321–338. [CrossRef]

19. World Health Organization (WHO). *WHO Treatment Guidelines for Drug-Resistant Tuberculosis 2016 Update*; World Health Organization (WHO): Geneva, Switzerland, 2016.

20. Walker, J.; Tadena, N. J&J Tuberculosis Drug Gets Fast-Track Clearance. *Wall St. J.* **2013**. Available online: https://www.wsj.com/articles/SB10001424127887323320404578213421059138236 (accessed on 27 November 2019).

21. Mahajan, R. Bedaquiline: First FDA-approved tuberculosis drug in 40 years. *Int. J. Appl. Basic Med. Res.* **2013**, *3*, 1–2. [CrossRef]

22. European Medicines Agency (EMA). *Deltyba Delamanid Summary of the European Public Assessment Report (EPAR) for Deltyba*; EMA: Amsterdam, The Netherlands, 2014; pp. 1–3. Available online: https://www.ema.europa.eu/en/medicines/human/EPAR/deltyba (accessed on 15 November 2019).

23. Ryan, N.J.; Lo, J.H. Delamanid: First global approval. *Drugs* **2014**, *74*, 1041–1045. [CrossRef]

24. US FDA. *FDA Approves New Drug for Treatment-Resistant Forms of Tuberculosis That Affects the Lungs*; US FDA: Silver Spring, MD, USA, 2019.

25. Baptista, R.; Fazakerley, D.M.; Beckmann, M.; Baillie, L.; Mur, L.A.J. Untargeted metabolomics reveals a new mode of action of pretomanid (PA-824). *Sci. Rep.* **2018**, *8*, 5084. [CrossRef]

26. Thompson, A.M.; Bonnet, M.; Lee, H.H.; Franzblau, S.G.; Wan, B.; Wong, G.S.; Cooper, C.B.; Denny, W.A. Antitubercular Nitroimidazoles Revisited: Synthesis and Activity of the Authentic 3-Nitro Isomer of Pretomanid. *ACS Med. Chem. Lett.* **2017**, *8*, 1275–1280. [CrossRef]

27. Manjunatha, U.; Boshoff, H.I.; Barry, C.E. The mechanism of action of PA-824: Novel insights from transcriptional profiling. *Commun. Integr. Biol.* **2009**, *2*, 215–218. [CrossRef]

28. Reymond, J.-L.; van Deursen, R.; Blum, L.C.; Ruddigkeit, L. Chemical space as a source for new drugs. *MedChemComm* **2010**, *1*, 30–38. [CrossRef]

29. Berman, H.M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.N.; Weissig, H.; Shindyalov, I.N.; Bourne, P.E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242. [CrossRef]

30. Ekins, S.; Freundlich, J.S.; Choi, I.; Sarker, M.; Talcott, C. Computational databases, pathway and cheminformatics tools for tuberculosis drug discovery. *Trends Microbiol.* **2011**, *19*, 65–74. [CrossRef]

31. Macalino, S.J.; Gosu, V.; Hong, S.; Choi, S. Role of computer-aided drug design in modern drug discovery. *Arch. Pharm. Res.* **2015**, *38*, 1686–1701. [CrossRef]

32. Schwede, T.; Kopp, J.; Guex, N.; Peitsch, M.C. SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Res.* **2003**, *31*, 3381–3385. [CrossRef]

33. Webb, B.; Sali, A. Comparative Protein Structure Modeling Using MODELLER. *Curr. Protoc. Protein Sci.* **2016**, *86*, 1–37. [CrossRef]

34. Kim, D.E.; Chivian, D.; Baker, D. Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Res.* **2004**, *32*, W526–W531. [CrossRef]

35. Schrödinger. Prime. Available online: https://www.schrodinger.com/prime (accessed on 26 October 2019).

36. Zheng, W.; Zhang, C.; Bell, E.W.; Zhang, Y. I-TASSER gateway: A protein structure and function prediction server powered by XSEDE. *Future Gener. Comput. Syst.* **2019**, *99*, 73–85. [CrossRef]

37. Zhang, Y. I-TASSER server for protein 3D structure prediction. *BMC Bioinf.* **2008**, *9*, 40. [CrossRef]

38. Yang, J.; Zhang, Y. I-TASSER server: New development for protein structure and function predictions. *Nucleic Acids Res.* **2015**, *43*, W174–W181. [CrossRef]

39. Yang, J.; Yan, R.; Roy, A.; Xu, D.; Poisson, J.; Zhang, Y. The I-TASSER Suite: Protein structure and function prediction. *Nat. Methods* **2015**, *12*, 7–8. [CrossRef]

40. Roy, A.; Kucukural, A.; Zhang, Y. I-TASSER: A unified platform for automated protein structure and function prediction. *Nat. Protoc.* **2010**, *5*, 725–738. [CrossRef] [PubMed]

41. Zhang, Y. I-TASSER: Fully automated protein structure prediction in CASP8. *Proteins* **2009**, *77*, 100–113. [CrossRef] [PubMed]

42. Zimmermann, L.; Stephens, A.; Nam, S.Z.; Rau, D.; Kubler, J.; Lozajic, M.; Gabler, F.; Soding, J.; Lupas, A.N.; Alva, V. A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core. *J. Mol. Biol.* **2018**, *430*, 2237–2243. [CrossRef] [PubMed]

43. Hildebrand, A.; Remmert, M.; Biegert, A.; Soding, J. Fast and accurate automatic structure prediction with HHpred. *Proteins* **2009**, *77*, 128–132. [CrossRef] [PubMed]

44. Soding, J.; Biegert, A.; Lupas, A.N. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.* **2005**, *33*, W244–W248. [CrossRef] [PubMed]

45. Laskowski, R.A.; MacArthur, M.W.; Moss, D.S.; Thornton, J.M. PROCHECK: A program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **1993**, *26*, 283–291. [CrossRef]

46. Wiederstein, M.; Sippl, M.J. ProSA-web: Interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res.* **2007**, *35*, W407–W410. [CrossRef]

47. Eisenberg, D.; Luthy, R.; Bowie, J.U. VERIFY3D: Assessment of protein models with three-dimensional profiles. *Methods Enzymol.* **1997**, *277*, 396–404.

48. Colovos, C.; Yeates, T.O. Verification of protein structures: Patterns of nonbonded atomic interactions. *Protein Sci.* **1993**, *2*, 1511–1519. [CrossRef]

49. Hussein, H.A.; Borrel, A.; Geneix, C.; Petitjean, M.; Regad, L.; Camproux, A.C. PockDrug-Server: A new web server for predicting pocket druggability on holo and apo proteins. *Nucleic Acids Res.* **2015**, *43*, W436–W442. [CrossRef]

50. Volkamer, A.; Kuhn, D.; Rippmann, F.; Rarey, M. DoGSiteScorer: A web server for automatic binding site prediction, analysis and druggability assessment. *Bioinformatics* **2012**, *28*, 2074–2075. [CrossRef] [PubMed]

51. Schmidtke, P.; Le Guilloux, V.; Maupetit, J.; Tuffery, P. fpocket: Online tools for protein ensemble pocket detection and tracking. *Nucleic Acids Res.* **2010**, *38*, W582–W589. [CrossRef] [PubMed]

52. Le Guilloux, V.; Schmidtke, P.; Tuffery, P. Fpocket: An open source platform for ligand pocket detection. *BMC Bioinf.* **2009**, *10*, 168. [CrossRef] [PubMed]

53. Tian, W.; Chen, C.; Lei, X.; Zhao, J.; Liang, J. CASTp 3.0: Computed atlas of surface topography of proteins. *Nucleic Acids Res.* **2018**, *46*, W363–W367. [CrossRef] [PubMed]

54. Binkowski, T.A.; Naghibzadeh, S.; Liang, J. CASTp: Computed Atlas of Surface Topography of proteins. *Nucleic Acids Res.* **2003**, *31*, 3352–3355. [CrossRef]

55. Dundas, J.; Ouyang, Z.; Tseng, J.; Binkowski, A.; Turpaz, Y.; Liang, J. CASTp: Computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. *Nucleic Acids Res.* **2006**, *34*, W116–W118. [CrossRef]

56. Koes, D.R.; Camacho, C.J. PocketQuery: Protein-protein interaction inhibitor starting points from protein-protein interaction structure. *Nucleic Acids Res.* **2012**, *40*, W387–W392. [CrossRef]

57. Brady, G.P., Jr.; Stouten, P.F. Fast prediction and visualization of protein binding pockets with PASS. *J. Comput. Aided Mol. Des.* **2000**, *14*, 383–401. [CrossRef]

58. Halgren, T.A. Identifying and characterizing binding sites and assessing druggability. *J. Chem. Inf. Model.* **2009**, *49*, 377–389. [CrossRef]

