We are IntechOpen, the world's leading publisher of Open Access books
Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

154
Countries delivered to

Our authors are among the
TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Has Molecular Docking Ever Brought us a Medicine?

Mark Andrew Phillips, Marisa A. Stewart,
Darby L. Woodling and Zhong-Ru Xie

Additional information is available at the end of the chapter

Abstract

Molecular docking has been developed and improving for many years, but its ability to bring a medicine to the drug market effectively is still generally questioned. In this chapter, we introduce several successful cases including drugs for treatment of HIV, cancers, and other prevalent diseases. The technical details such as docking software, protein data bank (PDB) structures, and other computational methods employed are also collected and displayed. In most of the cases, the structures of drugs or drug candidates and the interacting residues on the target proteins are also presented. In addition, a few successful examples of drug repurposing using molecular docking are mentioned in this chapter. It should provide us with confidence that the docking will be extensively employed in the industry and basic research. Moreover, we should actively apply molecular docking and related technology to create new therapies for diseases.

Keywords: computational drug design, molecular docking, drug repurposing

1. Introduction

Molecular docking is one of many computational tools that can be used in drug discovery [1–4]. It is a form of structure-based drug discovery that quantifies the binding affinities between small molecules and macromolecular targets (proteins). The first step in molecular docking is choosing a drug target. Any macromolecule can be used as a target; some very common targets include enzymes and regulatory elements. Next, the three-dimensional structure must be determined or predicted; high resolution structures can be determined using X-rays, NMR, or electron microscopy (EM). Thousands of popular targets have solved structures available on the protein data bank (PDB) [5]. Many drug targets have known binding sites; if not, software that can predict potential binding sites for different ligands have been developed. Docking

studies can be performed using known ligands (naturally occurring molecules or known drugs) or novel ligands. Virtual screening (i.e. identifying novel ligands with molecular docking) provides an extremely useful (but time consuming) method of drug discovery because molecules can be designed to have high binding affinity to a very specific site. Docking studies are often validated using further computational methods, such as molecular dynamic simulation. The most successful candidates from computational trials can be tested *in vitro* or *in vivo*, and eventually progress to clinical trials (**Figure 1**).
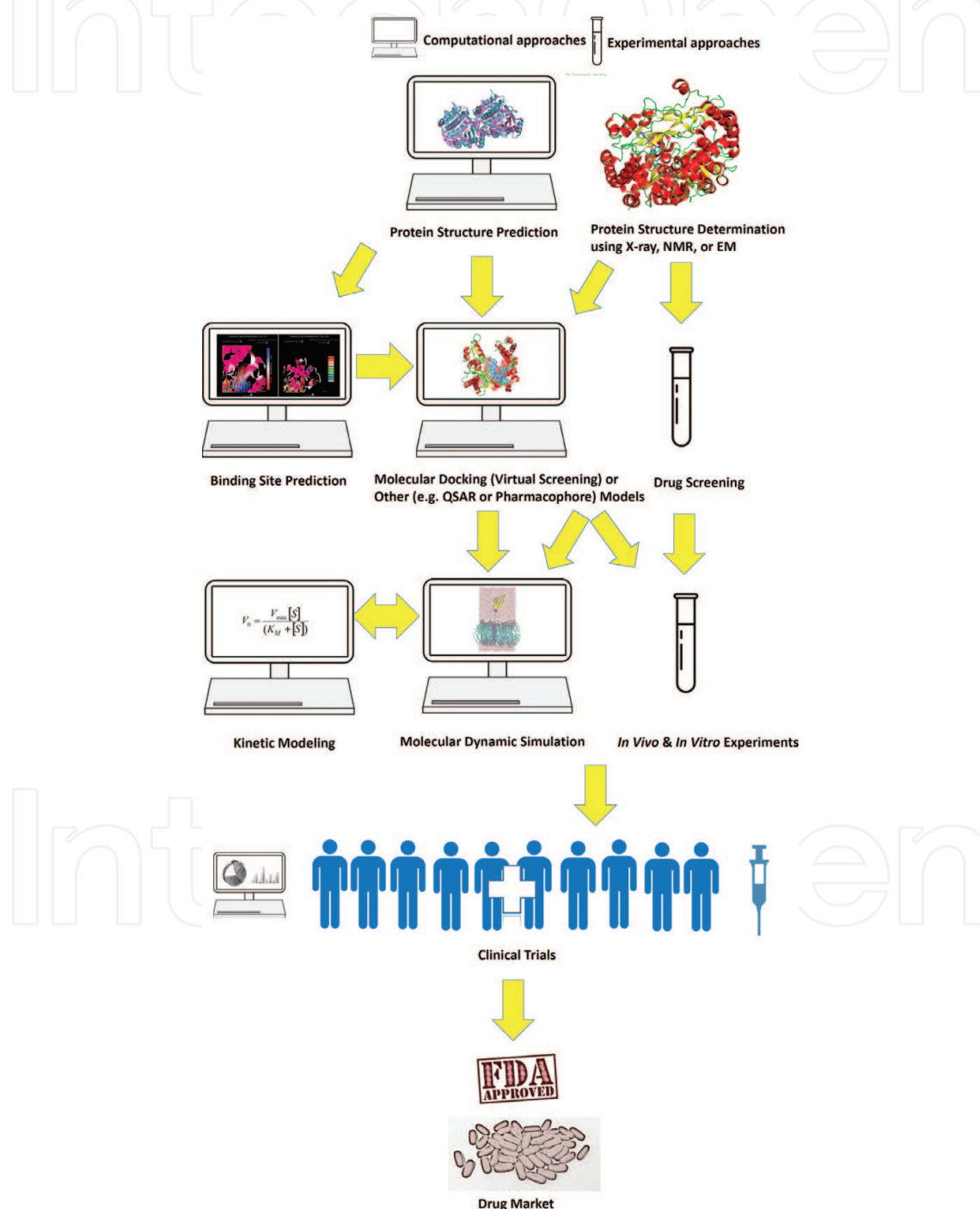

Figure 1. A brief flowchart of novel drug discovery procedure.

It is believed that a searching algorithm, which assists in thoroughly and efficiently exploring possible positions, orientations and conformations of potential drugs and the target proteins, and a scoring function, which assists in precisely and correctly identifying the most energetically favorable binding poses, are two most important components of a molecular docking programs. However, some other factors will affect the effectiveness and accuracy of molecular docking, such as the availability and quality of a determined or predicted structure of the target protein, the conformational changes of the target proteins after the drug binding, and the identification of potential binding sites. As those mentioned in previous chapters, many commercial and academic docking search algorithms, scoring functions, and software packages have been developed and improved in the past decades. However, it is still questioned if there are any successful stories in which molecular docking have helped to bring a drug to the market.

Although many molecular docking algorithms have been developed and improved for many decades, biomedical laboratories or pharmaceutical companies used to be hesitant to apply this technology to drug screening. Here are some possible reasons:

1. The "force fields" which describe the intra- and inter-molecular interaction energies were not accurate and precise enough to estimate or calculate the binding affinities between proteins and potential binding drugs.

2. The computer was not "fast" enough to calculate the interacting energy of many possible binding "conformations" of one or many possible binding compound(s) using a sophisticated model taking account into all the factors, components, and conditions of molecular interactions.

3. The number of binding complex structures was not large enough and the resolution of available structures was not good.

4. The searching/sampling algorithms to explore the possible binding orientations and conformations were not efficient to identify possible binding poses with reasonable time.

These reasons and concerns are all tightly cross-linked together, and, fortunately, have been dramatically improved in the past years. For example, the number of structures on PDB has increased from 47,605 to 133,759 since 2007 [6]. The resolution of determined structures has significantly improved. Therefore, the accuracy of both physics-based and knowledge-based scoring functions which assist researchers in identifying the most energy favorable binding poses and estimating binding affinities have been improved. The substantial improvement in both computer hardware and software also make it possible to screen a large number of natural and artificial compounds and search the best binding poses efficiently.

When we attempt to dock a compound to a target protein, often we need to use other computational methods before docking or in parallel. For instance, we may need to do structure prediction if the structure of the target protein has not yet been determined. The accumulated PDB structures with good resolution and the accurate structure prediction algorithms make it possible for researchers to obtain reliable structural models to perform molecular docking experiments. The enhanced quantity, quality, and diversity of protein-compound complex structures provide solid basis for creation of accurate binding site

prediction methods, and they help reduce the searching surface area on the target proteins for docking algorithms [7–9]. Other computational methods such as pharmacophore and quantitative structure-activity relationship (QSAR) models can be used prior to the molecular docking to reduce computational load and time [10–12]. In summary, the technology of molecular docking has matured and been applied in different stages of the drug discovery process. The successful stories have not been mentioned often and are not widely known. They will be introduced in this chapter.

2. Identification of medicine for HIV

The human immunodeficiency virus (HIV) epidemic around the world has pushed massive amounts of money into research that looks for ways to help treat and prevent this virus. Because bringing a drug into the market can take many years and cost astronomical amounts of money, it is of the utmost importance of researchers to use a cost effective ways to find these new therapeutics. Computational methods have been gradually becoming commonplace in drug design research. These methods have been either confirming established research, discovering new compounds, binding sites or conformations, and even allowing for the repurposing of the drug to treat other illnesses. HIV research has seen an influx of multiple computational methods being used to confirm discoveries of previous studies and establish new ones. Methods such as docking and molecular dynamics are saving researchers valuable time. These methods are also allowing research to make accurate and precise predictions of what is going on at the molecular level. While computational drug design methods are nowhere near replacing *in vitro* and *in vivo* testing, *in silico* testing is becoming increasingly popular for researchers to validate their research or act as a starting point for *in vitro* testing. This section will introduce how researchers used computational methods to help identify drugs for HIV-1 Protease and HIV-1 Integrase. It will also discuss how these methods are being utilized for future developments in this area of research, and how researchers were able to use the drugs Saquinavir and Nelfinavir toward treating a disease unrelated to HIV—Chagas.

2.1. Human immunodeficiency virus

Acquired immunodeficiency syndrome (AIDS) is acquired in humans by the retrovirus HIV [13]. HIV infects important helper T cells in the human immune system—specifically CD4+ T cells [14]. HIV is transmitted as positive-sense, single-stranded, enveloped RNA virus. There are currently two types of HIV that have been characterized as HIV-1 and HIV-2. HIV-1 was the first HIV virus discovered and it is more virulent and more infective than HIV-2 [15]. After the viral capsid has entered the cell, an enzyme called reverse transcriptase liberates the positive-sense RNA from the viral proteins and copies it into a complimentary DNA molecule [16]. The reverse transcriptase process is very prone to errors. This characteristic results in many mutations that make this component of HIV likely to encounter drug resistance. For this reason, HIV reverse transcriptase is an unlikely target for HIV therapeutics. The newly formed circular DNA strand and its complement form a double-stranded

viral DNA that is transported to the nucleus. The integration of the viral DNA into the host's genome is carried out by the integrase enzyme [16]. The HIV virus then may remain dormant or continue to assemble new HIV-1 virions. The plasma membrane of the host cell is the site for the production of new HIV-1 virions. The virion buds that are produced at the plasma membrane are cleaved by HIV-1 protease enzyme. Once the bud has been cleaved by HIV-1 protease, the internal components can assemble, and in turn create a virion capable of infecting other cells. The two targets that computational drug researchers have focused on significantly are HIV-1 integrase and HIV-1 protease.

2.2. HIV-1 integrase

Integrase (IN) is a retrovirus enzyme not exclusive to only HIV. This protein allows the genetic material of the virus to be integrated into the DNA of the host cell. Integration occurs after the double-stranded viral DNA is produced by reverse transcriptase. Once integration has commenced for a cell, there is no turning back. The cell is now considered a pro-virus, and it is now a permanent carrier of the virus. In general, retroviral integrases catalyze two reactions. Both reactions are catalyzed by the same active site on the enzyme and occur via transesterification.

2.2.1. HIV-1 integrase inhibitor—Raltegravir and its ensuing analogs

The most common inhibitors for integrase are referred to as integrase strand transfer inhibitors (INSTIs). Mg^{2+} and Mn^{2+} are critical cofactors in the integration phase [17], and inactivating these cofactors causes functional impairment of integrase. Most HIV-1 INSTIs contain a structural motif that coordinates the two divalent magnesium ions in the enzyme's active site [17]. Researchers screen over 250,000 compounds to yield potent inhibitors [18]. The most active inhibitors seemed to contain a distinct beta-diketo acid (DKA) moiety [19]. This moiety had the ability to coordinate metal ions within the IN active site. There was similar antiviral activity when the DKA pharmacophore was transferred to a naphthyridine carboxamide core [20]. A class of N-alkyl hydroxypyridinone carboxylic acids was the result of the success with the diketo acid structural analogs. These new analogs had a good pharmacokinetic profile in rats [21]. The drug, MK-0518, also known as Raltegravir, became the most promising pyrimidinone carboxamide derivative. Raltegravir was the first integrase inhibitor to progress to Phase III clinical trials. While there have been multiple resistant mutations for both treatment-experienced and treatment-naïve patients, Raltegravir still proved to be an effective IN inhibitor [22]. In October 2007, Raltegravir became the first FDA-approved IN inhibitor (**Table 1**).

