An extensive survey of molecular docking tools and their applications using text mining and deep curation strategies.

Kamal Rawal[1#], Tanishka Khurana[1], Himanshu Sharma[1], Sadika Verma[1], Simmi Gupta[1], Chahat Kubba[1], Ulrich Strych[2], Peter Hotez[2, 3], Maria Elena Bottazzi[2, 3]

1. **Department of Biotechnology, Jaypee Institute of Information and Technology, Noida, India.**
2. **Texas Children's Hospital Center for Vaccine Development, Departments of Pediatrics and Molecular Virology and Microbiology, National School of Tropical Medicine, Baylor College of Medicine, Houston, TX, USA.**
3. **Department of Biology, Baylor University, Waco, Texas, USA.**

#Corresponding Author:
Dr. Kamal Rawal
Jaypee Institute of Information Technology, A-10, Sector -62, NOIDA-201307, Uttar Pradesh, India
Email address: kamal.rawal@gmail.com

ABSTRACT

The technology of docking molecules *in-silico* has evolved significantly in recent years and has become a crucial component of the drug discovery tool process that includes virtual screening, lead optimization, and side-effect predictions. To date over 43,000 abstracts/papers have been published on docking, thereby highlighting the importance of this computational approach in the context of drug development. Considering the large amount of genomic and proteomic consortia active in the public domain, docking can exploit this data on a correspondingly 'large scale' to address a variety of research questions. Over 160 robust and accurate molecular docking tools based on different algorithms have been made available to users across the world. Further, 109 scoring functions have been reported in the literature till date. Despite these advancements, there continue to be several bottlenecks during the implementation stage. These problems or issues range from choosing the right docking algorithm, selecting a binding site in target proteins, performance of the given docking tool, integration of molecular dynamics information, ligand-induced conformational changes, use of solvent molecules, choice of docking pose, and choice of databases. Further, so far, not always have experimental studies been used to validate the docking results. In this review, basic features and key concepts of docking have been highlighted, with particular emphasis on its applications such as drug repositioning and prediction of side effects. Also, the use of docking in conjunction with wet lab experimentations and epitope predictions has been summarized. Attempts have been made to systematically address the above-mentioned challenges using expert-curation and text mining strategies. Our work shows the use of machine-assisted literature mining to process and analyze huge amounts of available information in a short time frame. With this work, we also propose to build a platform that combines human expertise (deep curation) and machine learning in a collaborative way and thus helps to solve ambitious problems (i.e. building fast, efficient docking systems by combining the best tools or to perform large scale docking at human proteome level).

Website and other links: We have created web based forms and a website so that scientists, developers and users of molecular docking tools can share their experiences and expertise to build a comprehensive resource on molecular docking. In addition, the collected information shall be used to update the molecular docking website and future versions of this manuscript.

The website(s) associated with this paper contain additional information in the form of tables and figures. The information provided on the website(s) is updated on periodic basis.

A) https://tinyurl.com/sci-net2000

B) https://tinyurl.com/docking-tools

C) https://tinyurl.com/networks-docking

D) https://tinyurl.com/docking-review

Keywords:

Side effect prediction; adverse drug reactions prediction; drug repositioning; drug repurposing; drug indication prediction, docking, tools, software, database, benchmarking, wet lab validations, collaborative writing.

INTRODUCTION

A major challenge in the healthcare field is to devise a systematic strategy to integrate diverse biological datasets to provide insight into disease, pathogenesis or discover new and safe drugs/vaccines against complex diseases. The process encompasses a period of intense research, typically involving a span of 10-15 years and a huge investment of sometimes more than $1 billion per product [Hughes et al. 2011]. Given the experimental difficulties of attaining knowledge on the ligand-target interaction at the molecular level, numerous high performing computational platforms and a wealth of structural data are now being increasingly used for enhancing the efficiency and speed of the drug discovery process. As it has been said, substantial progress has been witnessed in recent years for studying protein-ligand interactions over the traditional paradigm. The computational technique known as “docking” has permeated all aspects of the drug discovery process such as virtual screening, lead optimization, and side effect predictions and essentially acts as a complementary tool to predict the structure of a specific complex formed by two given interacting proteins. Docking holds a significant promise to screen potential drugs as well as drug targets and elucidate biomolecular interactions. Its applications (at larger scale) can be seen through public projects such as OpenZika (http://openzika.ufg.br/), which involves the screening of potential compounds against the models of Zika protein structures. The mechanistic approach of docking can also play a pivotal role in predicting