59. Capra, J.A.; Laskowski, R.A.; Thornton, J.M.; Singh, M.; Funkhouser, T.A. Predicting protein ligand binding sites by combining evolutionary sequence conservation and 3D structure. *PLoS Comput. Biol.* **2009**, *5*, e1000585. [CrossRef]

60. Jendele, L.; Krivak, R.; Skoda, P.; Novotny, M.; Hoksza, D. PrankWeb: A web server for ligand binding site prediction and visualization. *Nucleic Acids Res.* **2019**, *47*, W345–W349. [CrossRef] [PubMed]

61. Laskowski, R.A.; Watson, J.D.; Thornton, J.M. ProFunc: A server for predicting protein function from 3D structure. *Nucleic Acids Res.* **2005**, *33*, W89–W93. [CrossRef]

62. Morris, G.M.; Huey, R.; Lindstrom, W.; Sanner, M.F.; Belew, R.K.; Goodsell, D.S.; Olson, A.J. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* **2009**, *30*, 2785–2791. [CrossRef] [PubMed]

63. Trott, O.; Olson, A.J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **2010**, *31*, 455–461. [CrossRef] [PubMed]

64. Allen, W.J.; Balius, T.E.; Mukherjee, S.; Brozell, S.R.; Moustakas, D.T.; Lang, P.T.; Case, D.A.; Kuntz, I.D.; Rizzo, R.C. DOCK 6: Impact of new features and current docking performance. *J. Comput. Chem.* **2015**, *36*, 1132–1156. [CrossRef]

65. Verdonk, M.L.; Cole, J.C.; Hartshorn, M.J.; Murray, C.W.; Taylor, R.D. Improved protein–ligand docking using GOLD. *Proteins* **2003**, *52*, 609–623. [CrossRef]

66. Schrödinger. Glide. Available online: https://www.schrodinger.com/glide (accessed on 26 October 2019).

67. Schrödinger. Induced Fit. Available online: https://www.schrodinger.com/induced-fit (accessed on 26 October 2019).

68. Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. A fast flexible docking method using an incremental construction algorithm. *J. Mol. Biol.* **1996**, *261*, 470–489. [CrossRef]

69. Davis, I.W.; Baker, D. RosettaLigand docking with full ligand and receptor flexibility. *J. Mol. Biol.* **2009**, *385*, 381–392. [CrossRef]

70. Wu, G.; Robertson, D.H.; Brooks, C.L., 3rd; Vieth, M. Detailed analysis of grid-based molecular docking: A case study of CDOCKER-A CHARMm-based MD docking algorithm. *J. Comput. Chem.* **2003**, *24*, 1549–1562. [CrossRef]

71. Bitencourt-Ferreira, G.; de Azevedo, W.F., Jr. Docking with SwissDock. *Methods Mol. Biol.* **2019**, *2053*, 189–202.

72. Grosdidier, A.; Zoete, V.; Michielin, O. SwissDock, a protein-small molecule docking web service based on EADock DSS. *Nucleic Acids Res.* **2011**, *39*, W270–W277. [CrossRef] [PubMed]

73. Koes, D.R.; Camacho, C.J. Pharmer: Efficient and exact pharmacophore search. *J. Chem. Inf. Model.* **2011**, *51*, 1307–1314. [CrossRef] [PubMed]

74. Dassault Systèmes BIOVIA. *Catalyst*. Available online: https://www.3dsbiovia.com/products/collaborative-science/biovia-discovery-studio/pharmacophore-and-ligand-based-design.html (accessed on 26 October 2019).

75. Schneidman-Duhovny, D.; Dror, O.; Inbar, Y.; Nussinov, R.; Wolfson, H.J. PharmaGist: A webserver for ligand-based pharmacophore detection. *Nucleic Acids Res.* **2008**, *36*, W223–W228. [CrossRef] [PubMed]

76. Inte:Ligand. LigandScout. Available online: http://www.inteligand.com/ligandscout/ (accessed on 26 October 2019).

77. Zoete, V.; Daina, A.; Bovigny, C.; Michielin, O. SwissSimilarity: A Web Tool for Low to Ultra High Throughput Ligand-Based Virtual Screening. *J. Chem. Inf. Model.* **2016**, *56*, 1399–1404. [CrossRef]

78. Douguet, D. e-LEA3D: A computational-aided drug design web server. *Nucleic Acids Res.* **2010**, *38*, W615–W621. [CrossRef]

79. Dallakyan, S.; Olson, A.J. Small-molecule library screening by docking with PyRx. *Methods Mol. Biol.* **2015**, *1263*, 243–250.

80. Schrödinger. PHASE. Available online: https://www.schrodinger.com/phase (accessed on 26 October 2019).

81. Case, D.A.; Cheatham, T.E., 3rd; Darden, T.; Gohlke, H.; Luo, R.; Merz, K.M., Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R.J. The Amber biomolecular simulation programs. *J. Comput. Chem.* **2005**, *26*, 1668–1688. [CrossRef]

82. Salomon-Ferrer, R.; Case, D.A.; Walker, R.C. An overview of the Amber biomolecular simulation package. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2013**, *3*, 198–210. [CrossRef]

83. Brooks, B.R.; Brooks, C.L., 3rd; Mackerell, A.D., Jr.; Nilsson, L.; Petrella, R.J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; et al. CHARMM: The biomolecular simulation program. *J. Comput. Chem.* **2009**, *30*, 1545–1614. [CrossRef]

84. Miller, B.T.; Singh, R.P.; Klauda, J.B.; Hodoscek, M.; Brooks, B.R.; Woodcock, H.L., 3rd. CHARMMing: A new, flexible web portal for CHARMM. *J. Chem. Inf. Model.* **2008**, *48*, 1920–1929. [CrossRef]

85. Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A.E.; Berendsen, H.J. GROMACS: Fast, flexible, and free. *J. Comput. Chem.* **2005**, *26*, 1701–1718. [CrossRef] [PubMed]

86. Abraham, M.J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J.C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1–2*, 19–25. [CrossRef]

87. Phillips, J.C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R.D.; Kale, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781–1802. [CrossRef] [PubMed]

88. Schrödinger. Desmond. Available online: https://www.schrodinger.com/desmond (accessed on 26 October 2019).

89. Zoete, V.; Cuendet, M.A.; Grosdidier, A.; Michielin, O. SwissParam: A fast force field generation tool for small organic molecules. *J. Comput. Chem.* **2011**, *32*, 2359–2368. [CrossRef] [PubMed]

90. Jo, S.; Kim, T.; Iyer, V.G.; Im, W. CHARMM-GUI: A web-based graphical user interface for CHARMM. *J. Comput. Chem.* **2008**, *29*, 1859–1865. [CrossRef] [PubMed]

91. Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; et al. CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* **2010**, *31*, 671–690. [CrossRef] [PubMed]

92. Vanommeslaeghe, K.; MacKerell, A.D., Jr. Automation of the CHARMM General Force Field (CGenFF) I: Bond perception and atom typing. *J. Chem. Inf. Model.* **2012**, *52*, 3144–3154. [CrossRef] [PubMed]

93. Vanommeslaeghe, K.; Raman, E.P.; MacKerell, A.D., Jr. Automation of the CHARMM General Force Field (CGenFF) II: Assignment of bonded parameters and partial atomic charges. *J. Chem. Inf. Model.* **2012**, *52*, 3155–3168. [CrossRef]

94. Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graph.* **1996**, *14*, 27–38. [CrossRef]

95. Helguera, A.M.; Combes, R.D.; Gonzalez, M.P.; Cordeiro, M.N. Applications of 2D descriptors in drug design: A DRAGON tale. *Curr. Top. Med. Chem.* **2008**, *8*, 1628–1655. [CrossRef]

96. Tetko, I.V.; Gasteiger, J.; Todeschini, R.; Mauri, A.; Livingstone, D.; Ertl, P.; Palyulin, V.A.; Radchenko, E.V.; Zefirov, N.S.; Makarenko, A.S.; et al. Virtual computational chemistry laboratory—Design and description. *J. Comput. Aided Mol. Des.* **2005**, *19*, 453–463. [CrossRef] [PubMed]

97. Schrödinger. Canvas. Available online: https://www.schrodinger.com/canvas (accessed on 26 October 2019).

98. Landrum, G. RDKit: Open-Source Cheminformatics. Available online: http://www.rdkit.org (accessed on 26 October 2019).