To bring a single drug to the market, it can cost upwards of $2 billion [23]. Even with this, only one in three drugs will generate enough revenue to cover the cost of the research and development of the drug [24]. Pharmaceutical researchers and executives can see the allure of modifying current leads on drugs, rather than trying to design a new drug. "Me-too" drugs [18] can create an optimized drug and create vital marketplace competition, but many argue that slight modifications are producing negligible improvements [25]. "Me-too" drug emergence has seen a surge in the HIV-1 integrase inhibitor market. While Raltegravir has become the known and widely used anti-HIV drug, amino acid mutations have already conferred robust viral resistance of the drug [26]. This viral drug resistance normally occurs when one

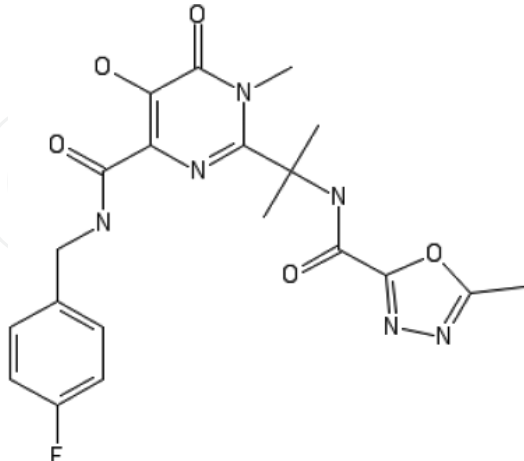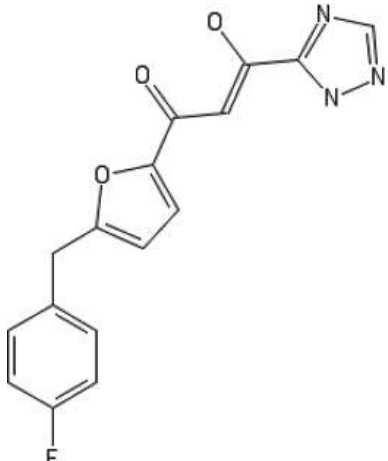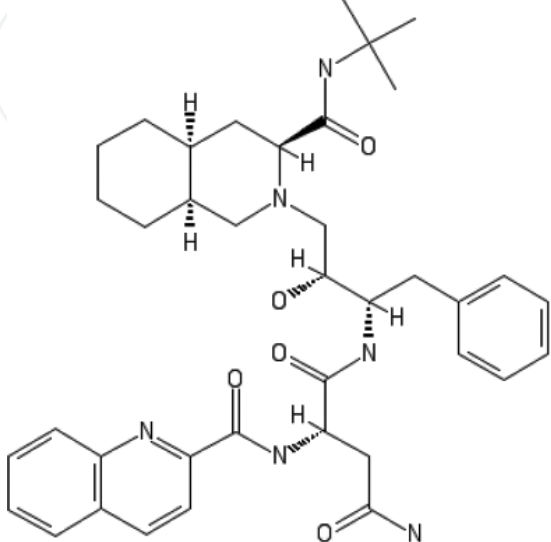**Table of various HIV-1 integrase and HIV-1 protease inhibitors**

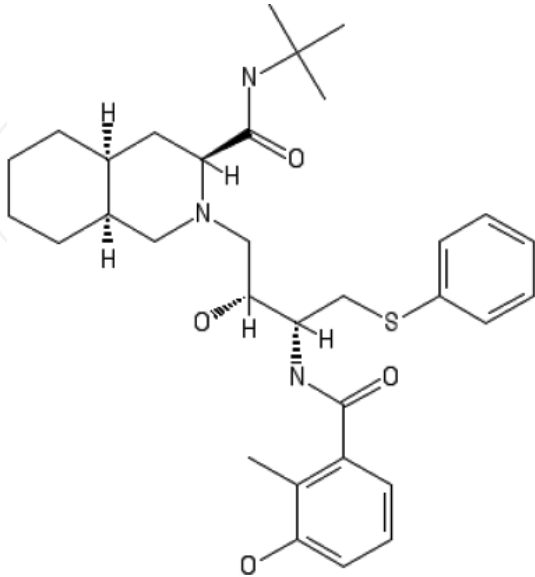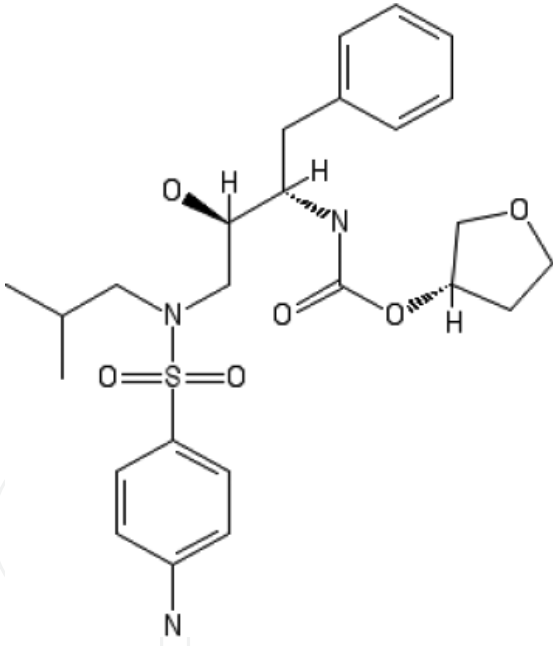| Drug	Type of inhibitor	Affected residues	Structure
Raltegravir	HIV-integrase	D64 T66 E92 D116 Y143 Q148 E152 N155	
S-1360	HIV-1 integrase	D64 T66 D116 Y143 Q148 E152 N155	
Saquinavir	HIV-1 protease	G84 I84 L90	

Table of various HIV-1 integrase and HIV-1 protease inhibitors

Drug	Type of inhibitor	Affected residues	Structure
Nelfinavir	HIV-1 protease	D30 I84 L90	
Amprenavir	HIV-1 protease	I84 L90 I54 V82	

Table 1. The various HIV-1 protease and HIV-1 integrase inhibitors and their structures. The affected residues in HIV-1 protease and HIV-1 integrase binding pocket are shown as well.

of three amino acids—Y143, Q148, or N155—mutate in conjunction with at least one other mutation [27]. The strongest antiviral resistant mutation seems to be the Q148H integrase mutant ($IC_{50} > 700$ nM), and G140S has been shown to restore the poor replication ability of Q148H to wild-type levels [28]. Even though Raltegravir has seen this resistance profile, pharmaceutical companies still spend lots of money on "me-too" research and the

development on this drug. There should be a distinction made between me-too drugs and second-generation drugs [18]. A second-generation inhibitor needs to exhibit a new mode of action. Secondly, a second-generation drug needs to show significantly improved potency or decreased toxicity. A major problem with second-generation drugs is cross resistance, so these drugs should maintain potency, but avoid this cross-resistance.

2.2.2. Using docking studies to predict the binding mode of S-1360

It is very important to predict a bioactive conformation of a ligand, but the task becomes difficult when the receptor site has a region with unusual conformational flexibility. With the numerous crystal structures available for HIV-1 integrase, there are numerous differences in the active site regions in the core domains of IN. S-1360 was one of the first beta-diketoacid IN inhibitors to enter clinical studies. Dayam and Neamati sought to predict the bioactive (active site bound) conformations of S-1360 [29]. To achieve this, the researchers performed extensive docking studies with three different crystal structures. The study was extended to include 5CITEP and a bis-diketoacid (BDKA).

To predict the binding mode of S-1360, 104 unique conformations within a 20 kcal/mol energy range were generated using catConf module of the Catalyst. All 104 conformations of S-1360 were docked into the active sites A, B, and C (PDB: 1QS4, 1BIS, and 1BL3, respectively). Based on GOLD fitness scores, 10 conformers with highest scores were selected for further analysis. The researchers noted that S-1360 adopted very different binding orientations inside the active sites for A, B, and C. In the A active site, the bound conformation with the highest GOLD fitness score was found 102 times of 200 conformations. S-1360 occupies a space near D64, D116, N120, and Mg^{2+} ion. In active site B, the highly favorable conformation of S-1360 is found 62 times. The triazole and the diketoacid moiety of S-1360 occupy a deep cavity surrounded by I151, N155, V75, and Q62. The groups show favorable van der Waals and electrostatic interactions with D64, I151, E152, and N155 (**Figure 2**). The highly favorable binding conformations of the C active site are found 172 times. The researchers also compared the best binding orientations of S-1360 in the active sites from the three different crystal structures of HIV-1 integrase. Dayam and Neamati observed that S-1360 in the A active site achieves a planar conformation and interacts with various residues throughout the active site. In this orientation, S-1360 forms H-bonding interactions with K159 and N120. The two oxygen
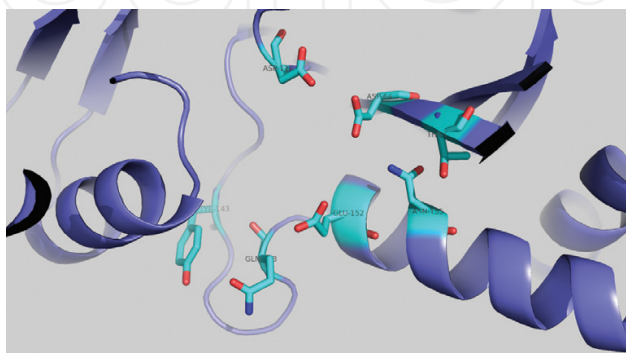

Figure 2. 1BL3 active site with the residues that contribute to antiviral resistance highlighted.

atoms from the furan ring and keto group for coordinate bonds with the Mg^{2+} ions. This conformation appears to be stable because it occurs 102 out of 200 times. The binding site conformation of S-1360 inside the C active site is also very stable because of its 172 appearances out of 200. While these conformations of A and C are stable, this conformation is not in line with experimentally observed results. S-1360 selectively inhibits strand transfer reactions of HIV-1 integrase, but S-1360 in A and C did not interact with amino acids in the strand transfer (ST) cavity. However, S-1360 did form strong interactions with various amino acids and Mg^{2+} ion in the cavity where 3′-processing of IN is believed to be carried out.

Docking in this study was performed using version 1.2 of the Genetic Optimization for Ligand Docking (GOLD) software. This uses a genetic algorithm to explore the ligand conformational flexibility with partial flexibility of the active site [30]. GOLD was tested on a dataset of over 300 complexes. GOLD succeeded in more than 70% of cases in reproducing the experimental bound conformations of the ligand [31]. GOLD requires that users define the specific binding site. For this study, Dayam and Neamati defined a 20 Å radius active site. D64 was selected as the center of the active site. The GOLD program then searches for a cavity within the defined area. The program also considers all the solvent accessible atoms in the defined area as active site atoms. "All docking runs were carried out using standard default settings with a population size of 100, a maximum number of 100,000 operations, a mutation, and crossover rate of 95" [29]. At the end of each run, GOLD reported all the predicted bound conformations based on their fitness score. The fitness score consists of H-bonding, complex energy, and ligand internal energy.

2.2.3. Validating the resistance profiles of "me-too" Raltegravir analogs using docking studies

Serrao et al. sought to validate the resistance profiles of me-too Raltegravir analogs [18]. There are minor variations in the *in vitro* activity of the numerous me-too integrase inhibitors. The researchers believed that the development of me-too compounds could possibly yield a relatively low amount of clinical success due to their similarities [18]. It is still possible for a Raltegravir me-too analog to become a second-generation integrase inhibitor. To elucidate this viewpoint, the researchers utilized the molecular docking program GOLD version 3.2 to conduct a docking study. Serrao et al. used the structure of 1BL3 complexed with an Mg^{2+} ion, and various me-too compounds.

Serrao et al. proposed that residues essential to the compounds' interaction with HIV-1 integrase would be prime candidates for resistance mutation. "Raltegravir makes direct interactions with three residues encompassing the [IN] catalytic motif (D64, D116, E152)" [18]. The researchers wanted to predict the interaction residues of Raltegravir's analogs in a similar way. They wanted to show that the compounds would have little success in viral eradication. Because S-1360 was the one of the first clinical IN inhibitor candidates, the researchers thought it would be interesting to look at the interactions between S-1360 and 1BL3 and compare with that of Raltegravir. The researchers found that there are identical interactions between the two drugs (D64, T66, D116, Y143, Q148, E152, and N155). Raltegravir showed an additional interaction with E92. While this observation has been confirmed by clinical experiments, the E92Q mutation has conferred upwards of a sevenfold viral resistance to Raltegravir [32–34].

The researchers' data could significantly validate the reliability of their docking technique. The researchers then moved on to describing the interactions between HIV-1 integrase and each most potent analog of Raltegravir. On the several compounds that were used in this follow up study, most all of them interacted in the same binding pocket that Raltegravir is active in. If the researchers' predictions are correct, these candidate drugs will fail to replace Raltegravir. The researcher's note, while there is always the possibility for me-too drugs to evolve into block-buster drugs, the studied HIV-1 integrase "drugs appear to have a small chance of improving the clinical outlook of HIV patients with Raltegravir viral strains" [18].

2.3. HIV-1 protease

An essential element in the HIV life cycle is HIV-1 protease. It is a retroviral aspartyl protease. HIV-1 protease is a homodimer, with each subunit made up of 99 amino acids [35]. Gag and Pol polyproteins are cleaved by this protease [36]. When these are cleaved at the appropriate places, a mature and infectious HIV virion is produced. When an effective HIV protease is blocked, the HIV virus is not infectious [37]. HIV's ability to replicate and infect additional cells can be disrupted by mutation of the HIV protease active site or inhibition [38]. For this reason, HIV protease has seen a massive amount of research money in developing HIV-1 protease inhibitors.

HIV-1 protease is a homodimeric enzyme. Two aspartic acid residues that are essential for catalysis [39], D25 and D25, are located on each monomer. Asp-Thr-Gly sequence is present in HIV-1 protease, but this is conserved among other mammalian aspartic protease enzymes. There are extended beta-sheet regions on each monomer, and these are known as "the flap". This makes up the hydrophobic substrate binding cavity with the two aspartyl residues on the bottom. HIV-1 proteases are highly selective, and very catalytically active in hydrolyzing peptide bonds. While the mechanism is similar to many known features of aspartic proteases, the full detailed mechanism of this enzyme has not been fully understood [40].

2.3.1. Saquinavir and Nelfinavir—HIV-1 protease inhibitors and their ensuing resistance

The ideal HIV-1 protease inhibitor should be potent and specific for HIV-1 protease compared to other mammalian aspartic acid proteases [41]. The drugs should also have good bioavailability and duration in human bodies. There were no known inhibitors of HIV-1 protease when it was first determined to be a good target for antiviral therapy. A good starting place to look was the type of enzyme that HIV-1 protease was, an aspartic acid protease.