adverse drug reactions (ADRs) for early screening of hazardous drug molecules, which is initiated by intended, on-target binding or promiscuous binding of drugs to an off-target protein. Highly publicized examples of phase IV failures including rosiglitazone ("Avandia") [Nissen et al. 2010] and rofecoxib ("Vioxx") [Karha et al. 2004] are indicative of the fact that the current approach of the pharmaceutical industry involving the use of in vitro toxicity panels to assay small molecule binding is inadequate [Blomme et al. 2015] and there exists a necessity to explore docking technologies in order to develop safer medicines. Another field where docking finds its application is drug repositioning in which already existing compounds can be repurposed to new potential therapeutic targets. The technique has become progressively mainstream in recent years and is believed to be of particular use in speeding up drug discovery by inspecting new uses of existing, accepted drugs [Ekins et al. 2017]. This review thus provides basic insights into the specific features and concepts of docking, with particular emphasis on applications of docking in the field of side effect prediction and drug repositioning, so as to develop a more rational and targeted therapy. We also discuss the role of software tools and online web services and provide a critical analysis to compare their performance on benchmark datasets along with the challenges of current docking models. To make this review comprehensive and accurate, we used Perl and Python based text mining/machine learning systems (developed in-house) to assist expert curators to analyse and curate a large number of papers/abstracts [Kuhl et al. 1984]. Further, to keep this review updated and to build an ambitious large-scale docking pipeline using the expertise of practitioners/users of molecular docking and tool developers, we have initiated an international collaborative effort using network sciences involving multiple organizations and researchers as co-authors of future versions of this paper. This initiative based upon the principles of network sciences, is expected to improve research quality, advance efficiency of the scientific production, and foster breakthroughs in a shorter time. Here, we also discuss our ongoing collaborative efforts to discover new vaccine targets using network sciences and the use of docking combined with experimental techniques in the area of Chagas Disease.

2. BACKGROUND

2.1: The illustrious history of docking engines (algorithms)

Following the advent of docking algorithms in the 1980s [Billeter et al. 1987] along with the advancement of techniques such as X-ray crystallography, nuclear magnetic resonance spectroscopy and high-throughput protein purification, molecular docking has now become the most commonly used method among the various rational approaches that are currently being pursued for drug discovery and development [Lemmon et al. 2012]. Simulated docking processes aim to predict the interaction of known structures (i.e. receptors, enzymes) with one or more ligands using computational procedures, principally to exploit their novel relationships to discover the chemical entities that exhibit strong binding energies for the active site of the relevant target molecule [Chhabra et al. 2007]. This is achieved by testing various poses (binding conformations between ligand and protein) which are subsequently ranked via a scoring function [DeLuca et al. 2015]. Protein-ligand docking can broadly be divided into three classes namely rigid body docking (where receptor and ligand conformations are fixed), semi-flexible ligand docking (the ligand's internal bond rotation is allowed and receptor is held fixed or the receptor is considered as flexible and the ligand is treated as a fixed molecule) and flexible docking (both molecules are considered flexible) [Halperin et al. 2002]. Rigid docking has been used in the majority of the docking software. It is relatively less demanding with respect to computing power when searching the space of the docked conformations. Whereas, flexible docking is computationally demanding and provides better results since its conjecture about the binding geometries of ligands surpass rigid-receptor docking [Camacho et al. 2002]. The representative set of docking tools used in each type of docking has been summarized in **Table 1**. Computational biologists have used a wide variety of computational techniques in docking studies/tools which includes evolutionary programming, fast Fourier transform, genetic algorithms, guided differential evolution, incremental construction, fragment-based approaches, multiple copy approach, matching algorithm, molecular dynamics, Monte Carlo simulations, simulated annealing, and Tabu search (See **Table 2**). Each technique offers unique advantages to the user for conducting docking studies. In the present work, we describe features of a variety of docking tools, along with their disadvantages so that a user is able to select the right algorithm for their research work.