99. Masand, V.H.; Rastija, V. PyDescriptor: A new PyMOL plugin for calculating thousands of easily understandable molecular descriptors. *Chemom. Intell. Lab. Syst.* **2017**, *169*, 12–18. [CrossRef]

100. Moriwaki, H.; Tian, Y.S.; Kawashita, N.; Takagi, T. Mordred: A molecular descriptor calculator. *J. Cheminform.* **2018**, *10*, 4. [CrossRef] [PubMed]

101. Tosco, P.; Balle, T. Open3DQSAR: A new open-source software aimed at high-throughput chemometric analysis of molecular interaction fields. *J. Mol. Model.* **2011**, *17*, 201–208. [CrossRef]

102. Dong, J.; Yao, Z.J.; Zhu, M.F.; Wang, N.N.; Lu, B.; Chen, A.F.; Lu, A.P.; Miao, H.; Zeng, W.B.; Cao, D.S. ChemSAR: An online pipelining platform for molecular SAR modeling. *J. Cheminform.* **2017**, *9*, 27. [CrossRef]

103. BioSolveIT. SeeSAR version 9.2. Available online: https://www.biosolveit.de/SeeSAR/ (accessed on 26 October 2019).

104. Schrödinger. QikProp. Available online: https://www.schrodinger.com/qikprop (accessed on 26 October 2019).

105. SimulationsPlus. ADMET Predictor. Available online: https://www.simulations-plus.com/software/admetpredictor/ (accessed on 26 October 2019).

106. ACD/Labs. Percepta Platform. Available online: https://www.acdlabs.com/products/percepta/ (accessed on 26 October 2019).

107. Miteva, M.A.; Violas, S.; Montes, M.; Gomez, D.; Tuffery, P.; Villoutreix, B.O. FAF-Drugs: Free ADME/tox filtering of compound collections. *Nucleic Acids Res.* **2006**, *34*, W738–W744. [CrossRef]

108. Rasolohery, I.; Moroy, G.; Guyon, F. PatchSearch: A Fast Computational Method for Off-Target Detection. *J. Chem. Inf. Model.* **2017**, *57*, 769–777. [CrossRef]

109. Dassault Systèmes BIOVIA. DS TOPKAT. Available online: https://www.3dsbiovia.com/products/collaborative-science/biovia-discovery-studio/qsar-admet-and-predictive-toxicology.html (accessed on 26 October 2019).

110. Dassault Systèmes BIOVIA. DS ADMET. Available online: https://www.3dsbiovia.com/products/collaborative-science/biovia-pipeline-pilot/component-collections/adme-tox.html (accessed on 26 October 2019).

111. Poroikov, V.; Filimonov, D.; Lagunin, A.; Gloriozova, T.; Zakharov, A. PASS: Identification of probable targets and mechanisms of toxicity. *SAR QSAR Environ. Res.* **2007**, *18*, 101–110. [CrossRef]

112. Daina, A.; Michielin, O.; Zoete, V. SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Sci. Rep.* **2017**, *7*, 42717. [CrossRef]

113. Cruciani, G.; Carosati, E.; De Boeck, B.; Ethirajulu, K.; Mackie, C.; Howe, T.; Vianello, R. MetaSite: Understanding metabolism in human cytochromes from the perspective of the chemist. *J. Med. Chem.* **2005**, *48*, 6970–6979. [CrossRef]

114. Tcheremenskaia, O.; Benigni, R.; Nikolova, I.; Jeliazkova, N.; Escher, S.E.; Batke, M.; Baier, T.; Poroikov, V.; Lagunin, A.; Rautenberg, M.; et al. OpenTox predictive toxicology framework: Toxicological ontology and semantic media wiki-based OpenToxipedia. *J. Biomed. Semant.* **2012**, *3*, S7. [CrossRef] [PubMed]

115. Smiesko, M.; Vedani, A. VirtualToxLab: Exploring the Toxic Potential of Rejuvenating Substances Found in Traditional Medicines. *Methods Mol. Biol.* **2016**, *1425*, 121–137.

116. Vedani, A.; Dobler, M.; Smiesko, M. VirtualToxLab—A platform for estimating the toxic potential of drugs, chemicals and natural products. *Toxicol. Appl. Pharmacol.* **2012**, *261*, 142–153. [CrossRef] [PubMed]

117. Vedani, A.; Smiesko, M.; Spreafico, M.; Peristera, O.; Dobler, M. VirtualToxLab—In silico prediction of the toxic (endocrine-disrupting) potential of drugs, chemicals and natural products. Two years and 2000 compounds of experience: A progress report. *ALTEX* **2009**, *26*, 167–176. [CrossRef] [PubMed]

118. Vedani, A.; Dobler, M.; Spreafico, M.; Peristera, O.; Smiesko, M. VirtualToxLab—In silico prediction of the toxic potential of drugs and environmental chemicals: Evaluation status and internet access protocol. *ALTEX* **2007**, *24*, 153–161. [CrossRef]

119. Cheng, F.; Li, W.; Zhou, Y.; Shen, J.; Wu, Z.; Liu, G.; Lee, P.W.; Tang, Y. Correction to "admetSAR: A Comprehensive Source and Free Tool for Assessment of Chemical ADMET Properties". *J. Chem. Inf. Model.* **2019**. [CrossRef]

120. Yang, H.; Lou, C.; Sun, L.; Li, J.; Cai, Y.; Wang, Z.; Li, W.; Liu, G.; Tang, Y. admetSAR 2.0: Web-service for prediction and optimization of chemical ADMET properties. *Bioinformatics* **2019**, *35*, 1067–1069. [CrossRef]

121. Cheng, F.; Li, W.; Zhou, Y.; Shen, J.; Wu, Z.; Liu, G.; Lee, P.W.; Tang, Y. admetSAR: A comprehensive source and free tool for assessment of chemical ADMET properties. *J. Chem. Inf. Model.* **2012**, *52*, 3099–3105. [CrossRef]

122. Rudik, A.; Bezhentsev, V.; Dmitriev, A.; Lagunin, A.; Filimonov, D.; Poroikov, V. Metatox-Web application for generation of metabolic pathways and toxicity estimation. *J. Bioinform. Comput. Biol.* **2019**, *17*, 1940001. [CrossRef] [PubMed]

123. Rudik, A.V.; Bezhentsev, V.M.; Dmitriev, A.V.; Druzhilovskiy, D.S.; Lagunin, A.A.; Filimonov, D.A.; Poroikov, V.V. MetaTox: Web Application for Predicting Structure and Toxicity of Xenobiotics' Metabolites. *J. Chem. Inf. Model.* **2017**, *57*, 638–642. [CrossRef]

124. Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.V.; Eiglmeier, K.; Gas, S.; Barry, C.E., 3rd; et al. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* **1998**, *393*, 537–544. [CrossRef] [PubMed]

125. Rosenkrands, I.; King, A.; Weldingh, K.; Moniatte, M.; Moertz, E.; Andersen, P. Towards the proteome of *Mycobacterium tuberculosis*. *Electrophoresis* **2000**, *21*, 3740–3756. [CrossRef]

126. Jungblut, P.R.; Schaible, U.E.; Mollenkopf, H.J.; Zimny-Arndt, U.; Raupach, B.; Mattow, J.; Halada, P.; Lamer, S.; Hagens, K.; Kaufmann, S.H. Comparative proteome analysis of *Mycobacterium tuberculosis* and Mycobacterium bovis BCG strains: Towards functional genomics of microbial pathogens. *Mol. Microbiol.* **1999**, *33*, 1103–1117. [CrossRef] [PubMed]

127. Kruh, N.A.; Troudt, J.; Izzo, A.; Prenni, J.; Dobos, K.M. Portrait of a pathogen: The *Mycobacterium tuberculosis* proteome in vivo. *PLoS ONE* **2010**, *5*, e13938. [CrossRef]

128. Keren, I.; Minami, S.; Rubin, E.; Lewis, K. Characterization and transcriptome analysis of *Mycobacterium tuberculosis* persisters. *MBio* **2011**, *2*, e00100–e00111. [CrossRef]

129. Rachman, H.; Strong, M.; Ulrichs, T.; Grode, L.; Schuchhardt, J.; Mollenkopf, H.; Kosmiadi, G.A.; Eisenberg, D.; Kaufmann, S.H. Unique transcriptome signature of *Mycobacterium tuberculosis* in pulmonary tuberculosis. *Infect. Immun.* **2006**, *74*, 1233–1242. [CrossRef]

130. Xu, D. Protein databases on the internet. *Curr. Protoc. Protein Sci.* **2012**. [CrossRef]

131. Kim, S.; Thiessen, P.A.; Bolton, E.E.; Chen, J.; Fu, G.; Gindulyte, A.; Han, L.; He, J.; He, S.; Shoemaker, B.A.; et al. PubChem Substance and Compound databases. *Nucleic Acids Res.* **2016**, *44*, D1202–D1213. [CrossRef]

132. Williams, A.J. Public chemical compound databases. *Curr. Opin. Drug Discov. Dev.* **2008**, *11*, 393–404.

133. UniProt, C. Ongoing and future developments at the Universal Protein Resource. *Nucleic Acids Res.* **2011**, *39*, D214–D219.