When researchers were designing HIV-1 protease inhibitors, it was noted that there was a stereocenter in the drug that correlated with the drug's activity. The transition state hydroxyl group needed to be in the R-stereochemistry or else the drug completely lost its activity. This discovery led researchers to identify Ro-31-8959, or Saquinavir, as a prime candidate for further studies because of this characteristic. Saquinavir has an IC_{50} < 0.37 nM for HIV-1 protease and does not inhibit other aspartic acid proteases, making it highly potent. While the drug is potent, it shows poor oral bioavailability—only 4% [41]. Researchers attribute this to the high molecular weight of the drug and the large number of amide bonds. Agouron

Pharmaceuticals and Lilly Research Laboratories collaborated to produce Nelfinavir [42]. The structure of Nelfinavir is very similar to the structure of Saquinavir, but Nelfinavir contains a couple of changes. Labile components in Saquinavir were replaced with a hydroxytoluene amide group, however, this modification resulted in reduced potency. Drug developers replaced the phenyl group with a phenylthio group. This phenylthio group was better able to fill the hydrophobic pocket of the HIV-1 protease active site [43]. With an IC_{50} = 2 nM, Nelfinavir is also a very potent HIV-1 protease inhibitor.

As shown with HIV-1 integrase inhibitors, resistance persists to be a pressing problem in the treatment plans for HIV-1. Because Saquinavir and Nelfinavir have similar structures, there are different, yet highly overlapping sets of amino acids substitution mutations that confer to drug resistance. The mutations that affect the binding site for Saquinavir are G84, I84, or L90. For Nelfinavir, the only difference from the Saquinavir mutation is D30 instead of G48 [41]. While these amino acids affect the binding pocket, there are other overlapping sets of amino acids that when mutated elsewhere in the HIV-1 protease enzyme confers antiviral resistance. These sites include L10, M46, L63, A71, and N88. Because many of the HIV-1 protease inhibitors on the market right now are very similar in structure, it is not surprising that there is a high degree of cross-resistance between the drugs.

2.3.2. Predicting HIV-1 protease resistance with docking studies

There are several different methods to interpret the resistant behavior of HIV-1 from genotypic data. A physics-based approach of docking has seen an influx of use by researchers in evaluating the energy interactions of the protein-inhibitor complexes. This technique has been widely used to look at the interactions between HIV-1 protease and its inhibitors. In 2005, Jenwitheesuk and Samudrala completed a study that used a protein-inhibitor docking approach to determine the correlation between experimentally and computer calculated protease inhibitor binding affinities [44]. The researchers also supplemented their findings with a molecular dynamics protocol [45]. This was used in part because most docking programs utilize a rigid protein protocol. HIV-1 protease has special flaps that are in motion upon binding. Since the structure of target protein is rigid, the opening and closing of the flaps is not performed [46]. This protocol was used to simulate the flexible nature between the ligand and the enzyme. The researchers used the X-ray crystal structures of various wild-type HIV-1 protease-inhibitor complexes. For Saquinavir and Nelfinavir, the researchers selected 1HXB and 1OHR, respectively (**Figures 3** and **4**). The researchers then substituted the wild-type side chains with a mutant side chain.

When preparing the inhibitor structure, the researchers treated them as an all atom entry. By doing so this filled the empty valences with hydrogen. All the rotatable bonds in the inhibitors were also allowed to rotate freely. The researchers used AutoDock version 3.0.5 with a Lamarckian genetic algorithm to carry out docking calculations. Genetic algorithms use the idea of natural genetics and biological evolution. There are specific values describing the ligand with respect to the protein (translation, orientation, and conformation). These are described at state variables and in the genetic algorithm (GA), each state variable corresponds to a gene. In genetic algorithms, the genotype is from the ligand's state, and the phenotype
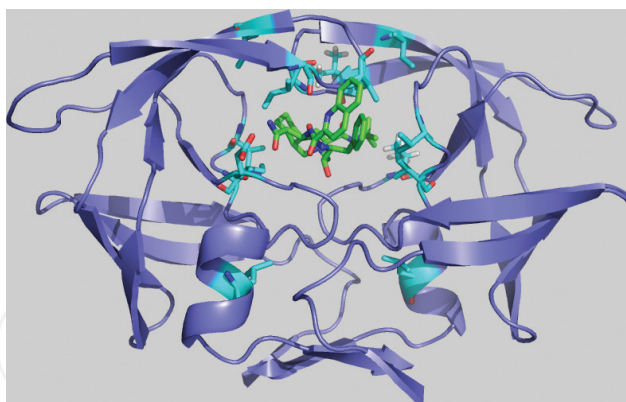
Figure 3. 1HXR mutated with interactions between the binding pocket and Saquinavir.
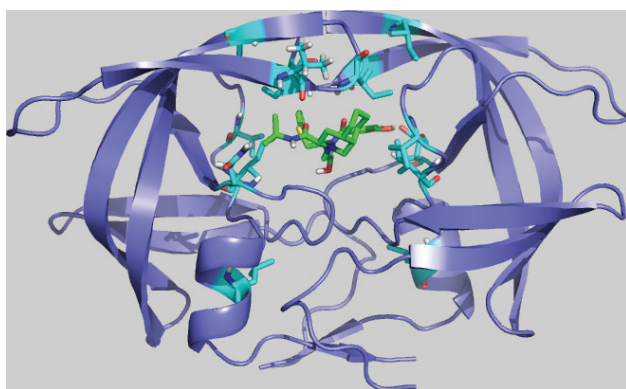

Figure 4. 1OHR mutated with amino acid interactions and Nelfinavir in the binding pocket.

comes from the atomic coordinates [46]. When molecular docking is performed, the fitness of the gene is referred to as the total interaction energy between the ligand and the protein. The GA comes into play by mating random pairs of individuals to induce crossover. In this scenario, some offspring undergo random mutation. The genes are selected from the current generation based off their fitness scores. This process is repeated for multiple generations to produce a ligand and protein interaction that has the most fitness. In the research conducted by Jenwitheesuk and Samudrala [45], there were a total of 27,000 generations. AutoDock generates the energy terms for inter-molecular energy, internal energy of the ligand, and torsional free energy. When the researchers determined the final docked energy of the protein ligand complex, the inter-molecular energy, and the internal energy of the ligand was added.

In the results of this study [45], Jenwitheesuk and Samudrala saw a significant improvement in the correlation coefficient when supplementing their docking procedure with MD simulation to provide a flexible nature of the protein (correlation coefficient changed from 0.38 to 0.87). The researchers were also able to see that their docking with dynamic protocol was 64% accurate for phenotypically resistant profiles and 83% accurate for phenotypically susceptible groups. There was a previous study done by Shenderovich et al. [47]. While this study followed a similar protocol to the one followed by Jenwitheesuk and Samudrala, Shenderovich et al. only used 50 HIV-1 protease sequences. Jenwitheesuk and Samudrala used 1792 HIV-1

protease sequences. This larger sample size could include all of the reported resistant mutations. Jenwitheesuk and Samudrala also added a protein-inhibitor relaxation feature to their protocol. Their protocol was also able to consider the rearrangement of the side chain on the active site surface. The relatively short MD simulation of 0.1 ps had a significant effect on the flap region (which moved away from the binding pocket—RMSD = 0.54 Å), yet was not long enough to affect the main chain of the protein. Using this protocol, the resistance and susceptibility predictions from Nelfinavir and Saquinavir were 86 and 94%, respectively [45].

This study looked at the two key mutations discussed earlier—Asp30Asn and Gly48Val. In this study, docking with the molecular dynamics implementations always failed to identify as a cause of drug resistance. This suggests that researchers should not rely solely on one method or system in making decisions about therapeutic regimens without consulting other methods, resources, and techniques. This study was still able to determine other mutations around the binding pocket. The docking with MD simulation implementation could identify mutations that correspond with high levels of resistance of Amprenavir (another kind of HIV-1 protease inhibitor)—I50V and a combination of I84V + L90 M and I54V + V82A + I84V + L90 M. These mutations are cross resistant with Nelfinavir and Saquinavir.

2.4. Repurposing HIV-1 protease inhibitors

American trypanosomiasis, or Chagas disease, is caused by the protist *Trypanosoma cruzi.* Many times there are no early signs of infection but over the course of the infection symptoms can range from a mild fever, swollen lymph nodes, or headaches. If the infection progresses further, the symptoms can include enlarged ventricles of the heart, which will ultimately lead to heart failure. This infection is most common in Mexico, Central America, and South America, and an estimated 6.6 million people are living with this parasite [48] The most common ways that the disease is spread are eating contaminated food, from mother to her fetus, and blood or organ transfusions [49]. While the knowledge of this parasite has grown remarkably, there have been no medications to treat Chagas disease in the last 40 years [50].

Over the years, there has been a recent interest in drug repurposing (also known as drug repositioning). The process involves using known and approved medications—and sometimes discontinued drugs from other drug trials—and using them for a new clinical applications other than their intended treatment. Drug repurposing is gaining popularity due to the fact that within the past few decades there has been a significant decline in the number of safe and effective drugs being developed for the pharmaceutical market. Pharmaceutical companies are not inclined to fund research or product design because development of a new drug is a long and costly process [51]. One of the major benefits of trying to repurpose drugs is the reduced cost of researching and developing a novel drug from scratch.

Bellera et al. present computer-aided identification of approved drugs Clofazimine, Benidipine, and Saquinavir as potential trypanocidal compounds [50]. The major drug target is cruzipain (Cz). Cz is the major cysteine protease of the parasite. This protease is essential for replication of the intracellular form of the parasite. Bellera et al. compiled a 147 compound dataset. This data set was balanced with 77 Cz inhibitors and 70 non-inhibitors. The researchers then used docking studies on Saquinavir, Benidipine, Clofazimine, and the inactive verapamil. The protein to be

used in the docking studies was 1ME4. This protein was a crystal structure of one reversible inhibitor that was complexed with Cz. The compounds were docked according to the Lamarckian genetic algorithm. The active site was defined as a $19 \times 15 \times 15$ Å^3 grid. The researchers performed 100 docking runs for each compound. The docking active site was treated as a rigid molecule and the ligands were treated as flexible. The researchers used Autodock 4.2 to analyze the results of their docking study. The binding results from the docking studies correlated with experimental evidence. The scores for Saquinavir, Benidipine, and Coldazimine were −12.76, −8.42, and −7.36 kcal/mol, respectively. However, the inactive verapamil compound was only −6.37 kcal/mol [50].

3. Identification of medicine for cancer

Cancer is one of the most devastating and destructive diseases that is known to be a persistent public health threat. As of the year 2016, cancer is the second leading cause of death in the United States. There were an estimated 1,685,210 new cases and 595,690 deaths resulting from cancer [52]. Along with the high rate of incidence exacerbating the pressure already felt by researchers to discover a cure, the mechanisms of the disease add another level of complexity that must be outmaneuvered. Many cancer cells lack molecular targets making it extremely difficult for anticancer chemotherapeutics to be fully effective. Toxicity against normal tissues can develop from anticancer therapy, which leads to unwanted side effects. Due to the adverse effects, many anticancer chemotherapeutics are given at suboptimal doses which typically results in failure of therapy, drug resistance, and metastatic disease [53]. The complications associated with cancer demonstrate the critical need for the development of new anticancer therapies that are successful with minimal undesired reactions. In order to aid in the task, many researchers are turning to *in silico* methods to expedite the process. Molecular docking is one of the most popular and reliable softwares available for drug discovery, design, and repurposing. Many researchers utilize molecular docking in cancer research because it provides great insight into protein-ligand interactions, ligand binding mechanisms, and knowledge of the optimal orientation of the ligand bound to its target to form the most stable complex. Molecular docking is an essential computational method that has demonstrated a promising future for the evolution of more effective and potent anticancer therapies.

3.1. Docking for identifying novel proteasome inhibitors and understanding the binding mechanisms

A variety of cancer therapeutics already exists and is available to patients; many of these therapies attempt to have a specific molecular target in order to eradicate the cancerous cells. One protein that receives extensive attention due to its pivotal biological role in eukaryotic cells is the proteasome. There are two major types of proteasomes such as the 20S proteasome, which is responsible for intracellular protein degradation and the 26S proteasome complex, which functions in the ubiquitin pathway as an ATP-dependent proteasome [54]. The 26S proteasome has three proteolytic activities including peptidyl glutamyl peptide hydrolase

(PGPH) in the β1subunit, trypsin-like (T-L) in the β2 subunit and CT-L activities in the β5 subunit [55].

Degradation of proteins in the cytoplasm and nucleus of eukaryotic cells can affect: regulation of cellular pathways particularly cell growth and proliferation, apoptosis, DNA repair, transcription, immune responses, and signaling processes [56]. Inhibition of proteasomes has therefore become an attractive target for anticancer therapies. The drug Bortezomib was developed by Millennium Pharmaceuticals Inc. and received regular approval by the Food and Drug Administration in 2005 as the first proteasome inhibitor to be used for the treatment of multiple myeloma [57]. Bortezomib is a peptide boronate inhibitor of the proteasome and it selectively binds to the protein to inhibit its chymotryptic-like activity [58]. The anticancer effects demonstrated by Bortezomib are mainly observed by the inhibition of the transcription factor NF*k*B and the promotion of apoptosis in rapidly dividing cells. While Bortezomib is considered a successful cancer treatment, many reports of adverse side effects have driven researchers to develop a more potent and selective proteasome inhibitor [59].