158 **2.2: Ab-initio Vs Knowledge Based Docking**

159 Traditionally, energy landscapes are used in solving protein structures. The outlook becomes
160 extremely complicated when we consider interactions of two molecules and intend to find global
161 minima [Vakser et al. 1996, Ruvinsky et al. 1996, Vakser et al. 2008]. Current protocols are
162 based upon concepts of physics (steric complementarity) [Katchalski-Katzir et al. 1992, Vakser
163 et al. 1997, O'Toole N et al. 2008, Vakser et al. 2008, Ruvinsky et al. 2008, Vakseret et al. 2008
164] and on the techniques borrowed from computer science and other engineering disciplines which
165 includes pattern recognition, optimization, machine learning, etc. In knowledge-based docking
166 approaches, strategies are adopted from comparative modelling systems. These includes
167 approaches based on comparison/alignment of sequences [Aloy et al. 2003, Kundrotas et al.
168 2008, Rodrigues et al. 2013], sequences and structures (i.e. threading) [Lu et al. 2002, Guerler et
169 al. 2013, Szilagyi et al. 2014], or only on the structures [Szilagyi et al. 2014, Günther et al. 2007,
170 Zhang et al. 2012, Ghoorah et al. 2011, Tuncbag et al. 2012, Sinha et al. 2010, Kundrotas et al.
171 2013] because the structures of the protein to be docked are assumed to be known by the very
172 definition of docking. In a 2012 research study, it was reported that, in spite of the limited
173 number of protein-protein complexes in the Protein Data Bank, docking templates can be found
174 for complexes representing almost all known protein-protein interactions, provided the
175 components themselves have a known structure or can be homology-built [Kundrotas et al.
176 2012]. In 2005, an approach named TM-align was described to identify the best structural
177 alignment between protein pairs that combines the TM-score rotation matrix and Dynamic
178 Programming (DP) which built a foundation for template-based docking [Zhang et al. 2005]. The
179 translational, rotational and conformational degree of freedom facilitates a large number of
180 binding modes between the ligand and protein molecules. Therefore, various sampling
181 algorithms have been deployed to overcome the infeasibility of computational generation of
182 attainable conformations [Sherman et al. 2006]. The process is supported with the structural and
183 affinity information available in the databases such as Protein Data Bank (PDB) [Schneidman-
184 Duhovny et al. 2005], ZINC [Irwin et al. 2005], PubChem [Wang et al. 2009], DrugBank,
185 PDBBIND [Wang et al. 2004], ChemDB [Chen et al. 2005], AffinDB [Block et al. 2006], PLD
186 [Puvanendrampillai et al. 2003] and CREDO [Schreyer et al. 2009], which aids the development
187 and validation of these algorithms.
188