134. Reddy, T.B.; Riley, R.; Wymore, F.; Montgomery, P.; DeCaprio, D.; Engels, R.; Gellesch, M.; Hubble, J.; Jen, D.; Jin, H.; et al. TB database: An integrated platform for tuberculosis research. *Nucleic Acids Res.* **2009**, *37*, D499–D508. [CrossRef] [PubMed]

135. Galagan, J.E.; Sisk, P.; Stolte, C.; Weiner, B.; Koehrsen, M.; Wymore, F.; Reddy, T.B.; Zucker, J.D.; Engels, R.; Gellesch, M.; et al. TB database 2010: Overview and update. *Tuberculosis* **2010**, *90*, 225–235. [CrossRef] [PubMed]

136. Kapopoulou, A.; Lew, J.M.; Cole, S.T. The MycoBrowser portal: A comprehensive and manually annotated resource for mycobacterial genomes. *Tuberculosis* **2011**, *91*, 8–13. [CrossRef]

137. Rosenthal, A.; Gabrielian, A.; Engle, E.; Hurt, D.E.; Alexandru, S.; Crudu, V.; Sergueev, E.; Kirichenko, V.; Lapitskii, V.; Snezhko, E.; et al. The TB Portals: An Open-Access, Web-Based Platform for Global Drug-Resistant-Tuberculosis Data Sharing and Analysis. *J. Clin. Microbiol.* **2017**, *55*, 3267–3282. [CrossRef] [PubMed]

138. Laskowski, R.A.; Jablonska, J.; Pravda, L.; Varekova, R.S.; Thornton, J.M. PDBsum: Structural summaries of PDB entries. *Protein Sci.* **2018**, *27*, 129–134. [CrossRef] [PubMed]

139. Liu, R.; Li, X.; Lam, K.S. Combinatorial chemistry in drug discovery. *Curr. Opin. Chem. Biol.* **2017**, *38*, 117–126. [CrossRef]

140. Ruddigkeit, L.; van Deursen, R.; Blum, L.C.; Reymond, J.L. Enumeration of 166 billion organic small molecules in the chemical universe database GDB-17. *J. Chem. Inf. Model.* **2012**, *52*, 2864–2875. [CrossRef]

141. Shivanyuk, A.; Ryabukhin, S.; Bogolyubsky, A.V.; Mykytenko, D.M.; Chuprina, A.; Heilman, W.; Kostyuk, A.N.; Tolmachev, A. Enamine real database: Making chemical diversity real. *Chem. Today* **2007**, *25*, 58–59.

142. Williams, A.J. ChemSpider: Integrating Structure-Based Resources Distributed across the Internet. In *Enhancing Learning with Online Resources, Social Networking, and Digital Libraries*; American Chemical Society: Washington, DC, USA, 2010; Volume 1060, pp. 23–39.

143. Sterling, T.; Irwin, J.J. ZINC 15—Ligand Discovery for Everyone. *J. Chem. Inf. Model.* **2015**, *55*, 2324–2337. [CrossRef]

144. Gaulton, A.; Hersey, A.; Nowotka, M.; Bento, A.P.; Chambers, J.; Mendez, D.; Mutowo, P.; Atkinson, F.; Bellis, L.J.; Cibrian-Uhalte, E.; et al. The ChEMBL database in 2017. *Nucleic Acids Res.* **2017**, *45*, D945–D954. [CrossRef]

145. Voigt, J.H.; Bienfait, B.; Wang, S.; Nicklaus, M.C. Comparison of the NCI open database with seven large chemical structural databases. *J. Chem. Inf. Comput Sci.* **2001**, *41*, 702–712. [CrossRef]

146. Pierleoni, A.; Martelli, P.L.; Fariselli, P.; Casadio, R. eSLDB: Eukaryotic subcellular localization database. *Nucleic Acids Res.* **2007**, *35*, D208–D212. [CrossRef]

147. Sprenger, J.; Lynn Fink, J.; Karunaratne, S.; Hanson, K.; Hamilton, N.A.; Teasdale, R.D. LOCATE: A mammalian protein subcellular localization database. *Nucleic Acids Res.* **2008**, *36*, D230–D233. [CrossRef]

148. Peabody, M.A.; Laird, M.R.; Vlasschaert, C.; Lo, R.; Brinkman, F.S. PSORTdb: Expanding the bacteria and archaea protein subcellular localization database to better reflect diversity in cell envelope structures. *Nucleic Acids Res.* **2016**, *44*, D663–D668. [CrossRef]

149. Hendlich, M.; Bergner, A.; Gunther, J.; Klebe, G. Relibase: Design and development of a database for comprehensive analysis of protein-ligand interactions. *J. Mol. Biol.* **2003**, *326*, 607–620. [CrossRef]

150. Chen, X.; Lin, Y.; Liu, M.; Gilson, M.K. The Binding Database: Data management and interface design. *Bioinformatics* **2002**, *18*, 130–139. [CrossRef]

151. Liu, T.; Lin, Y.; Wen, X.; Jorissen, R.N.; Gilson, M.K. BindingDB: A web-accessible database of experimentally determined protein-ligand binding affinities. *Nucleic Acids Res.* **2007**, *35*, D198–D201. [CrossRef]

152. Salwinski, L.; Miller, C.S.; Smith, A.J.; Pettit, F.K.; Bowie, J.U.; Eisenberg, D. The Database of Interacting Proteins: 2004 update. *Nucleic Acids Res.* **2004**, *32*, D449–D451. [CrossRef]

153. Oughtred, R.; Stark, C.; Breitkreutz, B.J.; Rust, J.; Boucher, L.; Chang, C.; Kolas, N.; O'Donnell, L.; Leung, G.; McAdam, R.; et al. The BioGRID interaction database: 2019 update. *Nucleic Acids Res.* **2019**, *47*, D529–D541. [CrossRef]

154. Jensen, L.J.; Kuhn, M.; Stark, M.; Chaffron, S.; Creevey, C.; Muller, J.; Doerks, T.; Julien, P.; Roth, A.; Simonovic, M.; et al. STRING 8—A global view on proteins and their functional interactions in 630 organisms. *Nucleic Acids Res.* **2009**, *37*, D412–D416. [CrossRef]

155. Franca, T.C. Homology modeling: An important tool for the drug discovery. *J. Biomol. Struct. Dyn.* **2015**, *33*, 1780–1793. [CrossRef]

156. McGinnis, S.; Madden, T.L. BLAST: At the core of a powerful and diverse set of sequence analysis tools. *Nucleic Acids Res.* **2004**, *32*, W20–W25. [CrossRef]

157. Ye, J.; McGinnis, S.; Madden, T.L. BLAST: Improvements for better sequence analysis. *Nucleic Acids Res.* **2006**, *34*, W6–W9. [CrossRef]

158. Madden, T. The BLAST Sequence Analysis Tool. In *The NCBI Handbook [Internet]*, 2nd ed.; National Center for Biotechnology Information (US): Bethesda, MD, USA, 2013.

159. Papadopoulos, J.S.; Agarwala, R. COBALT: Constraint-based alignment tool for multiple protein sequences. *Bioinformatics* **2007**, *23*, 1073–1079. [CrossRef]

160. Sievers, F.; Higgins, D.G. Clustal Omega, Accurate Alignment of Very Large Numbers of Sequences. In *Multiple Sequence Alignment Methods*; Russell, D.J., Ed.; Humana Press: Totowa, NJ, USA, 2014; pp. 105–116.

161. Lassmann, T.; Sonnhammer, E.L. Kalign—An accurate and fast multiple sequence alignment algorithm. *BMC Bioinf.* **2005**, *6*, 298. [CrossRef]

162. Schmidtke, P.; Barril, X. Understanding and predicting druggability. A high-throughput method for detection of drug binding sites. *J. Med. Chem.* **2010**, *53*, 5858–5867. [CrossRef]

163. Kalyaanamoorthy, S.; Chen, Y.P. Structure-based drug design to augment hit discovery. *Drug Discov. Today* **2011**, *16*, 831–839. [CrossRef]

164. Hetenyi, C.; van der Spoel, D. Blind docking of drug-sized compounds to proteins with up to a thousand residues. *FEBS Lett.* **2006**, *580*, 1447–1450. [CrossRef]

165. Shi, Y.; Colombo, C.; Kuttiyatveetil, J.R.; Zalatar, N.; van Straaten, K.E.; Mohan, S.; Sanders, D.A.; Pinto, B.M. A Second, Druggable Binding Site in UDP-Galactopyranose Mutase from *Mycobacterium tuberculosis*? *Chembiochem* **2016**, *17*, 2264–2273. [CrossRef]

166. Stahura, F.L.; Bajorath, J. Virtual screening methods that complement HTS. *Comb. Chem. High Throughput Screen.* **2004**, *7*, 259–269. [CrossRef]