In the race to discover a more efficacious proteasome inhibitor, molecular docking has been an extremely beneficial tool utilized by researchers to expedite the exacting process. *In silico* high throughput screening of multiple chemical libraries identified the compound PI-083 as a potential inhibitor due to its potency (IC$_{50}$ = 1 μM). Molecular docking of PI-083 to the 20S proteasome was performed by the GLIDE computer program, version 3.0 (Schrödinger, LLC, New York, NY). The GLIDE program used for the docking and grid generation was set using default options and parameters. The X-ray structure of yeast 20S proteasome complexed with Bortezomib revealed that the pyrazine ring in the Bortezomib forms a hydrogen bond with Asp114 from the β6 subunit of the proteasome. As visualized in **Figure 5**, Bortezomib also forms hydrogen bonds with T21, T1, G47, and A49 residues located in the active site. PI-083 possesses a pyridine ring and it was docked to a model derived from the Bortezomib-proteasome complex (PBD ID: 2F16). The docking studies revealed that PI-083 and Bortezomib have similar binding mechanisms to the active site of the CT-L enzyme within the proteasome [55]. The molecular docking combined with *in vivo* studies of PI-083 are indicative that the compound is successful in tumor suppression which insinuates a need for further clinical research in regards to PI-083 as an anticancer therapy.
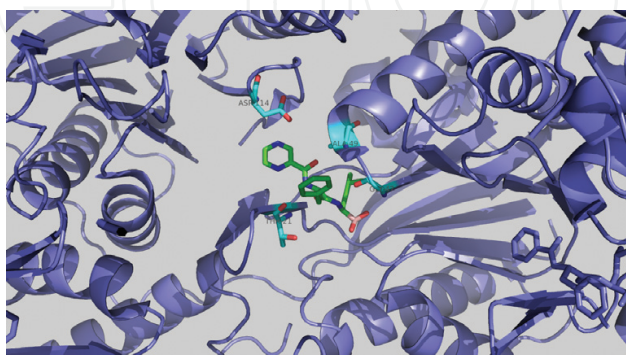

Figure 5. The Bortezomib ligand positioned in the active site of the yeast 20S proteasome crystal structure. The key residues T1, T21, G47, A49, and D114 in the active site are shown.

Molecular docking has not only been successful in identifying potential proteasome inhibitors but it has also been beneficial in understanding the binding mechanism of proteasome inhibitors to the proteasome. One study conducted by Zhang et al. was focused on MG132 (Z-Leu-Leu-Leu-al), which is a structural component of peptide aldehydes selective and potent against the proteasome. Using the Insight II software, the proteins and ligands were prepared for docking. MG132 was then covalently docked to the β5 subunit of the 20S proteasome using GOLD version 4.0. The results showed that the docking of MG132 proposed two binding modes with low docking energies. More thorough analysis and the use of molecular dynamics simulations revealed that binding mode I was more stable than mode II. The computational methods utilized in this study resulted in the generation of a model that was able to re-examine the correlation of the structure and activity of proteasome inhibitors, specifically the interactions that take place at the P2 and P4 sites [60]. Observing the binding mode is advantageous for the improvement of existing proteasome inhibitors but also for the development of more potent inhibitors.

Ma et al. used the binding mechanism of MG132 as a comparison for docking their own series of peptide aldehyde derivatives in which they synthesized. A total of 17 different peptide aldehydes were developed and are listed in **Table 2**. Eight of the peptides are in the Cbz class at the R4 position and the other nine peptides are in the Boc class at the R4 position.

Compounds	R4 position	P3 position	P2 position
1	Cbz	Asp(O^{t}Bu)	Phe
2	Cbz	Asp(O^{t}Bu)	Leu
3	Cbz	Glu(O^{t}Bu)	Phe
4	Cbz	Glu(O^{t}Bu)	Leu
5	Cbz	Phe	Leu
6	Cbz	Arg(NO$_2$)	Leu
7	Cbz	Arg(Tos)	Leu
8	Cbz	Nap[a]	Leu
9	Boc	Asp(OBzl)	Phe
10	Boc	Asp(OBzl)	Leu
11	Boc	Glu(OBzl)	Phe
12	Boc	Glu(OBzl)	Leu
13	Boc	Pro	Phe
14	Boc	Pro	Leu
15	Boc	Ser(OBzl)	Leu
16	Boc	Thr(OBzl)	Leu
17	Boc	Tyr(OBzl)	Leu

Table 2. Peptide aldehyde derivatives for the inhibition of 20S proteasome activity.

The 17 peptide aldehydes were then docked using GOLD software 4.0 with the $\beta5$ of the 20S proteasome based on the crystal structure of the first known inhibitor MG101 complexed with the 20S proteasome. The results of the docking experiment indicated that the size and length of the P3 side chain is critical to the activity of the peptide aldehyde. Compounds 3 and 4 which are part of the Cbz series synthesized by Ma et al. possess Glu(O^{t}Bu) residues at the P3 site providing the most active inhibition. The results from docking indicated that when a phenyl ester was used to replace a tert-butyl ester at P3 in the Boc-series, the Asp(OBzl) residue in compound 10 exhibited more active inhibition than Glu(OBzl) residue in compound 12. Also in the Boc-series, Ser(OBzl) in compound 15 has the most suitable length side chain because it demonstrated the most active inhibition to CT-L active site [61]. The docking results generated from this study highlighted the importance of the P3-position substitutes are vital for inhibitor potency, which is essential for designing more effective proteasome inhibitors.

Peptide aldehydes are not the only compounds being considered as proteasome inhibitors for cancer therapeutics. Santoro et al. investigated whether or not cationic and anionic porphyrins can be used as inhibitors of the proteasome. Porphyrins are hydrophilic compounds that possess tumor localizing properties and are used in conjunction with red light for photodynamic therapy for the treatment of tumorous cells [62]. Cationic and anionic porphyrins were docked using AutoDock Vina to the 20S proteasome complexed with Bortezomib (PDB: 2F16). The cationic porphyrin H_2T4 demonstrated similar inhibitory activity in all three catalytic sites of the proteasome when observed during *in vivo* studies. Docking of planar H_2T4 with the 20S proteasome revealed the binding mechanism of the porphyrin to the proteasome. The results from the docking studies reconcile with the results of the inhibition studies, indicative that H_2T4 has the potential to be a proteasome inhibitor. Along with the active ability of the porphyrin to inhibit the proteasome, the molecules also possess low toxicity, making them an attractive class of compounds to continue to evaluate as a form of anticancer therapy [63].

3.2. Docking for identifying inhibitors of CAs

Besides proteasomes, several isoforms of carbonic anhydrases (CAs) have become an attractive anticancer drug target. Carbonic anhydrases are ubiquitous metalloenzymes broken up into four unrelated gene families; the α-CAs, β-CAs, γ-CAs, and δ-CAs. Mammals have 16 α-CAs isozymes that are different in their tissue distribution, catalytic activity, and subcellular localization [64]. The α-CAs are of particular interest because they have well established catalytic and inhibition mechanisms [65]. One α-CA in particular, CA IX, has potential to act as an anticancer drug target as it has the ability to act as a biological marker for certain tumors [66]. CA IX is an extracellular transmembrane-bound protein located in the gastrointestinal tract. When the enzyme is present in hypoxic conditions, CA IX is overexpressed and is observed to be associated with different types of cancer cells via the hypoxia inducible factor-1 (HIF-1). Overexpression also causes the environmental pH of a tumor to be lowered to acidic conditions [66]. The appeal of the CA IX as a potential anticancer drug is demonstrated by the fact that the enzyme has restricted expression in normal tissues (**Table 2**).

Amresh et al. used molecular docking and several others *in silico* methods to discover five potential CA IX inhibitors. AutoDock 4.2 was used to dock all the inhibitors to the crystal
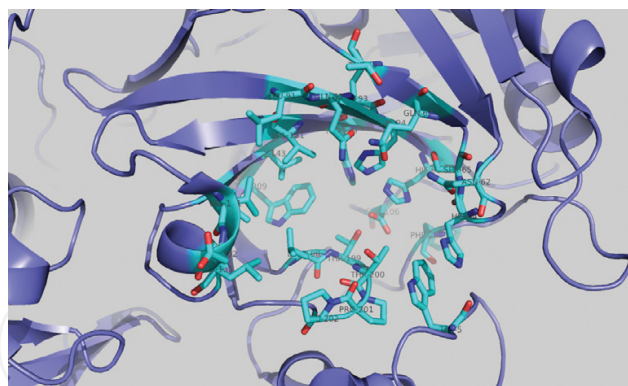
Figure 6. The critical residues in the active site located on the CA IX (PDB: 3IAI): L91, L93, L198, V121, L135, L141, V143, P201, P202, W5, W209, F245, H96, H119, E106, T199, T200, H94, D132, Q92, N62, H64, S65, Q67, T69, and Q92 are displayed.

structure of CA IX (PDB: 3IAI) visualized in **Figure 6**. Coulombic electrostatic potential, van der Waals interaction represented as a Lennard-Jones12-6 dispersion/repulsion term and hydrogen bonding were addressed when evaluating the binding energy during docking. Docking orientations within 2.0 Å in root-mean square deviation tolerance were the parameters set in order to obtain the most favorable free energy of binding. The inhibitors with the best docking poses and scores were then subjected to post-docking energy minimization on Discovery Studio 3.5. The final structures were analyzed using PyMOL visualization programs and the receptor-inhibitor complexes were used to develop the pharmacophore model for further evaluation [67]. Docking simulations were also performed in order to identify the residues present in the active site of CA IX that interact with the inhibitors. The docking study revealed that residues: L91, L93, L198, V121, L135, L141, V143, P201, P202, W5, W209, F245, H96, H119, E106, T199, T200, H94, D132, Q92, and V131 formed either hydrophobic or aromatic interactions with the inhibitor. N62, H64, S65, Q67, T69, and Q92 were identified as the hydrophilic residues in the active site as well [67]. The results of the docking studies established 10 novel compounds as CA IX inhibitors. Further analysis of the docking scores narrowed the list even further to the top five scoring compounds which were: ZINC03363328, ZINC08828920, ZINC12941947, ZINC03622539, and ZINC16650541 [67]. The information obtained from this study has demonstrated the value of molecular docking in identifying new CA IX inhibitors that provide a promising future as an anticancer therapy.

3.3. Docking for identifying inhibitors of EGFR

The epidermal growth factor receptor (EGFR) is another enticing biological target in the development of anticancer therapeutics. The EGFR is a family of tyrosine kinases that regulate many developmental, metabolic, and physiological processes. Binding of the epidermal growth factor to the family of kinases leads to homodimerization or heterodimerization of the EGFR. Mutations of EGFR gene, over expressed copies of the gene and EGFR protein overexpression lead to dysregulated TK activity which is observed in many tumors [68]. Overexpression of EGFR is frequently observed in breast, lung, ovarian, and prostate cancer

and is associated with aggressive tumor behavior [69]. The EGFR is the main activator in the downstream pathways for survival and growth signals such as p42/44 MAPK and PI3K/AKT pathways [70]. Inhibition of these pathways leads to apoptosis of cancer cells, making the EGFR a particularly promising area of cancer research.

The mutations G719S, L858R, T790M, G719S/T790M, and T790M/L858R are commonly seen in patients with cancer because they modify the EGFR kinase activity [71]. García-Godoy et al. used molecular docking in order to study the interactions of EGFR inhibitors on the wild-type EGFR and mutant EGFR. For the wild-type human EGFR, the EGFR (PDB: 4ZAU) was complexed with the ligand AZD9291. Docking was also conducted on the EGFR containing the G719S mutation and the L858R mutation. The EGFR (PBD ID: 2ITN) was used with the G719S mutation and the EGFR (PDB: 2 EB3) was used with the L858R mutation. Both EGFRs were in complex with AMP-PNP. Results of this docking study indicated that in both complexes, M793 was an important residue in facilitating interactions between the ligand and the active site [71]. In the final docking study, docking was performed on the EGFR double mutants T790M/L858R and T790M/G719S. In the instance where the EGFR mutant T790M/L858R was docked, the EGFR (PDB: 4JR5) was used and it was complexed with the ligand 3QY. The double mutant EGFR T790M/G719S (PDB: 3UG2) was also used and it was complexed with getfitinib (PDB: IRE). In both of the docking studies, the results revealed that there is a critical interaction between the ligand and the Met793 residue in the active site of the mutant EGFR [71]. Analysis of the results concluded that the interactions displayed in each case can be crucial evidence to why different cancer patients are more or less sensitive to certain treatments. This provides insight into how certain therapies should be considered circumstantial based on the mutation a patient may possess. The *in silico* methodology utilized in this study set a precedent for other researchers to use molecular docking to discover more drugs for EGFR inhibition.

Mahajan et al. discerned the value of the EGFR as a target for anticancer therapy; using molecular docking they were able to discover potential EGFR inhibitors. Screening of 50,000 compounds was performed by LigPrep (version3.3; Schrodinger, LLC, 2015) in order to prepare a library of drugs to be tested by several *in silico* methods. After the library was prepared with LigPrep, the compounds were then screened against EGFR drug target using e-Pharmacophore, docking, pharmacophore, substructure, and similarity search [72]. The protein used in the docking studies is complexed with the inhibitor tak-285 and it was chosen for the study because it has the best X-ray resolution (1.50 Å) of the EGFR structure (PDB: 3POZ). The downloaded protein was prepared for docking using the Protein Preparation Wizard. Docking the compounds was performed by the Glide module (version3.6; Schrodinger, LLC, 2015) software and the first round of docking studies used the high throughput virtual screening setting. After all compounds had been screened, the top 30% of the best scoring compounds were then re-docked using standard precision (SP) docking. Once those compounds had been docked, the top 30% of the best scoring compounds in SP docking were then re-docked using extra precision (XP) docking. A total of 1534 had been selected as compounds that bound to the EGFR with a respectable docking score [72]. Docking, along with e-Pharmacophore and pharmacophore *in silico* methods were able to narrow 50,000 compounds down to 200 compounds that showed potential for EGFR inhibition. Further computational methodology of the compounds revealed
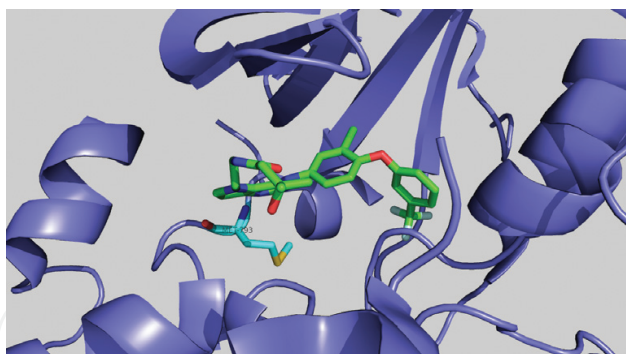
Figure 7. The tak-285 inhibitor complexed with the EGFR tyrosine kinase domain. The critical residue Met793 is shown in the active site of the EGFR tyrosine kinase.

that 87 out of the 200 compounds form an H-bond with M793, a critical residue in the inhibition of EGFR which can be visualized in **Figure 7**. Docking also revealed the structural similarity between the compounds and how the compounds orient themselves in the active site [72]. The 87 compounds were then categorized into 12 structural moieties which provided critical structural modification suggestions that would be beneficial in the development of more potent EGFR inhibitors [72].