2.3: Docking methods and scoring functions

The conformations obtained during docking are ranked via a scoring function, which accurately represents energetically favourable protein-ligand complexes and differentiates valid binding pose predictions from invalid ones. Three types of scoring functions are mainly employed for predicting target-ligand binding affinity. First, the force-field or molecular mechanics-based scoring functions, which can utilize the sum of van der Waals and electrostatic interactions and access the binding free energy of protein-ligand complexes; this scoring function has been used in DOCK [Raha et al. 2004]. The van der Waals energies are computed using Lennard-Jones potentials and electrostatic terms and are represented by coulomb interactions (with distance-dependent dielectric constant). Second, the empirical scoring function which is based upon evaluation of binding energy due to various energy components like hydrogen bonds, binding entropy, ionic interaction, and hydrophobic effect. Third, a knowledge-based scoring function in which statistical analysis of a co-crystallized ligand-protein complex is employed and contact frequencies and/or distances between a protein and its ligand is obtained [Beutler et al. 1994]. It evaluates the final score by promoting preferred contacts and penalizing repulsive interactions between each ligand-protein atom [Liu et al. 2015]. **Table 3** summarizes the above-mentioned scoring functions. Using text mining scripts, we found that over 107 scoring functions have been published till date (**Table 8**). In terms of choice of scoring functions, Feher proposed the use of a consensus scoring function rather than relying on a single system to improve the predictions [Feher et al. 2006]. In 2015, Chen et al. claimed that a weighted scoring system performs better than a consensus-based method. From the user's perspective, the choice of a rigid versus a flexible type of docking is dependent upon factors such as availability of computational hardware, the character of the target protein, the number of ligands and the number of target proteins used in the study, Also deserving consideration is whether the binding pocket will change the shape of the binding site, etc. (**Table 4**) [Chen et al. 2015]. In addition, the user also faces questions about the choice of software for conducting these docking simulations. When we searched for the top-ranking docking algorithms in a web or literature search, AutoDock and GOLD appeared as the top-ranking tools based on the number of citations and the popularity in internet searches. Though these are popular programs, they are not necessarily more accurate than others. As we can see from the comparative analysis in **Table 5** and the **Table 7**, each program offers unique advantages, but also has several limitations. Therefore, the user is always

advised to carefully review the details for each program and also consider other relevant tools (**Rosetta -http://boinc.bakerlab.org/**) [Li et al. 2006]. The next important factor to consider is the availability of a reliable target protein structure. PDB and the structure databases serve as starting points to search protein targets but the user must consider the quality and reliability of the structure using meta-information such as details of X-crystallography experiments namely resolution and conditions under which protein crystal was obtained. User may also consider employing molecular docking (MD), energy minimization or clustering to improve the structure before start of docking experiments [Huang et al. 2010]. MD offers several benefits which includes depiction of mechanism of action of compounds [Gohlke et al. 2000], confirming experimental findings such as ensuring the stability of protein and the candidate compound binding [Milan et al. 2015] and modelling the potency of multi-target drugs through in-silico tests [Li et al. 2014], but one issue which has come into focus lately is that an increasing number of available protein sequences does not have existing PDB entries, with the ratio of the former to the latter showing an alarming trend for the worse. It has been reported that in 2012, only one in 200 entries in UniProt had a corresponding PDB entry; the figure for 2007 was 100 [Buturak et al. 2014]. Therefore, in case the user is looking to increase the search space of target proteins for a given ligand(s), it will be advisable to include large scale automated 3D structure prediction programs before undertaking docking studies [Lee et al. 2014]. Conventionally docking programs restrict the search to small size binding sites (pockets) and small number of interacting residues otherwise the search time becomes impractically long and complex. Therefore, the user is encouraged to list docking sites during the preparatory phase. In case the target site is not known (blind docking), researchers split the docking box into multiple boxes, or repeat the search several times using different seeds, and then merge the results manually. Tools such as QuickVina-W [Trott et al. 2010] are useful in situations where target sites are not known beforehand. A novel virtual screening tool namely 'SQM/COSMO filter' (featuring semi-empirical quantum mechanics (SQM), Cabrera et al. (2011) has evidently outperformed the most widely used scoring tools. There have also been calls for changing current approaches since comparison of binding sites of proteins is more useful than comparing entire sequences and structures of the protein [Cabrera et al. 2011].

2.4: Protein-Protein Docking

In recent times, docking is moving from the standard drug-ligand interaction to study protein-protein interactions as well. The large part of this interest is driven by CAPRI (Critical Assessment of Predicted Interactions); an experiment involving separate groups of predictors and assessors (http://www.ebi.ac.uk/msd-srv/capri/)[Janin et al. 2013]. CAPRI is a blind prediction experiment which uses unpublished crystal or NMR structures of complexes, communicated on a confidential basis by their authors to the CAPRI management. The predictor group build models of based upon their algorithms and assessors evaluate their predictions in context of experimental information. Though the principles behind protein-protein docking are similar to protein-ligand docking, specialised programs are being developed due to the increased complexity of the system. On one hand, protein-protein docking programs need to deal with the conformational changes between unbound and bound structures, but on the other hand the inaccuracies of the interacting modelled structures present challenges. Over the past decade, protein–protein docking has significantly evolved from initial *ab-initio* docking [Katchalski-Katzir et al. 1992, Vakser et al. 1997] to interface-guided docking [de Vries et al. 2007].