167. Steindl, T.M.; Schuster, D.; Wolber, G.; Laggner, C.; Langer, T. High-throughput structure-based pharmacophore modelling as a basis for successful parallel virtual screening. *J. Comput. Aided Mol. Des.* **2006**, *20*, 703–715. [CrossRef]

168. Halperin, I.; Ma, B.; Wolfson, H.; Nussinov, R. Principles of docking: An overview of search algorithms and a guide to scoring functions. *Proteins* **2002**, *47*, 409–443. [CrossRef] [PubMed]

169. Leach, A.R.; Gillet, V.J.; Lewis, R.A.; Taylor, R. Three-dimensional pharmacophore methods in drug discovery. *J. Med. Chem.* **2010**, *53*, 539–558. [CrossRef] [PubMed]

170. Wermuth, C.G.; Ganellin, C.R.; Lindberg, P.; Mitscher, L.A. Glossary of terms used in medicinal chemistry (IUPAC Recommendations 1998). *Pure Appl. Chem.* **1998**, *70*, 1129. [CrossRef]

171. Hein, M.; Zilian, D.; Sotriffer, C.A. Docking compared to 3D-pharmacophores: The scoring function challenge. *Drug Discov. Today Technol.* **2010**, *7*, e229–e236. [CrossRef]

172. Hessler, G.; Baringhaus, K. The scaffold hopping potential of pharmacophores. *Drug Discov. Today Technol.* **2010**, *7*, e263–e269. [CrossRef]

173. Pagadala, N.S.; Syed, K.; Tuszynski, J. Software for molecular docking: A review. *Biophys. Rev.* **2017**, *9*, 91–102. [CrossRef]

174. Cross, J.B.; Thompson, D.C.; Rai, B.K.; Baber, J.C.; Fan, K.Y.; Hu, Y.; Humblet, C. Comparison of several molecular docking programs: Pose prediction and virtual screening accuracy. *J. Chem. Inf. Model.* **2009**, *49*, 1455–1474. [CrossRef]

175. Cummings, M.D.; DesJarlais, R.L.; Gibbs, A.C.; Mohan, V.; Jaeger, E.P. Comparison of automated docking programs as virtual screening tools. *J. Med. Chem.* **2005**, *48*, 962–976. [CrossRef]

176. Annamala, M.K.; Inampudi, K.K.; Guruprasad, L. Docking of phosphonate and trehalose analog inhibitors into *M. tuberculosis* mycolyltransferase Ag85C: Comparison of the two scoring fitness functions GoldScore and ChemScore, in the GOLD software. *Bioinformation* **2007**, *1*, 339–350. [CrossRef]

177. Perola, E.; Walters, W.P.; Charifson, P.S. A detailed comparison of current docking and scoring methods on systems of pharmaceutical relevance. *Proteins* **2004**, *56*, 235–249. [CrossRef]

178. Xu, W.; Lucke, A.J.; Fairlie, D.P. Comparing sixteen scoring functions for predicting biological activities of ligands for protein targets. *J. Mol. Graph. Model.* **2015**, *57*, 76–88. [CrossRef]

179. Billones, J.B.; Carrillo, M.C.; Organo, V.G.; Sy, J.B.; Clavio, N.A.; Macalino, S.J.; Emnacen, I.A.; Lee, A.P.; Ko, P.K.; Concepcion, G.P. In silico discovery and in vitro activity of inhibitors against *Mycobacterium tuberculosis* 7,8-diaminopelargonic acid synthase (Mtb BioA). *Drug Des. Devel. Ther.* **2017**, *11*, 563–574. [CrossRef]

180. Dassault Systèmes BIOVIA. Discovery Studio. Available online: https://www.3dsbiovia.com/products/collaborative-science/biovia-discovery-studio/ (accessed on 26 October 2019).

181. Huang, S.Y.; Grinter, S.Z.; Zou, X. Scoring functions and their evaluation methods for protein-ligand docking: Recent advances and future directions. *Phys. Chem. Chem. Phys.* **2010**, *12*, 12899–12908. [CrossRef]

182. Ericksen, S.S.; Wu, H.; Zhang, H.; Michael, L.A.; Newton, M.A.; Hoffmann, F.M.; Wildman, S.A. Machine Learning Consensus Scoring Improves Performance Across Targets in Structure-Based Virtual Screening. *J. Chem. Inf. Model.* **2017**, *57*, 1579–1590. [CrossRef]

183. Li, D.D.; Meng, X.F.; Wang, Q.; Yu, P.; Zhao, L.G.; Zhang, Z.P.; Wang, Z.Z.; Xiao, W. Consensus scoring model for the molecular docking study of mTOR kinase inhibitor. *J. Mol. Graph. Model.* **2018**, *79*, 81–87. [CrossRef]

184. Charifson, P.S.; Corkery, J.J.; Murcko, M.A.; Walters, W.P. Consensus scoring: A method for obtaining improved hit rates from docking databases of three-dimensional structures into proteins. *J. Med. Chem.* **1999**, *42*, 5100–5109. [CrossRef]

185. Clark, R.D.; Strizhev, A.; Leonard, J.M.; Blake, J.F.; Matthew, J.B. Consensus scoring for ligand/protein interactions. *J. Mol. Graph. Model.* **2002**, *20*, 281–295. [CrossRef]

186. Harrison, A.J.; Yu, M.; Gardenborg, T.; Middleditch, M.; Ramsay, R.J.; Baker, E.N.; Lott, J.S. The structure of MbtI from *Mycobacterium tuberculosis*, the first enzyme in the biosynthesis of the siderophore mycobactin, reveals it to be a salicylate synthase. *J. Bacteriol.* **2006**, *188*, 6081–6091. [CrossRef]

187. Manos-Turvey, A.; Bulloch, E.M.; Rutledge, P.J.; Baker, E.N.; Lott, J.S.; Payne, R.J. Inhibition studies of *Mycobacterium tuberculosis* salicylate synthase (MbtI). *ChemMedChem* **2010**, *5*, 1067–1079. [CrossRef]

188. Vasan, M.; Neres, J.; Williams, J.; Wilson, D.J.; Teitelbaum, A.M.; Remmel, R.P.; Aldrich, C.C. Inhibitors of the salicylate synthase (MbtI) from *Mycobacterium tuberculosis* discovered by high-throughput screening. *ChemMedChem* **2010**, *5*, 2079–2087. [CrossRef]

189. Pini, E.; Poli, G.; Tuccinardi, T.; Chiarelli, L.R.; Mori, M.; Gelain, A.; Costantino, L.; Villa, S.; Meneghetti, F.; Barlocco, D. New Chromane-Based Derivatives as Inhibitors of *Mycobacterium tuberculosis* Salicylate Synthase (MbtI): Preliminary Biological Evaluation and Molecular Modeling Studies. *Molecules* **2018**, *23*, 1506. [CrossRef]

190. Chiarelli, L.R.; Mori, M.; Barlocco, D.; Beretta, G.; Gelain, A.; Pini, E.; Porcino, M.; Mori, G.; Stelitano, G.; Costantino, L.; et al. Discovery and development of novel salicylate synthase (MbtI) furanic inhibitors as antitubercular agents. *Eur. J. Med. Chem.* **2018**, *155*, 754–763. [CrossRef]

191. McGann, M. FRED pose prediction and virtual screening accuracy. *J. Chem. Inf. Model.* **2011**, *51*, 578–596. [CrossRef]

192. Korb, O.; Stützle, T.; Exner, T.E. PLANTS: Application of Ant Colony Optimization to Structure-Based Drug Design. In *ANTS 2006: Ant Colony Optimization and Swarm Intelligence*; Dorigo, M., Gambardella, L.M., Birattari, M., Martinoli, A., Poli, R., Stützle, T., Eds.; Springer: Heidelberg/Berlin, Germany, 2016; pp. 247–258.