3.4. Repurposing approved drugs to anticancer applications

Molecular docking for drug repurposing is another effective and beneficial method that many researchers utilize in order to discover new indications for already existing drugs. The technique is especially favorable when assessing different pharmaceuticals as potential anticancer therapies. Avastin, which was originally developed for metastatic colon cancer and non-small cell lung cancer, has now been approved for metastatic breast cancer. Rituxan, which was intended for non-Hodgkin's Lymphoma has been repurposed for chronic lymphocytic leukemia and rheumatoid arthritis [51]. Molecular docking to make predictions of the physical interactions between the ligand and the target has been a successful practice in drug repurposing.

Avastin and Rituxan are not the only two drugs that have been repurposed for anticancer therapeutics. Oliva et al. used molecular docking to aid in the study of repurposing the FDA approved psychotropic drug Chlorpromazine. Evidence had shown that Chlorpromazine had antiproliferative activity against colon and brain tumors [73]. The drug accomplished this by inhibiting cytochrome c oxidase (CcO), which is the terminal electron acceptor enzyme of the mitochondrial respiratory chain and is composed of 13 subunits [74, 75] . Cytochrome c oxidase subunit 4 isoform 1 (COX4-1) was the focus of the study because in patients with glioblastoma, increased expression of COX4-1 has been associated with Temozolomide chemoresistance [73]. *In vitro* studies indicated that Chlorpromazine inhibited CcO when COX4-1 is expressed, however the binding mechanism was not well understood. Using Schrödinger Suite 2015 (Schrödinger, LLC, New York, NY, 2015), two human CcO homology models were constructed based on the mouse CcO crystal structure (PDB: 2Y69) using the Prime program. The Chlorpromazine ligand was prepared using the LigPrep program and the docking studies were

conducted using the Glide program. The docking results showed that Chlorpromazine binds to a hydrophobic pocket formed by residues from COX4 and transmembrane helices of COX1. L129, K122, M119, and Y126 were identified as being important residues that displayed interactions with the Chlorpromazine [73]. The results also indicated that the Chlorpromazine overlaps with residues of COX11, preventing the subunit from interacting with the rest of the CcO complex [73]. The study provides critical evidence on the repurposing for Chlorpromazine as a treatment for chemoresistant gliomas and persuades future research on Chlorpromazine as an anticancer therapy.

In silico methods have been played an essential role in the battle against many of world's most devastating diseases. Cancer is debilitating, painful, and in some cases lethal; there is a massive urgency for researchers to find a cure so patients no longer have to suffer. Molecular docking has been on the forefront for the development, design, and discovery for new anticancer therapeutics. Among other things, one of the most important features of molecular docking is that it provides researchers with the opportunity to examine specific interactions between the ligand and the molecular target that are not well understood by *in vivo* and *in vitro* methods. Detailed knowledge of the binding interactions and mechanisms of the ligand to the target is critical for the production of new drugs or the improvement of the already existing drugs. Molecular docking is a dependable, economic, and an expeditious process that is of paramount importance in the advancement of anticancer therapeutics.

4. Identification of medicine for other prevalent diseases

4.1. Influenza

Influenza, commonly referred to as the flu, is a viral infection that can be mild or severe, depending on the strain, and the host it infects. Due to the rapidly mutating nature of the influenza virus, new vaccines must be made and administered annually. Each year, researchers must determine which strains of the influenza virus are most likely to become prevalent in the coming flu season; annual flu vaccines are manufactured based on those recommendations [76]. Unfortunately, there is always the threat that the virus may mutate after that decision has been made, rendering vaccines ineffective. In that case, flu outbreaks and even pandemics may occur. In a pandemic, vaccination will no longer be a feasible option, and antiviral agents will become a critical resource [77].

There are two types of antiviral drugs that have been used to treat influenza. The first marketed influenza antivirals were Adamantanes, specifically Amantadine and Rimantadine (**Figure 8A**, **B**). Adamantanes function by blocking the M2 proton channel [78]. This class of drugs was effective against influenza type A, but drug resistance developed rapidly [79, 80]. Hayden et al. conducted a study in which 17 Rimantadine-resistant influenza strains were recovered from 13 patients [81]. The M2 coding sequences of 17 resistant strains were then compared to 8 drug sensitive strains, and it was determined that all resistant strains had a nonsynonymous substitution in RNA segment 7. The most common mutation was S31N,
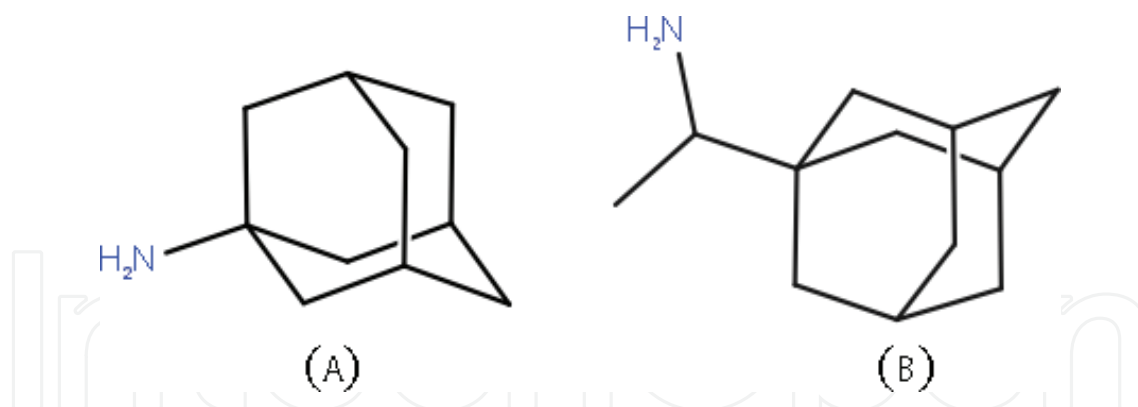
Figure 8. Two dimensional structures of the Adamantanes, (A) amantadine [SMILES: NC13CC2CC(CC(C1)C2)C3] and (B) Rimantadine [SMILES: NC(C)C13CC2CC(CC(C1)C2)C3].

which was found in 14 separate isolates. The other mutations found were A30V, A30T, and V27A. By 2009, all strains of influenza A had become resistant to Adamantanes [82].

The second class of influenza drugs is neuraminidase inhibitors. Neuraminidase, also referred to as sialidase, is an enzyme involved in the release of viral progeny. At the end of the viral replication cycle, neuraminidase cleaves O-sialic acid, also called NeuAc5 (N-acetyl-alpha-neuraminate), during the budding process that releases viral progeny that then infect other cells. Because inhibition of this enzyme greatly reduces the spread of the virus throughout the body, it is an attractive drug target [83]. There are currently two neuraminidase inhibitors on the market: Zanamivir (Relenza) and Oseltamivir (Tamiflu). Zanamivir (4-guanidino-Neu5-Ac2en) was created using computer-assisted rational design based on the X-ray diffraction structure of influenza neuraminidase, which was first solved by Varghese et al. (now PDB: 7NN9) [84]. In further studies, Colman et al. characterized the active site of this protein, identifying a large pocket containing "an unusually large number of charged residues," including R119 and E120[1] [86]. Von Itzstein et al. used GRID software to analyze the active site of influenza neuraminidase and its interactions with various novel inhibitors [87]. The inhibitor with the most energetically favorable interactions was 4-guanidino-Neu5Ac2en, now known as Zanamivir. It was noted that one of the terminal amino groups of Zanamivir's guanidyl group interacted with the glutamic acid 119 carboxyl group (**Figure 9A**, **B**). Von Itzstein et al. went on to conduct Zanamivir trials on influenza infected ferrets and mice, which validated the results of their computational studies [87]. Hayden et al. conducted randomized double blind trials that concluded Zanamivir was both effective and safe for use to treat influenza A and B [88]. The drug became FDA approved in 1999 and has since been used in conjunction with annual vaccines to prevent and minimize influenza outbreaks [89].

4.2. Malaria

Malaria is an infectious disease caused by a parasitic protist and spread by mosquitoes. There are several different species of this parasite; the most deadly, and most prevalent is *Plasmodium*

---

[1]Colman et al. (1983) refers to Arg 119 and Glu 120 as Arg 118 and Glu 119. This text uses the more up to date numbering used in [85].
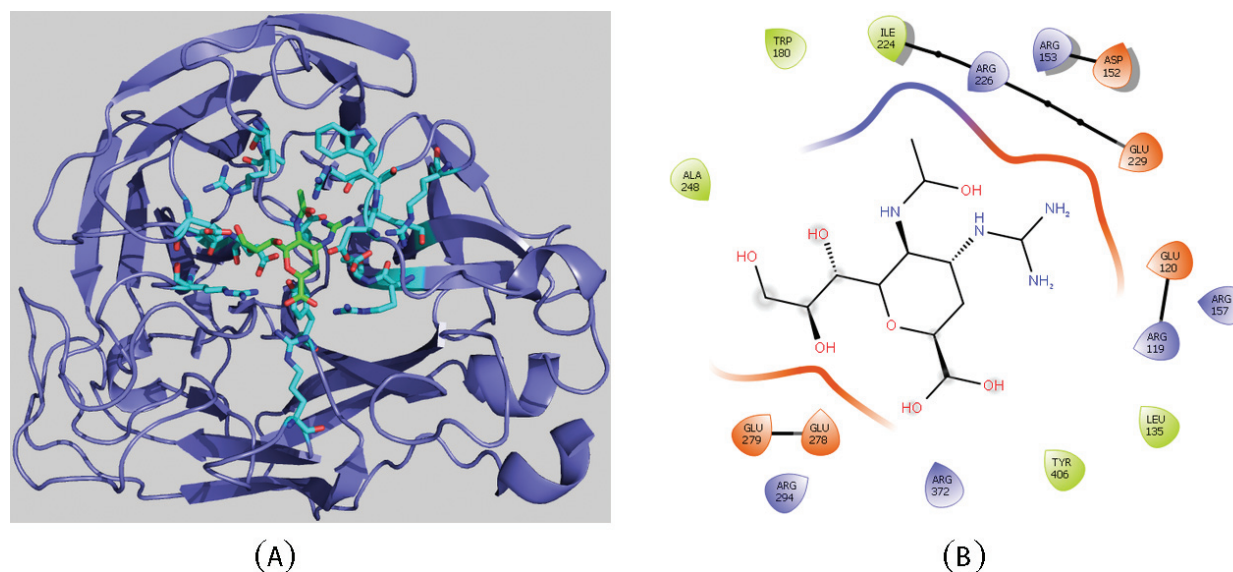
Figure 9. (A) (PDB ID: 5 L17) this structure shows Zanamivir bound to influenza a neuraminidase. Zanamivir, shown with green carbons, interacts with R119, E120, L135, D152, R153, W180, I224, R226, E229, E278, E279, R294, R372, and Y406 (cyan carbons). (B) Ligand interaction diagram showing a closer look at how these residues interact with the ligand. Note the interaction between amino groups and acidic residues (primarily glutamic acid) and the interactions between hydroxyl groups and basic residues (primarily arginine).

falciparum. This disease can cause flu-like symptoms, and can be fatal if left untreated [90]. Malaria is typically treated with quinine drugs such as Chloroquine, Hydroxychloroquine, or Amodiaquine (**Figure 10A**, **B**), which function by interfering with heme polymerization [91]. Interference with this function leads to increased levels of hemoglobin and ferriprotoporphyrin IX (FPIX), which can be toxic to the parasite. *P. falciparum* has developed resistance to chloroquine (and similar drugs); in resistant cells, quinine drugs are actively transported out of the parasitic vacuole [92]. This form of resistance has become widespread, resulting in a need for new antimalarial drugs.

In *P.* falciparum, M18 aspartyl aminopeptidase (*Pf*M18AAP) and its interactions with membrane proteins are essential for parasite survival, making it an attractive antimalarial drug target. Using molecular docking and other computational methods, Kumari et al. determined structural requirements for *Pf*M18AAP inhibitors using GOLD v5.2 and the Schrödinger Maestro 9.1 GLIDE program [12]. This study selected and screened just under 30,000 compounds
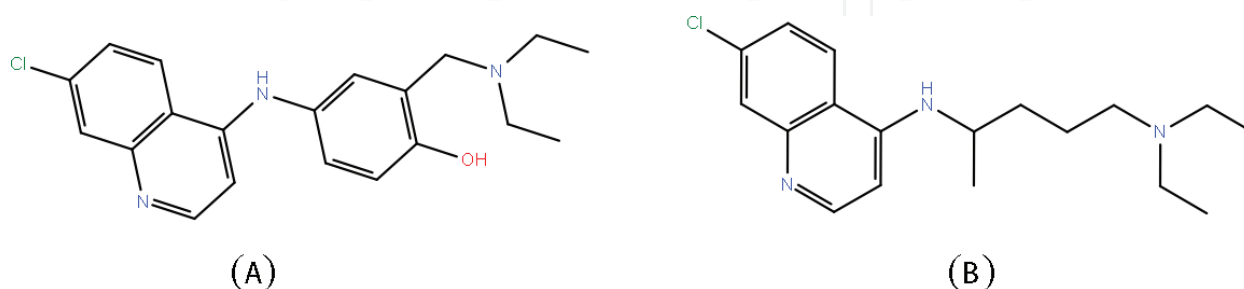

Figure 10. Two-dimensional structures of the common quinine drugs, (A) Amodiaquine [SMILES: Clc1cc2nccc(c2cc1) Nc3cc(c(O)cc3)CN(CC)CC] and (B) Chloroquine [SMILES: Clc1cc2nccc(c2cc1)NC(C)CCCN(CC)CC].

for binding activity. Based on the results, it was concluded that the best inhibitors had one hydrogen donor, one hydrophobic group, and two aromatic rings. Molecular docking and pharmacophore modeling have been used to search for novel inhibitors using those criteria.