A 2009 CAPRI study reported that there are 3 classes of methods for protein-protein docking. The global method, based on Fast Fourier Transformation (e.g. ZDOCK, PatchDock tools), the medium range method, based on Monte Carlo minimization (e.g. Rosetta-dock tool) and a restraint-based method, where prior information on the interface residues is available (e.g. the HADDOCK tool). Resources such as Dockground [Douguet et al. 2006, Gao et al. 2007] and benchmark datasets from Weng's group are playing an important role in this domain of docking [Huang et al. 2013]. Ruvinsky et al. (2012) presented a systematic large-scale analysis of conformational changes in the side chains during protein-protein interaction. Following on the same work, they developed a tool named "HingeProt" which separates proteins into their rigid parts and the hinge regions connecting them. The method is useful in flexible protein-protein and protein-ligand docking, flexible docking of protein structures into cryo-EM maps, and refinement of low-resolution EM structures. Tools such as DOT program finds low-energy docked structures for two proteins by performing a systematic search over six degrees of freedom by incorporating Poisson–Boltzmann electrostatic energy and a van der Waals energy, each represented as a grid-based correlation function [Mandell et al 2001]. Apart from these, methods have been developed for discretization of the conformational space into rotameric states [Beglov et al. 2011, Ruvinsky et al. 2012]. Protein-Protein docking problems become more

complicated when docking of modelled protein structures is performed. This is because models are considered to be less accurate than experimentally determined structures. Tovchigrechko et al. (2002) presented a prediction system on co-crystallized complexes for low resolution docking of protein models. In a recent work, Anishchenko et al. (2014) contributed that meaningful results in these circumstances can be obtained through carefully curated sets of structures with levels of distortion typical for modelled proteins. It is to be noted that each of these methods is suitable for different families of proteins [Tovchigrechko et al. 2002, Anishchenko et al. 2014].

2.5: Protein-Nucleic acid Docking

Though protein–RNA and protein–DNA interactions are important cellular processes, but the computational community has paid relatively little attention to protein–nucleic acid docking, and specifically protein–RNA docking. There are few tools such as HADDOCK [Dominguez et al. 2003], GRAMM [Katchalski-Katzir et al. 1992], HEX [Ritchie et al. 2000, Kemp et al. 2000], PatchDock [Schneidman-Duhovny et al. 2005] and FTDock [Gabb et al. 1997] which were originally developed for protein-protein docking and later adapted to accept nucleic acid as an input. Further, the lack of availability of scoring functions to assess protein-RNA interactions compounds the problem. Recent efforts by different scientific groups aim to improve protein-nucleic acid docking systems [Puton et al. 2012, Zheng et al. 2007, Perez-Cano et al. 2010, Huang et al. 2014] (http://genesilico.pl/NPDock).

2.6: Issues in comparative analysis of docking tools

A plethora of docking tools has been developed in the past 20 years and the number of new tools is steadily increasing (**Table 5 and 7**). A comprehensive understanding of the advantages and limitations of each docking program is fundamentally important to conduct more reasonable docking studies and docking-based virtual screening but comparing them is very difficult. This is due to the following factors: First, during our review process we were not able to download or install many of the published tools due to several reasons such as broken hyperlinks, obsolete websites, and issues during installations. Second, only a few studies have been conducted to assess the relative performance of docking algorithms/scoring functions [Bissantz et al. 2000, Charifson et al. 1999, Keseru et al. 2001] and most of these studies focused on the use of only a few methods. Third, scientists have different points of view regarding the performance of the