193. Koshland, D.E. Application of a Theory of Enzyme Specificity to Protein Synthesis. *Proc. Natl. Acad. Sci. USA* **1958**, *44*, 98–104. [CrossRef]

194. Sotriffer, C.A. Accounting for induced-fit effects in docking: What is possible and what is not? *Curr. Top. Med. Chem.* **2011**, *11*, 179–191. [CrossRef]

195. Hartkoorn, R.C.; Sala, C.; Neres, J.; Pojer, F.; Magnet, S.; Mukherjee, R.; Uplekar, S.; Boy-Rottger, S.; Altmann, K.H.; Cole, S.T. Towards a new tuberculosis drug: Pyridomycin-nature's isoniazid. *EMBO Mol. Med.* **2012**, *4*, 1032–1042. [CrossRef]

196. Rozwarski, D.A.; Vilcheze, C.; Sugantino, M.; Bittman, R.; Sacchettini, J.C. Crystal structure of the *Mycobacterium tuberculosis* enoyl-ACP reductase, InhA, in complex with NAD+ and a C16 fatty acyl substrate. *J. Biol. Chem.* **1999**, *274*, 15582–15589. [CrossRef]

197. Rozwarski, D.A.; Grant, G.A.; Barton, D.H.; Jacobs, W.R., Jr.; Sacchettini, J.C. Modification of the NADH of the isoniazid target (InhA) from *Mycobacterium tuberculosis*. *Science* **1998**, *279*, 98–102. [CrossRef]

198. Amaro, R.E.; Li, W.W. Emerging methods for ensemble-based virtual screening. *Curr. Top. Med. Chem.* **2010**, *10*, 3–13. [CrossRef]

199. Brindha, S.; Sundaramurthi, J.C.; Velmurugan, D.; Vincent, S.; Gnanadoss, J.J. Docking-based virtual screening of known drugs against murE of *Mycobacterium tuberculosis* towards repurposing for TB. *Bioinformation* **2016**, *12*, 359–367. [CrossRef]

200. Wishart, D.S.; Knox, C.; Guo, A.C.; Cheng, D.; Shrivastava, S.; Tzur, D.; Gautam, B.; Hassanali, M. DrugBank: A knowledgebase for drugs, drug actions and drug targets. *Nucleic Acids Res.* **2008**, *36*, D901–D906. [CrossRef]

201. Schmidt, M.F.; Korb, O.; Howard, N.I.; Dias, M.V.B.; Blundell, T.L.; Abell, C. Discovery of Schaeffer's Acid Analogues as Lead Structures of *Mycobacterium tuberculosis* Type II Dehydroquinase Using a Rational Drug Design Approach. *ChemMedChem* **2013**, *8*, 54–58. [CrossRef]

202. Lovell, S.C.; Word, J.M.; Richardson, J.S.; Richardson, D.C. The penultimate rotamer library. *Proteins* **2000**, *40*, 389–408. [CrossRef]

203. Korb, O.; Stutzle, T.; Exner, T.E. Empirical scoring functions for advanced protein-ligand docking with PLANTS. *J. Chem. Inf. Model.* **2009**, *49*, 84–96. [CrossRef]

204. Bhabha, G.; Biel, J.T.; Fraser, J.S. Keep on moving: Discovering and perturbing the conformational dynamics of enzymes. *Acc. Chem. Res.* **2015**, *48*, 423–430. [CrossRef]

205. Goh, C.S.; Milburn, D.; Gerstein, M. Conformational changes associated with protein-protein interactions. *Curr. Opin. Struct. Biol.* **2004**, *14*, 104–109. [CrossRef]

206. Hospital, A.; Goni, J.R.; Orozco, M.; Gelpi, J.L. Molecular dynamics simulations: Advances and applications. *Adv. Appl. Bioinform. Chem.* **2015**, *8*, 37–47.

207. Lee, Y.; Jeong, L.S.; Choi, S.; Hyeon, C. Link between allosteric signal transduction and functional dynamics in a multisubunit enzyme: S-adenosylhomocysteine hydrolase. *J. Am. Chem. Soc.* **2011**, *133*, 19807–19815. [CrossRef]

208. McCammon, J.A.; Gelin, B.R.; Karplus, M. Dynamics of folded proteins. *Nature* **1977**, *267*, 585–590. [CrossRef]

209. Larsson, P.; Hess, B.; Lindahl, E. Algorithm improvements for molecular dynamics simulations. *WIREs Comput. Mol. Sci.* **2011**, *1*, 93–108. [CrossRef]

210. Orozco, M.; Orellana, L.; Hospital, A.; Naganathan, A.N.; Emperador, A.; Carrillo, O.; Gelpi, J.L. Coarse-grained representation of protein flexibility. Foundations, successes, and shortcomings. *Adv. Protein Chem. Struct. Biol.* **2011**, *85*, 183–215.

211. Linge, J.P.; Williams, M.A.; Spronk, C.A.; Bonvin, A.M.; Nilges, M. Refinement of protein structures in explicit solvent. *Proteins* **2003**, *50*, 496–506. [CrossRef]

212. Anandakrishnan, R.; Drozdetski, A.; Walker, R.C.; Onufriev, A.V. Speed of conformational change: Comparing explicit and implicit solvent molecular dynamics simulations. *Biophys. J.* **2015**, *108*, 1153–1164. [CrossRef]

213. MacKerell, A.D.; Bashford, D.; Bellott, M.; Dunbrack, R.L.; Evanseck, J.D.; Field, M.J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; et al. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **1998**, *102*, 3586–3616. [CrossRef]

214. Cornell, W.D.; Cieplak, P.; Bayly, C.I.; Gould, I.R.; Merz, K.M.; Ferguson, D.M.; Spellmeyer, D.C.; Fox, T.; Caldwell, J.W.; Kollman, P.A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197. [CrossRef]

215. Oostenbrink, C.; Villa, A.; Mark, A.E.; van Gunsteren, W.F. A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. *J. Comput. Chem.* **2004**, *25*, 1656–1676. [CrossRef]

216. Jorgensen, W.L.; Maxwell, D.S.; Tirado-Rives, J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236. [CrossRef]

217. Kaminski, G.A.; Friesner, R.A.; Tirado-Rives, J.; Jorgensen, W.L. Evaluation and Reparametrization of the OPLS-AA Force Field for Proteins via Comparison with Accurate Quantum Chemical Calculations on Peptides†. *J. Phys. Chem. B* **2001**, *105*, 6474–6487. [CrossRef]

218. Daggett, V.; Levitt, M. Protein Unfolding Pathways Explored Through Molecular Dynamics Simulations. *J. Mol. Biol.* **1993**, *232*, 600–619. [CrossRef]

219. Alonso, H.; Bliznyuk, A.A.; Gready, J.E. Combining docking and molecular dynamic simulations in drug design. *Med. Res. Rev.* **2006**, *26*, 531–568. [CrossRef]

220. Papaleo, E. Integrating atomistic molecular dynamics simulations, experiments, and network analysis to study protein dynamics: Strength in unity. *Front. Mol. Biosci.* **2015**, *2*, 28. [CrossRef]

221. Prada-Gracia, D.; Gomez-Gardenes, J.; Echenique, P.; Falo, F. Exploring the free energy landscape: From dynamics to networks and back. *PLoS Comput. Biol.* **2009**, *5*, e1000415. [CrossRef]

222. Wahab, H.A.; Choong, Y.S.; Ibrahim, P.; Sadikun, A.; Scior, T. Elucidating isoniazid resistance using molecular modeling. *J. Chem. Inf. Model.* **2009**, *49*, 97–107. [CrossRef]

223. Schroeder, E.K.; Basso, L.A.; Santos, D.S.; de Souza, O.N. Molecular dynamics simulation studies of the wild-type, I21V, and I16T mutants of isoniazid-resistant *Mycobacterium tuberculosis* enoyl reductase (InhA) in complex with NADH: Toward the understanding of NADH-InhA different affinities. *Biophys. J.* **2005**, *89*, 876–884. [CrossRef]

224. Cruz, J.N.; Costa, J.F.S.; Khayat, A.S.; Kuca, K.; Barros, C.A.L.; Neto, A. Molecular dynamics simulation and binding free energy studies of novel leads belonging to the benzofuran class inhibitors of *Mycobacterium tuberculosis* Polyketide Synthase 13. *J. Biomol. Struct. Dyn.* **2019**, *37*, 1616–1627. [CrossRef]

225. Aggarwal, A.; Parai, M.K.; Shetty, N.; Wallis, D.; Woolhiser, L.; Hastings, C.; Dutta, N.K.; Galaviz, S.; Dhakal, R.C.; Shrestha, R.; et al. Development of a Novel Lead that Targets, *M. tuberculosis* Polyketide Synthase 13. *Cell* **2017**, *170*, 249–259. [CrossRef]

226. Nikolova, N.; Jaworska, J. Approaches to Measure Chemical Similarity—A Review. *QSAR Comb. Sci.* **2003**, *22*, 1006–1026. [CrossRef]

227. Johnson, M.A.; Maggiora, G.M. American Chemical Society. In *Concepts and Applications of Molecular Similarity*; Wiley: New York, NY, USA, 1990.