The lactate dehydrogenase enzyme of *P. falciparum* (*Pf*LDH) is a target of quinine drugs, and is another potential target for novel antimalarial drugs. This enzyme is important for glycolysis, and its inhibition can potentially result in death of the parasite [93]. Compounds similar to nicotinamide adenine dinucleotide (NADH) are believed to be excellent candidates for *Pf*LDH inhibition [94]. Penna-Coutinho et al. used molecular docking (with software MolDock) to select potential drug candidates [95]. NADH and 50 potential drug candidates were docked to *Pf*LDH in complex with Oxamate (PDB: 1LDG), the substrate that NADH binds to (**Figure 11**); the compounds that had the most similar docking score to NADH were selected for *in vitro* tests. The *in vitro* tests confirmed the activity of the highest scoring compounds, Itraconazole, Atorvastatin, and Posaconazole. In further tests, these same compounds inhibited parasite growth in mice infected with *Plasmodium berghei*, another species of the malaria parasite. These compounds require further testing, but could potentially progress to clinical trials and eventually be marketed as antimalarial drugs.

4.3. Zika

The Zika virus (ZIKV), named for the Ugandan forest in which it was originally found, was first isolated in monkeys [96]. ZIKV belongs to a genus of viruses known as flaviviruses; other viruses belonging to this genus are dengue fever, yellow fever, hepatitis, and West Nile. ZIKV can be transmitted by mosquitoes or sexual contact. Symptoms of the virus include fever, joint pain, and rash for up to 7 days. ZIKV has also been associated with Guillain-Barre syndrome [97], an autoimmune disease. The virus can also be transmitted from mother to fetus, which can result in severe birth defects. From 2007 to 2014, several small outbreaks of the virus were
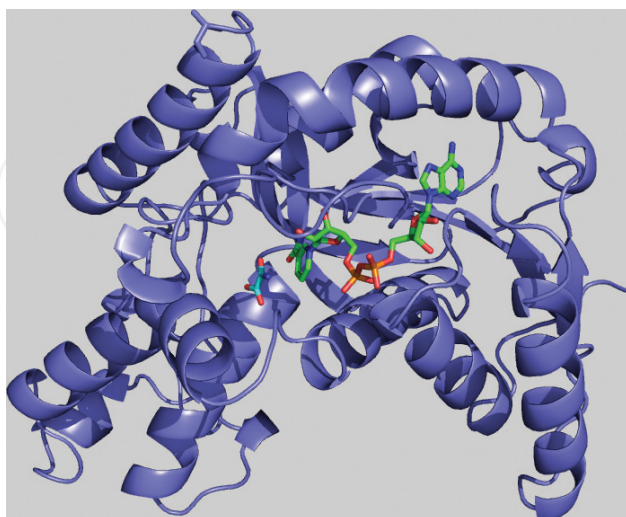

Figure 11. Structure of plasmodium falciparum lactate dehydrogenase in complex with Oxamate and NADH. NADH [SMILES: O=C(N)c1ccc[n+](c1)[C@@H]2O[C@@H]([C@@H](O)[C@H]2O)COP([O-])(=O)OP(=O)(O)OC[C@H]5O[C@@H] (n4cnc3c(ncnc34)N)[C@H](O)[C@@H]5O]; Oxamate [SMILES: C(=O)(C(=O)O)N].

reported [97–99]. In 2015, the first ZIKV epidemic began in Brazil. As outbreaks become more and more severe, it is becoming increasingly urgent to find a drug to treat ZIKV.

Non-structural protein 5 methyl transferase (NS5 MTase) is crucial for the maintained stability of a flaviviral genome, and the ability to evade immune response [100] which makes it an attractive target for antiviral activity. Zhang et al. used docking simulations (AutoDock 4.2) to determine potential designs for novel NS5 MTase inhibitors and binding sites [101]; the authors of this study found that dengue virus inhibitor compound 10 found by Lim et al. [102] (PDB: 3P8Z) may bind to ZIKV NS5MTase. Ramharack and Soliman utilized several different computational tools in their study. Preliminary methods included homology modeling, binding site prediction, and pharmacophore modeling [103]. To narrow down the results from these studies, they used molecular docking [AutoDock Vina]. Out of 31 compounds subjected to docking studies, 3 were chosen for the next step, molecular dynamic simulation. It was concluded that two of their compounds showed "substantial stability in complex with the target enzyme (ZIKV NS5)," [103].

Hepatitis C is another virus that is closely related to Zika. Hepatitis C is commonly treated with polymerase inhibitors (Ribavirin and Sofosbuvir). Sacramento et al. used docking simulations (MODELER 9.16) to model binding between Hepatitis C polymerase inhibitors and Zika RNA polymerase (PDB: 4WTG) [104]. These simulations, as well as their *in vitro* trials suggested that these drugs intended for treatment of Hepatitis C may be effective against Zika as well. A study by Elfiky supported these results through further docking simulations (SCIGRESS software with PDB: 2J7U) [105].

4.4. Tuberculosis

Tuberculosis (TB) and infectious disease caused by *Mycobacterium tuberculosis*. Human infection with TB dates back all the way to ancient Egypt, India, and China [106]. TB is spread through the air, usually by a cough or sneeze from an infected person. TB kills nearly 2 million people each year, mostly in Africa [107]. The most effective treatments for non-resistant TB are Isoniazid and Rifampin. Unfortunately, TB drug resistance has become extensive [107]. There are three categories of resistant TB strains: multidrug resistant (MDR), extensively drug-resistant (XDR), and totally drug-resistant (TDR). In order to be classified as MDR TB, the strain must be resistant to both Isoniazid and Rifampin [108]. A TB strain is classified as XDR if it is resistant to Isoniazid, Rifampin, and "is also resistant to three or more of the six classes of second line TB drugs," [108]. TDR strains are resistant to all known TB drugs [109]. Dramatic increases of drug resistance have prompted researchers to seek new drug targets; in order to reduce research costs and get results as quickly as possible, many are turning to docking simulations for preliminary trials.

Shikimate kinase is a protein involved in an amino acid biosynthesis pathway in *M. tuberculosis* [110]. Interruption of this pathway prevents synthesis of essential amino acids, leading to incomplete proteins, which leads to cell death. Vianna and de Azevedo used docking simulations (MOLDOCK) to identify novel SK inhibitors; these compounds were compared to staurosporine, which has demonstrated SK inhibition *in vitro* [111]. The novel inhibitors were

docked to a number of structures for *Mt*SK (PDB: 2DFN, 1U8A, 1WE2, 1ZYU, 2G1K, 2IYQ, 2IYR, 2IYS, 2IYX, 2IYY, 2IYZ, and 3BAF).

Another response to drug resistance is drug repurposing. The advantage of drug repurposing is that potential drugs have already been shown not to have severe side effects, which speeds up the process and saves money. Studies of this nature often utilize molecular docking and other computational methods to save even more time and money by screening more potential drugs in a shorter time frame. Kahlous et al. selected 1991 FDA-approved (nonantibiotic) drugs and tested them for antibiotic activity against *Staphylococcus aureus* [PDB: 2XCS and 2XCT] by docking (OpenEye HYBRID) drug structures to known antibiotic targets [112]. These results were then compared to a variety of market antibiotics. The drug candidates were narrowed down to 34 potential candidates for further testing. Among the top candidates were Diclofenac (antiinflammatory), Drotaverine (antispasmodic), Flurbiprofen (antiinflammatory), Ibuprofen (antiinflammatory), and Niacin (vitamin B_3). Brindha et al. conducted a similar study, specifically targeting tuberculosis [113]. This study screened 1554 FDA-approved drugs (Schrödinger GLIDE) for their ability to bind to protein kinase B of *M. tuberculosis* (PDB: 2FUM), a known antibiotic target. Fourteen of these drugs were determined suitable for further exploration as TB drugs. The top three candidates from this study were Flavin adenine dinucleotide (treats vitamin B_2 deficiency), Valrubicin (treats bladder cancer), and Arcarbose (treat/manage type II diabetes).

5. Summary and discussion

While *in silico* methods must be reaffirmed by *in vitro* and *in vivo* testing, computational methods have been gaining popularity in the drug design industry by proving they are critical in the discovery of medications. Several drugs that are currently available to the public for the treatment of different diseases have been developed based on *in silico* approaches. For example, Zanamivir, used to treat influenza, was developed using computer-assisted design [84]. Through these studies, Zanamivir was identified as the inhibitor having the most energetically favorable interactions with influenza neuraminidase. The results from the docking study of Zanamivir were convincing enough to move forward with *in vivo* testing; the results of *in vivo* studies reaffirmed the results of the *in silico* tests [87]. Nelfinavir and Saquinavar are used in the treatment of HIV and were also developed by computational methods. Docking studies also revealed how the HIV protease developed resistance toward Nelfinavir and Saquinavar which was beneficial in improving the potency of the drugs [45]. Based on these successful examples, it is clear that computational methods are capable of developing new pharmaceuticals and provide evidence to other researchers that this is a reliable and effective technique in drug discovery.

The cost of bringing a drug to market and the amount of drug resistance profiles emerging are major factors that researchers need to address when designing a drug. It may cause more than 1 billion dollars and 10 years to bring a drug to the market [114]. As we have collected millions of pharmaceutical compounds in a database like Pubchem [115] and ChEMBL [116], we will

need several months or even years to screen them all manually or automatically in the lab, if it is possible we can obtain them all. More and more researchers are turning to computational methods to design drugs in an efficient manner that has the possibility to save money for pharmaceutical companies. While these *in silico* methods are not yet ready to replace *in vivo* and *in vitro* methods and have only brought a few medications to the market, such Raltegravir and Dorzolamide [117], they still provide a valuable insight into the molecular interactions between the ligand and protein. As seen above, there are situations in which computational methods are not always able to accurately determine the results. For this reason, many researchers use *in silico* methods in tandem with other research methods to verify or elucidate standing results. Each time these computational methods verify already established experimental results, their validity in the drug design market has the opportunity to go up. It seems that many researchers are starting to rely on computational results from molecular docking and other computational methods in their research. Often these methods cannot solely generate results that will create a novel drug. However, computational methods are slowly solidifying their place in the pharmaceutical industry as a necessary step toward designing new drugs.

To summarize the cases we reported above (Please see **Table 3**):

1. Most of these projects were designed to recognize a new inhibitor(s) to an enzyme which plays an essential role in a key metabolic/proliferation pathway or the infectious procedure of a pathogen.

2. One or more determined PDB structures of the target protein with good resolution were used, and, often, the key residues of the catalytic reaction, binding/inhibition mechanisms, and drug resistances were revealed based on the docking results.

3. Other computational methods or tools were also used in sequence or in parallel, such as structure prediction, binding site prediction, pharmacophore model, QSAR model, and MD simulation.

4. Drug repurposing has received more and more attention.

Enzymes and membrane proteins (receptors) are two major drug targets. According to previous studies, there is severe bias on the number of determined structures deposit on PDB [118, 119]. A large proportion of solved structures belong to soluble proteins, especially enzymes. It not only made structures of enzymes easier to obtain for molecular docking, but it also made scoring functions/force fields of molecular docking and other related computational approaches to be more accurate for enzymes than membrane proteins. However, we have noticed the importance of membrane receptors, glycol-proteins and non-structure proteins. How to create a reliable strategy to determine or predict the structures of these important drug targets remains a big challenge in molecular docking.

Drug resistance is also a major issue in the failures of treatment of both cancers and infectious diseases. Due to the advancement of docking calculation, we will be able to predict the possible drug resistances and side effects before the treatment or even the drug approval in the future. Therefore, the back-up drugs should be developed and utilized even before drug

Disease	Target protein	PDB ID	Docking Software	Drug(s)	Purpose	Other computational method(s) used
HIV	Integrase	1QS4, 1BIS, 1BLE	GOLD	S-1360	To predict the binding mode	
HIV	Protease	1HXB, 1OHR	AutoDock	Saquinavir, Nelfinavir	To predict the drug resistance	MD simulation
Cancer	Proteasome	2F16	Glide, GOLD	PI-083, MG132, peptide aldehydes	To identify new drug, to understand the binding mechanism	MD simulation
Cancer	Carbonic anhydrases IX	3IAI	AutoDock	ZINC03363328, ZINC08828920, ZINC12941947, ZINC03622539, ZINC1665054	To discover inhibitors	Post-docking energy minimization
Cancer	EGFR	2ITN et al.	AutoDock, jMetalCpp	AMPPNP, Dacomitinib, et al.	To study the effects of mutations	Optimization algorithms
Cancer	CcO	2Y69	Glide	Chlorpromazine	Drug repurposing	
Influenza	Neuraminidase	7NN9	GRID	Zanamivir	To analyze the active site	
Malaria	M18 aspartyl aminopeptidase	4EME	GOLD, Glide	CHEMBL588000 et al.	To identify new drugs	Pharmacophore and QSAR models
Malaria	Lactate dehydrogenase	1LDG	MolDock	Itraconazole, Atorvastatin, Posaconazole	To select potential drugs	
Zika	NS5MTase	3P8Z	AutoDock	New candidates	To identify new drugs	MD simulation
Zika	RNA Polymerase	4WTG	Modeler	Ribavirin, Sofosbuvir	To model and compare ligand binding	
TB	Shikimate kinase	2DFN et al.	MOLDOCK	New candidates	To identify new inhibitors	
TB	Shikimate kinase	2XCS, 2XCT, 2FUM	OpenEye HYBRID, Glide	Diclofenac et al.	Drug repurosing	

Table 3. Summary of the cases presented.

resistance occurs. The improved reliability of molecular docking also facilitates the precision medicine.

As we see in **Figure 1**, the computational approaches play key roles in different steps of the drug discovery process: obtaining the protein structures, binding site prediction, virtual drug screening, binding verification, binding affinity estimation, prediction of drug resistances, binding kinetic modeling, and so on. Molecular docking assists in achieving many objectives in the steps mentioned above effectively and efficiently. Often, it is cheaper, faster than performing

experiments in the biological labs, and we can do even more than conventional approaches. For example, we can predict the potential side effects or drug resistances. Other computational tools such as (3D structure or binding site) prediction models, molecular dynamic simulation, and kinetic modeling have been also well established and applied in different steps of drug discovery to provide more information of target protein or drug efficacy, narrow down the searching spaces/reduce computational load, and/or validate the results of docking. Moreover, drug repurposing is another important application of molecular docking that helps us to enhance the cost- and time-effectiveness of drug development.