tools since the examined properties vary in each of these studies (quality of the top-ranked pose, quality of all plausible poses, binding free energy prediction, and virtual screening utility). Fourth, the use of approximations during the docking process may lead to variable results such as inhomogeneous docking paces ranging from a few seconds to a few hours. Last, most of the docking tools have been calibrated and validated on small protein–ligand data sets instead of large datasets [Kramer et al. 1999, Diller et al. 2001, Pang et al. 2001, Paul and Rognan et al. 2002, Verdonk et al. 2003, Nissink et al. 2002] (**Table 5 and 7**) (to predict ligand binding poses) and the scoring functions (to rank the binding affinities). In addition, one must keep in mind the diversity of protein structure/domains and therefore expect variations in performance of docking tools/scoring functions due to the differences between protein families. For example, LeDock performs well for docking studies for eukaryotic proteases and pepsin families but performs poorly for retroviral proteases and phosphate binding proteins.

2.7: Online Web Services for docking

The availability of web-enabled docking servers takes computational load from the user's computer thereby helping scientists particularly those with minimal or no background in computers. Over the years several web servers have been developed to handle different aspects of docking. For example, ZDock [Chen et al. 2003] and PatchDock [Schneidman-Duhovny et al. 2005] were developed to perform rigid-body docking. A program named ClusPro [Comeau et al. 2004, Comeau et al. 2004] can filter, cluster and rank docking solution candidates. SmoothDock (version of ClusPro) [Camacho et al. [2003] refines the representatives of the largest clusters. RosettaDock [Wang et al. 2005] allows search in the vicinity of a single given input solution candidate. The GRAMM-X [Tovchigrechko et al. 2006,Vakser et al. 2006] and Hex [Ritchie et al. 2000, Kemp et al. 2000] web servers perform rigid-body docking followed by an optimization of the rigid-body orientation. There is significant interest in this area as evidenced by a growing list of docking servers currently available, such as Docking Server (http://www.dockingserver.com/web), DockingAtUTMB(https://scsb.utmb.edu/facilities/software/), Pardock (http://www.scfbio-iitd.res.in/dock/pardock.jsp), PatchDock(http://bioinfo3d.cs.tau.ac.il/PatchDock/), MetaDock (http://dock.bioinfo.pl/), PPDock (http://140.112.135.49/ppdock/index.html) and MEDock (http://medock.ee.ncku.edu.tw/), PliP (projects.biotec.tu-dresden.de/plip-web), ClusPro [Comeau

et al. 2004], HADDOCK [de Vries et al. 2012], RosettaDock server [Lyskov et al. 2008], GRAMM-X [Tovchigrech at al. 2006], 3D-Garden [Lesk et al. 2008], HEX server [Macindoe et al. 2010], SwarmDock [Torchala et al. 2013], ZDOCK server [Pierce et al. 2014], PatchDock [Schneidman-Duhovny et al. 2015], ATTRACT[de Vries et al. 2015], pyDockSAXS [Jimenez-Garcia et al. 2015], Inter EvDock and NPDock [Tuszynska et al. 2015].

Further state-of-the-art web tools such as SwissDock [Gabb et al. 1997], GalaxySite [Vakser et al. 1997] or ProBiS [de Vries et al. 2007] can thus be applied in the evaluation of docking results. HexServer is based on Fast Fourier transform (FFT) and takes 15 s for each blind 6D docking calculations (http://hexserver.loria.fr/). It uses two graphics processors simultaneously and demonstrates the ability to produce up to 1,000 docking predictions [Macindo et al. 2010]. This server has played a role in clinical studies such as finding the role of mutations in the NOTCH pathway regulator MIB1 in left ventricular non-compaction cardiomyopathy. It has also played a role in several other studies. For instance, Demchuk et al. used this server to find several potential bindings. The 3D models of FtsZ-ligand complexes generated using the Hex 6.1 server facilitated the identification of benzimidazoles binding sites on FtsZ2-2 protein surface present in *Arabidopsis thaliana*. Paul et al. (2014) also applied HexServer to dock 44 potential inhibitors of oncogenes and transcription factors having anti-cancer properties; in which significant interactions were found in 21 docking cases. The GRAMM-X web server extends original FFT methodology by employing smoothed potentials, refinement stage, and knowledge-based scoring. A full docking protocol for a single complex on an average is completed in 2 minutes, running on 16 2.0 GHz Opteron processors. It is implemented in Python and C++ hence combining the fast prototyping power of Python with the numerical performance of C++ modules [Tovchigrechko et al. 2006]. It is freely accessible at (http://vakser.bioinformatics.ku.edu/resources/gramm/grammx).