228. Bacilieri, M.; Moro, S. Ligand-based drug design methodologies in drug discovery process: An overview. *Curr. Drug Discov. Technol.* **2006**, *3*, 155–165. [CrossRef]

229. Sukumar, N.; Das, S. Current trends in virtual high throughput screening using ligand-based and structure-based methods. *Comb. Chem. High Throughput Screen.* **2011**, *14*, 872–888. [CrossRef]

230. Cereto-Massague, A.; Ojeda, M.J.; Valls, C.; Mulero, M.; Garcia-Vallve, S.; Pujadas, G. Molecular fingerprint similarity search in virtual screening. *Methods* **2015**, *71*, 58–63. [CrossRef]

231. Bajusz, D.; Racz, A.; Heberger, K. Why is Tanimoto index an appropriate choice for fingerprint-based similarity calculations? *J. Cheminform.* **2015**, *7*, 20. [CrossRef]

232. Faulon, J.-L.; Bender, A. *Handbook of Chemoinformatics Algorithms*; Chapman & Hall/CRC: Boca Raton, FL, USA, 2010; 440p.

233. Haranczyk, M.; Holliday, J. Comparison of similarity coefficients for clustering and compound selection. *J. Chem. Inf. Model.* **2008**, *48*, 498–508. [CrossRef]

234. Al Khalifa, A.; Haranczyk, M.; Holliday, J. Comparison of nonbinary similarity coefficients for similarity searching, clustering and compound selection. *J. Chem. Inf. Model.* **2009**, *49*, 1193–1201. [CrossRef]

235. Ginn, C.M.R.; Willett, P.; Bradshaw, J. Combination of molecular similarity measures using data fusion. *Perspect. Drug Discov. Des.* **2000**, *20*, 1–16. [CrossRef]

236. Medina-Franco, J.L.; Maggiora, G.M.; Giulianotti, M.A.; Pinilla, C.; Houghten, R.A. A similarity-based data-fusion approach to the visual characterization and comparison of compound databases. *Chem. Biol. Drug Des.* **2007**, *70*, 393–412. [CrossRef]

237. Hu, G.; Kuang, G.; Xiao, W.; Li, W.; Liu, G.; Tang, Y. Performance evaluation of 2D fingerprint and 3D shape similarity methods in virtual screening. *J. Chem. Inf. Model.* **2012**, *52*, 1103–1113. [CrossRef]

238. Drwal, M.N.; Griffith, R. Combination of ligand- and structure-based methods in virtual screening. *Drug Discov. Today Technol.* **2013**, *10*, e395–e401. [CrossRef]

239. Hu, Y.; Bajorath, J. Extending the activity cliff concept: Structural categorization of activity cliffs and systematic identification of different types of cliffs in the ChEMBL database. *J. Chem. Inf. Model.* **2012**, *52*, 1806–1811. [CrossRef]

240. Hu, Y.; Stumpfe, D.; Bajorath, J. Advancing the activity cliff concept. *F1000Res* **2013**, *2*, 199. [CrossRef]

241. Stumpfe, D.; de la Vega de Leon, A.; Dimova, D.; Bajorath, J. Advancing the activity cliff concept, part II. *F1000Res* **2014**, *3*, 75. [CrossRef]

242. Verma, J.; Khedkar, V.M.; Coutinho, E.C. 3D-QSAR in drug design—A review. *Curr. Top. Med. Chem.* **2010**, *10*, 95–115. [CrossRef]

243. Exploring QSAR. *Environ. Sci. Technol.* **1995**, *29*, 444A. [CrossRef]

244. Bostrom, J.; Norrby, P.O.; Liljefors, T. Conformational energy penalties of protein-bound ligands. *J. Comput. Aided Mol. Des.* **1998**, *12*, 383–396. [CrossRef]

245. Perola, E.; Charifson, P.S. Conformational analysis of drug-like molecules bound to proteins: An extensive study of ligand reorganization upon binding. *J. Med. Chem.* **2004**, *47*, 2499–2510. [CrossRef]

246. Melo-Filho, C.C.; Braga, R.C.; Andrade, C.H. 3D-QSAR approaches in drug design: Perspectives to generate reliable CoMFA models. *Curr. Comput. Aided Drug Des.* **2014**, *10*, 148–159. [CrossRef]

247. Cherkasov, A.; Muratov, E.N.; Fourches, D.; Varnek, A.; Baskin, II; Cronin, M.; Dearden, J.; Gramatica, P.; Martin, Y.C.; Todeschini, R.; et al. QSAR modeling: Where have you been? Where are you going to? *J. Med. Chem.* **2014**, *57*, 4977–5010. [CrossRef]

248. Cramer, R.D.; Patterson, D.E.; Bunce, J.D. Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. *J. Am. Chem. Soc.* **1988**, *110*, 5959–5967. [CrossRef]

249. Klebe, G.; Abraham, U.; Mietzner, T. Molecular similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological activity. *J. Med. Chem.* **1994**, *37*, 4130–4146. [CrossRef]

250. Bajpai, A.; Agarwal, N.; Gupta, S.P. A comparative 2D QSAR study on a series of hydroxamic acid-based histone deacetylase inhibitors vis-a-vis comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA). *Indian J. Biochem. Biophys.* **2014**, *51*, 244–252.

251. Chhatbar, D.M.; Chaube, U.J.; Vyas, V.K.; Bhatt, H.G. CoMFA, CoMSIA, Topomer CoMFA, HQSAR, molecular docking and molecular dynamics simulations study of triazine morpholino derivatives as mTOR inhibitors for the treatment of breast cancer. *Comput. Biol. Chem.* **2019**, *80*, 351–363. [CrossRef]

252. Singh, S.; Supuran, C.T. 3D-QSAR CoMFA studies on sulfonamide inhibitors of the Rv3588c beta-carbonic anhydrase from *Mycobacterium tuberculosis* and design of not yet synthesized new molecules. *J. Enzym. Inhib. Med. Chem.* **2014**, *29*, 449–455. [CrossRef]

253. Punkvang, A.; Hannongbua, S.; Saparpakorn, P.; Pungpo, P. Insight into the structural requirements of aminopyrimidine derivatives for good potency against both purified enzyme and whole cells of *M. tuberculosis*: Combination of HQSAR, CoMSIA, and MD simulation studies. *J. Biomol. Struct. Dyn.* **2016**, *34*, 1079–1091. [CrossRef] [PubMed]

254. Schuster, D. 3D pharmacophores as tools for activity profiling. *Drug Discov. Today Technol.* **2010**, *7*, e205–e211. [CrossRef] [PubMed]

255. Tawari, N.R.; Degani, M.S. Predictive models for nucleoside bisubstrate analogs as inhibitors of siderophore biosynthesis in *Mycobacterium tuberculosis*: Pharmacophore mapping and chemometric QSAR study. *Mol. Divers.* **2011**, *15*, 435–444. [CrossRef]

256. Hohenberg, P.; Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev.* **1964**, *136*, B864–B871. [CrossRef]

257. Kohn, W.; Sham, L.J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **1965**, *140*, A1133–A1138. [CrossRef]

258. Sharma, S. *Molecular Dynamics Simulation of Nanocomposites Using BIOVIA Materials Studio, Lammps and Gromacs*, 1st ed.; Elsevier: Waltham, MA, USA, 2019.

259. Fiolhais, C.; Nogueira, F.; Marques, M.A.L. *A Primer in Density Functional Theory*; Springer: Berlin, Germany; New York, NY, USA, 2003.

260. Becke, A.D. Perspective: Fifty years of density-functional theory in chemical physics. *J. Chem. Phys.* **2014**, *140*, 18A301. [CrossRef]

261. Rabi, S.; Patel, A.H.G.; Burger, S.K.; Verstraelen, T.; Ayers, P.W. Exploring the substrate selectivity of human sEH and *M. tuberculosis* EHB using QM/MM. *Struct. Chem.* **2017**, *28*, 1501–1511. [CrossRef]

262. Ramalho, T.C.; Caetano, M.S.; Josa, D.; Luz, G.P.; Freitas, E.A.; da Cunha, E.F. Molecular modeling of *Mycobacterium tuberculosis* dUTpase: Docking and catalytic mechanism studies. *J. Biomol. Struct. Dyn.* **2011**, *28*, 907–917. [CrossRef]

263. Oliveira, C.G.; da, S.M.P.I.; Souza, P.C.; Pavan, F.R.; Leite, C.Q.; Viana, R.B.; Batista, A.A.; Nascimento, O.R.; Deflon, V.M. Manganese(II) complexes with thiosemicarbazones as potential anti-*Mycobacterium tuberculosis* agents. *J. Inorg. Biochem.* **2014**, *132*, 21–29. [CrossRef]

264. Chi, G.; Manos-Turvey, A.; O'Connor, P.D.; Johnston, J.M.; Evans, G.L.; Baker, E.N.; Payne, R.J.; Lott, J.S.; Bulloch, E.M. Implications of binding mode and active site flexibility for inhibitor potency against the salicylate synthase from *Mycobacterium tuberculosis*. *Biochemistry* **2012**, *51*, 4868–4879. [CrossRef]

265. Frisch, M.J.; Trucks, G.W.; Schlegel, H.B.; Scuseria, G.E.; Robb, M.A.; Cheeseman, J.R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G.A.; et al. *Gaussian 09, Revision E.01*; Gaussian, Inc.: Wallingford, CT, USA, 2009.