In the Era of "Big Data", the accumulated number of protein structures and upgraded computation software and hardware generally improved all related computational methods, not just molecular docking. Based on the progress of the knowledge on protein folding, structural flexibility and molecular recognition, molecular docking has matured. As the core technology of virtual drug discovery, molecular docking will be widely applied to many stages of the drug discovery process.

Author details

Mark Andrew Phillips, Marisa A. Stewart, Darby L. Woodling and Zhong-Ru Xie*

*Address all correspondence to: paulxie@uga.edu

Computational Drug Discovery Laboratory, School of Electrical and Computer Engineering, College of Engineering, University of Georgia, Athens, GA, USA

References

[1] Pagadala NS, Syed K, Tuszynski J. Software for molecular docking: A review. Biophysical Reviews. 2017;**9**(2):91-102

[2] Wong CF. Flexible receptor docking for drug discovery. Expert Opinion on Drug Discovery. 2015;**10**(11):1189-1200

[3] Ferreira LG, Dos Santos RN, Oliva G, Andricopulo AD. Molecular docking and structure-based drug design strategies. Molecules. 2015;**20**(7):13384-13421

[4] Yuriev E, Ramsland PA. Latest developments in molecular docking: 2010–2011 in review. Journal of Molecular Recognition. 2013;**26**(5):215-239

[5] Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The protein data bank. Nucleic Acids Research. 2000;**28**(1):235-242

[6] Berman HM, Henrick K, Nakamura H, Markley J, Bourne PE, Westbrook J. Realism about PDB. Nature Biotechnology. 2007;**25**(8):845-846

[7] Roche DB, Brackenridge DA, McGuffin LJ. Proteins and their interacting partners: An introduction to protein-ligand binding site prediction methods. International Journal of Molecular Sciences. 2015;**16**(12):29829-29842

[8] Xie ZR, Hwang MJ. Methods for predicting protein-ligand binding sites. Methods in Molecular Biology. 2015;**1215**:383-398

[9] Zhu X, Xiong Y, Kihara D: Large-scale binding ligand prediction by improved patch-based method patch-Surfer2.0. Bioinformatics. 2015;**31**(5):707-713

[10] Abdolmaleki A, Ghasemi JB, Ghasemi F. Computer aided drug design for multi-target drug design: SAR/QSAR, molecular docking and pharmacophore methods. Current Drug Targets. 2017;**18**(5):556-575

[11] Desaphy J, Azdimousa K, Kellenberger E, Rognan D. Comparison and druggability prediction of protein-ligand binding sites from pharmacophore-annotated cavity shapes. Journal of Chemical Information and Modeling. 2012;**52**(8):2287-2299

[12] Kumari M, Chandra S, Tiwari N, Subbarao N. 3D QSAR, pharmacophore and molecular docking studies of known inhibitors and designing of novel inhibitors for M18 aspartyl aminopeptidase of plasmodium falciparum. BMC Structural Biology. 2016;**16**(12)

[13] Weiss R. How does HIV cause AIDS? Science. 1993;**260**:1273-1279

[14] Cunningham AL, Donaghy H, Harman AN, Kim M, Turville SG. Manipulation of dendritic cell function by viruses. Current Opinion in Microbiology. 2010;**13**:524-529

[15] Gilbert PB, McKeague IW, Eisen G, Mullins C, Guéye-Ndiaye A, Mboup S, Kanki PJ. Comparison of HIV-1 and HIV-2 infectivity from a prospective cohort study in Senegal. Statistics in Medicine. 2003;**22**(4):573-593

[16] Zheng Y-H, Lovsin N, Peterlin BM. Newly identified host factors modulate HIV replication. Immunology Letters. 2005;**97**(2):225-234

[17] Fan X, Zhang F-H, Al-Safi RI, Zeng L-F, Shabaik Y, Debnath B, Sanchez TW, Odde S, Neamati N, Long Y-Q. Design of HIV-1 integrase inhibitors targeting the catalytic domain as well as its interaction with LEDGF/p75: A scaffold hopping approach using salicylate and catechol groups. Bioorganic & Medicinal Chemistry. 2011;**19**:4935-4952

[18] Serrao E, Odde S, Ramkumar K, Neamati N. Raltegravir, elvitegravir, and metoogravir: The birth of "me-too" HIV-1 integrase inhibitors. Retrovirology. 2009;**6**:25-25

[19] Hazuda a DJ, Felock a P, Witmer a M, Wolfe a A, Stillmock a K, Grobler a JA, Espeseth a A, Gabryelski a L, Schleif a W, Blau a C, et al. Inhibitors of strand transfer that prevent integration and inhibit HIV-1 replication in cells. Science. 2000;(5453):646

[20] Zhuang L, Wai JS, Embrey MW, Fisher TE, Egbertson MS, Payne LS, Guare JP, Jr., Vacca JP, Hazuda DJ, Felock PJ et al: Design and synthesis of 8-hydroxy-[1,6]naphthyridines as novel inhibitors of HIV-1 integrase in vitro and in infected cells. Journal of Medicinal Chemistry. 2003;**46**(4):453-456

[21] Summa V, Petrocchi A, Matassa VG, Gardelli C, Muraglia E, Rowley M, Paz OG, Laufer R, Monteagudo E, Pace P. 4,5-dihydroxypyrimidine carboxamides and N-alkyl-5-hydroxypyrimidinone carboxamides are potent, selective HIV integrase inhibitors with good pharmacokinetic profiles in preclinical species. Journal of Medicinal Chemistry. 2006;**49**(23):6646-6649

[22] Low A, Mohri H, Markowitz M. Frequency of naturally occurring polymorphisms associated with resistance to integrase inhibitors in a recently infected cohort. In: 14th Conference on Ret- roviruses and Opportunistic Infections: Feb 25–28, 2007; Los Angeles, CA. 2007

[23] Gilbert J, Henske P, Singh A. Rebuilding big pharma's business model. The Business & Medicine Report. 2003;**21**(10):73-80

[24] Grabowski H, Vernon J, DiMasi JA. Returns on research and development for 1990s new drug introductions. PharmacoEconomics. 2002;**20**(15):11-29

[25] Lee TH. "Me-too" products—Friend or foe? New England Journal of Medicine. 2004; **350**(3):211-212

[26] Al-Mawsawi LQ, Al-Safi RI, Neamati N. Anti-infectives clinical progress of HIV-1 integrase inhibitors. Expert Opinion on Emerging Drugs. 2008;**13**(2):213-225

[27] Cooper DA, Steigbigel RT, Gatell JM, Rockstroh JK, Katlama C, Yeni P, Lazzarin A, Clotet B, Kumar PN, Eron JE. Subgroup and resistance analyses of raltegravir for resistant HIV-1 infection. New England Journal of Medicine. 2008;**359**(4):355-365

[28] Delelis O, Malet I, Na L, Tchertanov L, Calvez V, Marcelin A-G, Subra F, Deprez E, Mouscadet J-F. The G140S mutation in HIV integrases from raltegravir-resistant patients rescues catalytic defect due to the resistance Q148H mutation. Nucleic Acids Research. 2009;**37**(4):1193-1201

[29] Dayam R, Neamati N. Active site binding modes of the beta-diketoacids: A multi-active site approach in HIV-1 integrase inhibitor design. Bioorganic & Medicinal Chemistry. 2004;**12**(24):6371-6381

[30] Jones G, Willett P, Glen RC, Leach AR, Taylor R. Development and validation of a genetic algorithm for flexible docking. Journal of Molecular Biology. 1997;**267**(3):727-748

[31] Nissink JWM, Murray C, Hartshorn M, Verdonk ML, Cole JC, Taylor R. A new test set for validating predictions of protein–ligand interaction. Proteins: Structure, Function, and Bioinformatics. 2002;**49**(4):457-471

[32] Fikkert V, Hombrouck A, Van Remoortel B, De Maeyer M, Pannecouque C, De Clercq E, Debyser Z, Witvrouw M. Multiple mutations in human immunodeficiency virus-1 integrase confer resistance to the clinical trial drug S-1360. AIDS. 2004;**18**(15):2019-2028

[33] Billich A. S-1360 Shionogi-GlaxoSmithKline. Current Opinion in Investigational Drugs (London, England: 2000). 2003;**4**(2):206-209

[34] Yoshinaga T, Sato A, Fujishita T, Fujiwara T. S-1360: In vitro activity of a new HIV-1 integrase inhibitor in clinical development. In: 9th conference on retroviruses and opportunistic infections, Seattle, WA: 2002. 55

[35] Davies DR. The structure and function of the aspartic proteinases. Annual Review of Biophysics and Biophysical Chemistry. 1990;**19**(1):189-215

[36] Huang X, Britto MD, Kear-Scott JL, Boone CD, Rocca JR, Simmerling C, Mckenna R, Bieri M, Gooley PR, Dunn BM. The role of select subtype polymorphisms on HIV-1 protease conformational sampling and dynamics. Journal of Biological Chemistry. 2014; **289**(24):17203-17214

[37] Kräusslich H-G, Ingraham RH, Skoog MT, Wimmer E, Pallai PV, Carter CA. Activity of purified biosynthetic proteinase of human immunodeficiency virus on natural substrates and synthetic peptides. Proceedings of the National Academy of Sciences. 1989; **86**(3):807-811

[38] Seelmeier S, Schmidt H, Turk V, Von Der Helm K. Human immunodeficiency virus has an aspartic-type protease that can be inhibited by pepstatin a. Proceedings of the National Academy of Sciences. 1988;**85**(18):6612-6616

[39] Goodman LS. Goodman and Gilman's the Pharmacological Basis of Therapeutics. vol. 1549. New York: McGraw-Hill; 1996

[40] Brik A, Wong C-H. HIV-1 protease: Mechanism and drug discovery. Organic & Biomolecular Chemistry. 2003;**1**(1):5-14

[41] Vacca JP, Condra JH. Clinically effective HIV-1 protease inhibitors. Drug Discovery Today. 1997;**2**(7):261-272

[42] Kalish V, Kaldor S, Shetty B, Tatlock J, Davies J, Hammond M, Dressman B, Fritz J, Appelt K, Reich S. Iterative protein structure-based drug design and synthesis of HIV protease inhibitors. European Journal of Medicinal Chemistry. 1995;**30**:201s-214s

[43] Vacca JP, Guare J, Desolms S, Sanders W, Giuliani E, Young S, Darke P, Sigal I, Schleif W. L-687,908, a potent hydroxyethylene containing HIV protease inhibitor. Journal of Medicinal Chemistry. 1991;**34**(3):1225-1228

[44] Jenwitheesuk E, Samudrala R. Prediction of HIV-1 protease inhibitor resistance using a protein-inhibitor flexible docking approach. Antiviral Therapy. 2005;**10**(1):157-166

[45] Jenwitheesuk E, Samudrala R. Improved prediction of HIV-1 protease-inhibitor binding energies by molecular dynamics simulations. BMC Structural Biology. 2003;**3**(1):2

[46] Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, Olson AJ. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. Journal of Computational Chemistry. 1998;**19**(14):1639-1662

[47] Shenderovich MD, Kagan RM, Heseltine PN, Ramnarayan K. Structure-based phenotyping predicts HIV-1 protease inhibitor resistance. Protein Science. 2003;**12**(8):1706-1718

[48] Vos T, Allen C, Arora M, Barber RM, Bhutta ZA, Brown A, Carter A, Casey DC, Charlson FJ, Chen AZ. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990-2015: A systematic analysis for the global burden of disease study 2015. The Lancet. 2016;**388**(10053):1545

[49] Organization WH. Chagas disease (American trypanosomiasis). Fact sheet No. 340. 2016. In. 2016

[50] Bellera CL, Balcazar DE, Vanrell MC, Casassa AF, Palestro PH, Gavernet L, Labriola CA, Gálvez J, Bruno-Blanch LE, Romano PS. Computer-guided drug repurposing: Identification of trypanocidal activity of clofazimine, benidipine and saquinavir. European Journal of Medicinal Chemistry. 2015;**93**:338-348

[51] Dudley JT, Deshpande T, Butte AJ. Exploiting drug-disease relationships for computational drug repositioning. Briefings in Bioinformatics. 2011;**12**(4):303-311

[52] Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. CA: a Cancer Journal for Clinicians. 2016;**66**(1):7-30

[53] Allen T. Ligand-targeted therapuetics in anticancer therapy. Nature Reviews Cancer. 2002;**2**(10):750-763

[54] Peters JM. Proteasomes: Protein degradation machines of the cell. Trends in Biochemical Sciences. 1994;**19**(9):377-382

[55] Kazi A, Lawrence H, Guida WC, McLaughlin ML, Springett GM, Berndt N, Yip RM, Sebti SM. Discovery of a novel proteasome inhibitor selective for cancer cells over non-transformed cells. Cell Cycle. 2009;**8**(12):1940-1951

[56] Guedes RA, Serra P, Salvador JAR, Guedes RC. Computational approaches for the discovery of human proteasome inhibitors: An overview. Molecules. 2016;**21**(7):927-953

[57] Kane RC, Farrell AT, Sridhara R, Pazdur R. United States Food and Drug Administration approval summary: Bortezomib for the treatment of progressive multiple myeloma after one prior therapy. Clinical Cancer Research. 2006;**12**(10):2955-2960

[58] Lashinger LM, Zhu K, Williams SA, Shrader M, Dinney CP, McConkey DJ. Bortezomib abolishes tumor necrosis factor-related apoptosis-inducing ligand resistance via a p21-dependent mechanism in human bladder and prostate cancer cells. Cancer Research. 2005;**65**(11):4902-4908

[59] Liu J, Zhang H, Xiao Z, Wang F, Wang X, Wang Y. Combined 3D-QSAR, molecular docking and molecular dynamics study on derivatives of peptide epoxyketone and tyropeptin-boronic acid as inhibitors against the beta5 subunit of human 20S proteasome. International Journal of Molecular Sciences. 2011;**12**(3):1807-1835

[60] Zhang S, Shi Y, Jin H, Liu Z, Zhang L, Zhang L. Covalent complexes of proteasome model with peptide aldehyde inhibitors MG132 and MG101: Docking and molecular dynamics study. Journal of Molecular Modeling. 2009;**15**(12):1481-1490