Cluspro (https://cluspro.bu.edu/login.php) was the first fully automated, web-based program employed for the prediction of protein structures. Billions of putative complexes can be evaluated through this docking algorithm. A filtering method is executed in a certain number of structures; only those with good electrostatic and desolvation free energies are further selected for clustering. When the algorithm was applied to a benchmark set of 2000 conformations, within the top 30 predictions, at least one experimentally relevant complex was predicted [Comeau et al. 2004]. The output generates a list of complexes that are ranked on the basis of

their clustering properties [Comeau et al. 2005]. The performance of ClusPro suggests that its success rate is around 71% when targets having a significant structural rearrangement upon binding are not included [Comeau et al. 2007]. The new version of ClusPro also incorporates the docking program PIPER, which effectively increases the number of near-native docked structures [Comeau et al. 2007]. SwissDock is freely available at http://www.swissdock.ch. It is dedicated to docking of small molecules on target proteins and uses the EADock DSS engine. Success rates for small and relatively rigid ligands with less than 10 flexible rotatable bonds have been reported by [Grosdidier et al. 2011]. It has been implemented in various studies, one of which involved screening molecules which can act as antibiofilm agents, needed for the purpose of inhibiting *Staphylococcus epidermidis* biofilm production [Al-Khafaji et al. 2014]. The Istar web server, freely available at http://istar.cse.cuhk.edu.hk/idock, provides a key computational method for large scale protein-ligand docking. The website facilitates filtering of ligands on the basis of desired molecular properties, monitoring job progress, and visualization of ligand conformations. Results show that it outperformed AutoDock Vina in terms of docking efficiency. Moreover, the use of Istar requires no manual processing of input proteins in most of the cases [Li et al. 2014]. The PharmMapper web server is another tool used for potential drug target prediction against any given small molecules via a 'reverse' pharmacophore mapping approach. [Xiaofeng Liu et al. 2010] searched the potential drug target proteins for tamoxifen via the PharmMapper server.

2.8: Distinct features of docking software and its performances on different datasets

Existing docking software can be classified based on its search algorithm, scoring function and several other factors. The following section summarizes popular and highly cited software along with its applications in the context of large-scale docking studies (Also see **Table 7**).

2.9: Large Scale Docking

Research projects employing docking to study the interaction at the whole proteome or genome level or using a large number of ligands can be labelled as 'large-scale'. Using this criterion, previously published studies by Gao et al. (~1,100 targets [Gao et al. 2008]), and Hui-fang et al., (1,714 targets and 8 compounds) [Hui-fang et al. 2010], or modeling networks [Szilagyi et al. 2014, Wass et al. 2011, Vakser et al. 2013, Mosca et al. 2013, Wodak et al. 2013, Zhang et al.

2012, Kundrotas et al. 2012, Kar et al. 2012, Kundrotas et al. 2010] may be designated as large-scale docking studies. On similar lines, Lee and Kim in 2012 generated a 2D matrix of docking scores among all the possible protein structures in yeast and humans for 35 well-known drugs. In 2016, our group built an automated docking pipeline to dock orlistat as well as other drugs against the 24,000 proteins in the human structural proteome to explain the therapeutics and side effects at a network level. In **Table 5,** we discuss various examples where tools such as GemDock and PsiDock were used to dock a large number of ligands and protein targets. Conventionally, protein interactions are studied using free docking methods [Zhu et al. 2008, Mosca et al. 2009], or template-based docking methods at larger scale [Zhang et al. 2012, Kundrotas et al. 2012, Kar et al. 2012].