266. Stephens, P.J.; Devlin, F.J.; Chabalowski, C.F.; Frisch, M.J. Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *J. Phys. Chem.* **1994**, *98*, 11623–11627. [CrossRef]

267. Becke, A.D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652. [CrossRef]

268. Indarto, A. *Theoretical Modelling and Mechanistic Study of the Formation and Atmospheric Transformations of Polycyclic Aromatic Compounds and Carbonaceous Particles*; Universal Publishers: Irvine, CA, USA, 2010.

269. Hamada, I. van der Waals density functional made accurate. *Phys. Rev. B* **2014**, *89*, 121103. [CrossRef]

270. Berland, K.; Cooper, V.R.; Lee, K.; Schroder, E.; Thonhauser, T.; Hyldgaard, P.; Lundqvist, B.I. van der Waals forces in density functional theory: A review of the vdW-DF method. *Rep. Prog. Phys.* **2015**, *78*, 066501. [CrossRef]

271. Grimme, S. Accurate description of van der Waals complexes by density functional theory including empirical corrections. *J. Comput. Chem.* **2004**, *25*, 1463–1473. [CrossRef]

272. Cohen, A.J.; Mori-Sanchez, P.; Yang, W. Insights into current limitations of density functional theory. *Science* **2008**, *321*, 792–794. [CrossRef]

273. Wilson, G.L.; Lill, M.A. Integrating structure-based and ligand-based approaches for computational drug design. *Future Med. Chem.* **2011**, *3*, 735–750. [CrossRef]

274. Polgar, T.; Keseru, G.M. Integration of virtual and high throughput screening in lead discovery settings. *Comb. Chem. High Throughput Screen.* **2011**, *14*, 889–897. [CrossRef]

275. Tanrikulu, Y.; Kruger, B.; Proschak, E. The holistic integration of virtual screening in drug discovery. *Drug Discov. Today* **2013**, *18*, 358–364. [CrossRef]

276. Tan, L.; Geppert, H.; Sisay, M.T.; Gutschow, M.; Bajorath, J. Integrating structure- and ligand-based virtual screening: Comparison of individual, parallel, and fused molecular docking and similarity search calculations on multiple targets. *ChemMedChem* **2008**, *3*, 1566–1571. [CrossRef]

277. Huang, S.Y.; Li, M.; Wang, J.; Pan, Y. HybridDock: A Hybrid Protein-Ligand Docking Protocol Integrating Protein- and Ligand-Based Approaches. *J. Chem. Inf. Model.* **2016**, *56*, 1078–1087. [CrossRef]

278. Lam, P.C.; Abagyan, R.; Totrov, M. Ligand-biased ensemble receptor docking (LigBEnD): A hybrid ligand/receptor structure-based approach. *J. Comput. Aided Mol. Des.* **2018**, *32*, 187–198. [CrossRef]

279. Mestres, J.; Knegtel, R.M.A. Similarity versus docking in 3D virtual screening. *Perspect. Drug Discov. Des.* **2000**, *20*, 191–207. [CrossRef]

280. Kruger, D.M.; Evers, A. Comparison of structure- and ligand-based virtual screening protocols considering hit list complementarity and enrichment factors. *ChemMedChem* **2010**, *5*, 148–158. [CrossRef]

281. Billones, J.B.; Carrillo, M.C.; Organo, V.G.; Macalino, S.J.; Sy, J.B.; Emnacen, I.A.; Clavio, N.A.; Concepcion, G.P. Toward antituberculosis drugs: In silico screening of synthetic compounds against *Mycobacterium tuberculosis*l,d-transpeptidase 2. *Drug Des. Devel Ther.* **2016**, *10*, 1147–1157. [CrossRef]

282. Fakhar, Z.; Govender, T.; Maguire, G.E.M.; Lamichhane, G.; Walker, R.C.; Kruger, H.G.; Honarparvar, B. Differential flap dynamics in l,d-transpeptidase2 from *Mycobacterium tuberculosis* revealed by molecular dynamics. *Mol. Biosyst.* **2017**, *13*, 1223–1234. [CrossRef]

283. Sandhu, P.; Akhter, Y. The drug binding sites and transport mechanism of the RND pumps from *Mycobacterium tuberculosis*: Insights from molecular dynamics simulations. *Arch. Biochem. Biophys.* **2016**, *592*, 38–49. [CrossRef]

284. Shah, P.; Mistry, J.; Reche, P.A.; Gatherer, D.; Flower, D.R. In silico design of *Mycobacterium tuberculosis* epitope ensemble vaccines. *Mol. Immunol.* **2018**, *97*, 56–62. [CrossRef]

285. Li, D.; Chi, B.; Wang, W.-W.; Gao, J.-M.; Wan, J. Exploring the possible binding mode of trisubstituted benzimidazoles analogues in silico for novel drug designtargeting Mtb FtsZ. *Med. Chem. Res.* **2017**, *26*, 153–169. [CrossRef]

286. Spitzer, R.; Jain, A.N. Surflex-Dock: Docking benchmarks and real-world application. *J. Comput. Aided Mol. Des.* **2012**, *26*, 687–699. [CrossRef]

287. Villoutreix, B.O.; Eudes, R.; Miteva, M.A. Structure-based virtual ligand screening: Recent success stories. *Comb. Chem. High Throughput Screen.* **2009**, *12*, 1000–1016. [CrossRef]

288. Talele, T.T.; Khedkar, S.A.; Rigby, A.C. Successful applications of computer aided drug discovery: Moving drugs from concept to the clinic. *Curr. Top. Med. Chem.* **2010**, *10*, 127–141. [CrossRef]

289. Clark, D.E. What has virtual screening ever done for drug discovery? *Expert Opin. Drug Discov.* **2008**, *3*, 841–851. [CrossRef]

290. Scior, T.; Bender, A.; Tresadern, G.; Medina-Franco, J.L.; Martinez-Mayorga, K.; Langer, T.; Cuanalo-Contreras, K.; Agrafiotis, D.K. Recognizing pitfalls in virtual screening: A critical review. *J. Chem. Inf. Model.* **2012**, *52*, 867–881. [CrossRef]

291. Baig, M.H.; Ahmad, K.; Roy, S.; Ashraf, J.M.; Adil, M.; Siddiqui, M.H.; Khan, S.; Kamal, M.A.; Provaznik, I.; Choi, I. Computer Aided Drug Design: Success and Limitations. *Curr. Pharm. Des.* **2016**, *22*, 572–581. [CrossRef]

292. Coupez, B.; Lewis, R.A. Docking and scoring—Theoretically easy, practically impossible? *Curr. Med. Chem.* **2006**, *13*, 2995–3003.

293. Geppert, H.; Vogt, M.; Bajorath, J. Current trends in ligand-based virtual screening: Molecular representations, data mining methods, new application areas, and performance evaluation. *J. Chem. Inf. Model.* **2010**, *50*, 205–216. [CrossRef]

294. Jain, A.N.; Nicholls, A. Recommendations for evaluation of computational methods. *J. Comput. Aided Mol. Des.* **2008**, *22*, 133–139. [CrossRef]

295. Lindorff-Larsen, K.; Maragakis, P.; Piana, S.; Shaw, D.E. Picosecond to Millisecond Structural Dynamics in Human Ubiquitin. *J. Phys. Chem. B* **2016**, *120*, 8313–8320. [CrossRef] [PubMed]

296. Noe, F. Beating the millisecond barrier in molecular dynamics simulations. *Biophys. J.* **2015**, *108*, 228–229. [CrossRef] [PubMed]

297. Shi, J.; Nobrega, R.P.; Schwantes, C.; Kathuria, S.V.; Bilsel, O.; Matthews, C.R.; Lane, T.J.; Pande, V.S. Atomistic structural ensemble refinement reveals non-native structure stabilizes a sub-millisecond folding intermediate of CheY. *Sci. Rep.* **2017**, *7*, 44116. [CrossRef]

298. Fujita, T. Recent Success Stories Leading to Commercializable Bioactive Compounds with the Aid of Traditional QSAR Procedures. *QSAR* **1997**, *16*, 107–112. [CrossRef]

299. Gao, Q.; Yang, L.; Zhu, Y. Pharmacophore based drug design approach as a practical process in drug discovery. *Curr. Comput. Aided Drug Des.* **2010**, *6*, 37–49. [CrossRef]

300. Sardari, S.; Dezfulian, M. Cheminformatics in anti-infective agents discovery. *Mini Rev. Med. Chem.* **2007**, *7*, 181–189. [CrossRef]

301. Topol, E.J. High-performance medicine: The convergence of human and artificial intelligence. *Nat. Med.* **2019**, *25*, 44–56. [CrossRef]