[61] Ma Y, Xu B, Fang Y, Yang Z, Cui J, Zhang L, Zhang L. Synthesis and SAR study of novel peptide aldehydes as inhibitors of 20S proteasome. Molecules. 2011;**16**(9):7551-7564

[62] Spikes JD, Jori G. Photodynamic therapy of tumours and other diseases using porphyrins. Lasers in Medical Science. 1987;**2**(1):3-15

[63] Santoro AM, Lo Giudice MC, D'Urso A, Lauceri R, Purrello R, Milardi D. Cationic porphyrins are reversible proteasome inhibitors. Journal of the American Chemical Society. 2012;**134**(25):10451-10457

[64] Supuran CT. Carbonic anhydrases: Novel therapeutic applications for inhibitors and activators. Nature Reviews Drug Discovery. 2008;**7**(2):168-181

[65] Krungkrai J, Supuran CT. The alpha-carbonic anhydrase from the malaria parasite and its inhibition. Current Pharmaceutical Design. 2008;**14**(7):631-640

[66] Pinard MA, Boone CD, Rife BD, Supuran CT, McKenna R. Structural study of interaction between brinzolamide and dorzolamide inhibition of human carbonic anhydrases. Bioorganic & Medicinal Chemistry. 2013;**21**(22):7210-7215

[67] Amresh P, Kumar K, Islam A, Hassan MI, Ahmad F. Receptor Chemoprint derived Pharmacophore model for development of CAIX Inihibitors. Carcinogenesis & Mutagenesis. 2013;**S8**(003)

[68] Gazdar AF. Activating and resistance mutations of EGFR in non-small-cell lung cancer: Role in clinical response to EGFR tyrosine kinase inhibitors. Oncogene. 2009;**28**:S24-S31

[69] Anderson NG, Ahmad T, Chan K, Dobson R, Bundred NJ. ZD1839 (Iressa), a novel epidermal growth factor receptor (EGFR) tyrosine kinase inhibitor, potently inhibits the growth of EGFR-positive cancer cell lines with or without erbB2 overexpression. International Journal of Cancer. 2001;**94**(6):774-782

[70] Engelman JA, Cantley LC. A sweet new role for EGFR in cancer. Cancer Cell. 2008;**13**(5): 375-376

[71] Garcia-Godoy MJ, Lopez-Camacho E, Garcia-Nieto J, Nebro AJ, Aldana-Montes JF. Molecular docking optimization in the context of multi-drug resistant and sensitive EGFR mutants. Molecules. 2016;**21**(11)

[72] Mahajan P, Suri N, Mehra R, Gupta M, Kumar A, Sing SK, Nargotra A. Discovery of novel small molecule EGFR inhibitory leads by structure and ligand-based virtual screening. Medicinal Chemistry Research. 2016;**26**(74):74-92

[73] Oliva CR, Zhang W, Langford C, Suto MJ, Griguer CE. Repositioning chlorpromazine for treating chemoresistant glioma through the inhibition of cytochrome c oxidase bearing the COX4-1 regulatory subunit. Oncotarget. 2017;**8**(23):37568-37583

[74] Brunori M, Antonini G, Malatesta F, Sarti P, Wilson MT. Cytochrome-*c* oxidase. European Journal of Biochemistry. 1987;**169**:1-8

[75] Tsukihara T, Aoyama H, Yamashita E, Tomizaki T, Yamaguchi H, Shinzawa-Itoh K, Nakashima R, Yaono R, Yoshikawa S. The whole structure of the 13-subunit oxidized cytochrome c oxidase at 2.8 A. Science. 1996;**272**(5265):1136-1144

[76] Gerdil C. The annual production cycle for influenza vaccine. Vaccine. 2003;**21**(16):1776-1779

[77] Hayden FG. Pandemic influenza: Is an antiviral response realistic? The Pediatric Infectious Disease Journal. 2004;**23**(11 Suppl):S262-S269

[78] Moorthy NS, Poongavanam V, Pratheepa V. Viral M2 ion channel protein: A promising target for anti-influenza drug discovery. Mini Reviews in Medicinal Chemistry. 2014;**14**(10):819-830

[79] Moscona A. Neuraminidase inhibitors for influenza. The New England Journal of Medicine. 2005;**353**(13):1363-1373

[80] Davies WL, Grunert RR, Haff RF, McGahen JW, Neumayer EM, Paulshock M, Watts JC, Wood TR, Hermann EC, Hoffmann CE. Antiviral activity of 1-Adamantanamine (amantadine). Science. 1964;**144**(3620):862-863

[81] Hayden FG, Belshe RB, Clover RD, Hay AJ, Oakes MG, Soo W. Emergence and apparent transmission of rimantadine-resistant influenza A virus in families. The New England Journal of Medicine. 1989;**321**(25):1696-1702

[82] Alves Galvao MG, Rocha Crispino Santos MA, Alves da Cunha AJ. Amantadine and rimantadine for influenza A in children and the elderly. Cochrane Database Systematic Review. 2014;**11**. CD002745

[83] von Itzstein M: The war against influenza: Discovery and development of sialidase inhibitors. Nature Reviews Drug Discovery. 2007;**6**(12):967-974

[84] Varghese JN, Laver WG, Colman PM. Structure of the influenza virus glycoprotein antigen neuraminidase at 2.9 a resolution. Nature. 1983;**303**(5912):35-40

[85] Gubareva LV, Sleeman K, Guo Z, Yang H, Hodges E, Davis CT, Baranovich T, Stevens J. Drug susceptibility evaluation of an influenza a(H7N9) virus by Analyzing recombinant neuraminidase proteins. The Journal of Infectious Diseases. 2017;**216**(suppl_4):S566-S574

[86] Colman PM, Varghese JN, Laver WG. Structure of the catalytic and antigenic sites in influenza virus neuraminidase. Nature. 1983;**303**(5912):41-44

[87] Vonitzstein M, Wu WY, Kok GB, Pegg MS, Dyason JC, Jin B, Phan TV, Smythe ML, White HF, Oliver SW, et al. Rational design of potent sialidase-based inhibitors of influenza-virus replication. Nature. 1993;**363**(6428):418-423

[88] Hayden FG, Osterhaus AD, Treanor JJ, Fleming DM, Aoki FY, Nicholson KG, Bohnen AM, Hirst HM, Keene O, Wightman K. Efficacy and safety of the neuraminidase inhibitor zanamivir in the treatment of influenzavirus infections. GG167 influenza study group. The New England Journal of Medicine. 1997;**337**(13):874-880

[89] Winquist AG, Fukuda K, Bridges CB, Cox NJ. Neuraminidase inhibitors for treatment of influenza A and B infections. In. Edited by Services DoHaH. In: Morbidity and Mortality Weekly Report: Centers for Disease Control and Prevention. Vol. 48. 1999. pp. 1-9

[90] Newton CR, Krishna S. Severe falciparum malaria in children: Current understanding of pathophysiology and supportive treatment. Pharmacology & Therapeutics. 1998;**79**(1):1-53

[91] Wellems TE. Malaria. How chloroquine works. Nature. 1992;**355**(6356):108-109

[92] Chinappi M, Via A, Marcatili P, Tramontano A. On the mechanism of Chloroquine resistance in plasmodium falciparum. PLoS One. 2010;**5**(11)

[93] Egan TJ, Ncokazi KK. Quinoline antimalarials decrease the rate of beta-hematin formation. Journal of Inorganic Biochemistry. 2005;**99**(7):1532-1539

[94] Dunn CR, Banfield MJ, Barker JJ, Higham CW, Moreton KM, Turgut-Balik D, Brady RL, Holbrook JJ. The structure of lactate dehydrogenase from plasmodium falciparum reveals a new target for anti-malarial design. Nature Structural Biology. 1996;**3**(11):912-915

[95] Penna-Coutinho J, Cortopassi WA, Oliveira AA, Franca TC, Krettli AU. Antimalarial activity of potential inhibitors of plasmodium falciparum lactate dehydrogenase enzyme selected by docking studies. PLoS One. 2011;**6**(7):e21237

[96] Sikka V, Chattu VK, Popli RK, Galwankar SC, Kelkar D, Sawicki SG, Stawicki SP, Papadimos TJ. The emergence of Zika virus as a Global Health security threat: A review and a consensus statement of the INDUSEM joint working group (JWG). Journal of Global Infectious Disease. 2016;**8**(1):3-15

[97] Cao-Lormeau VM, Blake A, Mons S, Lastere S, Roche C, Vanhomwegen J, Dub T, Baudouin L, Teissier A, Larre P, et al. Guillain-Barre syndrome outbreak associated with Zika virus infection in French Polynesia: A case-control study. Lancet. 2016;**387**(10027):1531-1539

[98] Duffy MR, Chen TH, Hancock WT, Powers AM, Kool JL, Lanciotti RS, Pretrick M, Marfel M, Holzbauer S, Dubray C, et al. Zika virus outbreak on Yap Island, Federated States of Micronesia. The New England Journal of Medicine. 2009;**360**(24):2536-2543

[99] Musso D. Zika virus transmission from French Polynesia to Brazil. Emerging Infectious Diseases. 2015;**21**(10):1887

[100] Dong HP, Fink K, Zust R, Lim SP, Qin CF, Shi PY. Flavivirus RNA methylation. Journal of General Virology. 2014;**95**:763-778

[101] Zhang C, Feng T, Cheng J, Li Y, Yin X, Zeng W, Jin X, Li Y, Guo F, Jin T. Structure of the NS5 methyltransferase from Zika virus and implications in inhibitor design. Biochemical and Biophysical Research Communications; 2016

[102] Lim SP, Sonntag LS, Noble C, Nilar SH, Ng RH, Zou G, Monaghan P, Chung KY, Dong H, Liu B, et al, Small molecule inhibitors that selectively block dengue virus methyltransferase. The Journal of Biological Chemistry. 2011;**286**(8):6233-6240

[103] Ramharack P, Soliman MES. Zika virus NS5 protein potential inhibitors: An enhanced in silico approach in drug discovery. Journal of Biomolecular Structure & Dynamics. 2017:1-16

[104] Sacramento CQ, de Melo GR, de Freitas CS, Rocha N, Hoelz LV, Miranda M, Fintelman-Rodrigues N, Marttorelli A, Ferreira AC, Barbosa-Lima G et al: The clinically approved antiviral drug sofosbuvir inhibits Zika virus replication. Scientific Reports. 2017;**7**:40920

[105] Elfiky AA. Zika viral polymerase inhibition using anti-HCV drugs both in market and under clinical trials. Journal of Medical Virology. 2016;**88**(12):2044-2051

[106] Hershkovitz I, Donoghue HD, Minnikin DE, Besra GS, Lee OY, Gernaey AM, Galili E, Eshed V, Greenblatt CL, Lemma E, et al. Detection and molecular characterization of 9,000-year-old mycobacterium tuberculosis from a Neolithic settlement in the eastern Mediterranean. PLoS One. 2008;**3**(10):e3426

[107] Keshavjee S, Farmer PE. Tuberculosis, drug resistance, and the history of modern medicine. The New England Journal of Medicine. 2012;**367**(10):931-936

[108] Centers for Disease Control and Prevention. Plan to Combat Extensively Drug-Resistant Tuberculosis: Recommendations of the Federal Tuberculosis Task Force. MMWR Recomm Report. 2009;**58**(RR-3):1-43

[109] Velayati AA, Masjedi MR, Farnia P, Tabarsi P, Ghanavi J, ZiaZarifi AH, Hoffner SE. Emergence of new forms of totally drug-resistant tuberculosis bacilli: Super extensively drug-resistant tuberculosis or totally drug-resistant strains in iran. Chest. 2009;**136**(2): 420-425

[110] Ratledge C, Patel PV, Mundy J. Iron transport in mycobacterium smegmatis: The location of mycobactin by electron microscopy. Journal of General Microbiology. 1982;**128**(7): 1559-1565

[111] Vianna CP, de Azevedo WF, Jr.: Identification of new potential mycobacterium tuberculosis shikimate kinase inhibitors through molecular docking simulations. Journal of Molecular Modeling. 2012;**18**(2):755-764

[112] Kahlous NA, Bawarish MA, Sarhan MA, Kupper M, Hasaba A, Rajab M. Using Chemoinformatics, bioinformatics, and bioassay to predict and explain the antibacterial activity of nonantibiotic Food and Drug Administration drugs. Assay and Drug Development Technologies. 2017;**15**(3):89-105

[113] Brindha S, Vincent S, Velmurugan D, Ananthakrishnan D, Sundaramurthi JC, Gnanadoss JJ. Bioinformatics approach to prioritize known drugs towards repurposing for tuberculosis. Medical Hypotheses. 2017;**103**:39-45

[114] DiMasi JA, Grabowski HG, Hansen RW. Innovation in the pharmaceutical industry: New estimates of R&D costs. Journal of Health Economics. 2016;**47**:20-33

[115] Kim S, Thiessen PA, Bolton EE, Chen J, Fu G, Gindulyte A, Han L, He J, He S, Shoemaker BA, et al. PubChem substance and compound databases. Nucleic Acids Research. 2016;**44**(D1):D1202-D1213

[116] Gaulton A, Bellis LJ, Bento AP, Chambers J, Davies M, Hersey A, Light Y, McGlinchey S, Michalovich D, Al-Lazikani B, et al. ChEMBL: A large-scale bioactivity database for drug discovery. Nucleic Acids Research. 2012;**40**(Database issue):D1100-D1107

[117] Kubinyi H. Chance favors the prepared mind–from serendipity to rational drug design. Journal of Receptor and Signal Transduction Research. 1999;**19**(1–4):15-39

[118] Peng K, Obradovic Z, Vucetic S: Exploring bias in the protein data Bank using contrast classifiers. Pacific symposium on Biocomputing Pacific Symposium on Biocomputing 2004:435-446

[119] Xie ZR, Liu CK, Hsiao FC, Yao A, Hwang MJ. LISE: A server using ligand-interacting and site-enriched protein triangles for prediction of ligand-binding sites. Nucleic Acids Research. 2013;**41**(Web Server issue):W292-W296