2.10: Limitations of Docking tools: Challenges and Opportunities

In several studies, it was observed that despite obtaining high docking scores or binding affinity (in-silico/in-vitro/in-vivo systems), potent lead for a commercial drug is difficult to find. The reasons are attributed to problems in protein structure [Hoelder et al. 2012], variations in environment of binding site, and variations in pH affecting target proteins in context of human body [Kitchen et al. 2004]. Similarly, many studies have shown poor correlations between docking scores and experimental binding affinities. A study was conducted to improve this correlation by implementing a multipose binding concept in the docking scoring scheme [Atkovska et al. 2014]. In many instances, researchers tend to over-interpret docking results. For example, some authors have claimed a particular ligand as agonist/inhibitor for a target protein only on the basis of docking scores without conducting confirmatory studies [Alves et al. 2007, Chen et al. 2012, Chen 2013, Hong et al. 2012]. McGaughey et al. (2007) have also shown that 2D and 3D ligand similarity-based methods outperform docking tools in virtual screening experiments. Molecular dynamics (MD) simulations can be used to validate docking results since MD is able to check movement of the protein-ligand complex over a period of time. This is important since changes in structure of protein/ligand during perturbation can modify final binding pose [McGaughey et al. 2007]. Although MD provides useful information to complement the docking prediction, very few studies have utilised MD [Cavalli et al. 2004, Park et al. 2004]. The presence of solvent (water) molecules plays an important role such as electrostatic screening [Schutz et al. 2001], catalysis and molecular recognition [Ben et al. 2001,

Bienstock et al. 2015] and well known docking packages incorporate water molecules explicitly to predict protein–ligand docking [Verdonk et al. 2005, Osterberg et al. 2002, Friesner et al. 2006]. But, very few methods exist that allow the prediction of hydration water positions at protein–protein interfaces [Ruyck et al. 2016]. Recently, researchers have developed several methods to incorporate solvation to improve docking predictions such as WaterMap protocol [Yang et al. 2013], SZMAP [Kumar et al. 2013], Ligand Hydration Methods [Forli et al. 2012], and WaterDock [Sridhar et al. 2017]. Besides receptor flexibility, ligand induced large scale conformational changes add a new set of challenges in front of computational biologists. To investigate one such problem, Dietzen et al. used normal mode analysis (NMA) in predicting the conformational changes observed upon small-molecule binding, albeit with limited success. In addition, specific parts of the protein structure such as ionizable residues [Yuriev et al. 2015], and protein pockets were also examined in the past. There have been number of studies which had explored the role of ligand structure (namely ionization and tautomerism parameters) to improve docking predictions. For instance, Natesan et al. (2012), introduced the concept of multi species approach into the QM/MM linear response method and used it for structural correlation of published inhibition data on mitogen-activated protein kinase (MAPK)-activated protein kinase (MK2) by 66 benzothiophene and pyrrolopyridine analogues with reasonable success [Natesan et al. 2012]. Continuing on similar lines, Feher and Williams et al. (2012) examined the sensitivity of docking programs to small changes in input files of ligands. They also demonstrated that part of the docking variation is due to numerical sensitivity and potentially chaotic effects in current docking algorithms and not solely due to incomplete ligand conformation and pose searching [Feher et al. 2012, Williams et al. 2012].

During docking, ligand flexibility is a major reason for failure of docking protocols to correctly predict the pose. Bohari and Sastry (2012) recommended that docking protocols perform optimally when a balanced type of hydrophilic and hydrophobic interaction or dominant hydrophilic interaction exists. Similarly, by using more than one docking program to predict the binding pose, correct poses were identified more accurately and there appears to be a certain ligand size that maximizes pose prediction accuracy because of optimum flexibility. In order to circumvent these issues, tools/methods such as S4MPLE have been designed [Beato et al. 2013, Des Jarlais et al. 1986].

Scoring functions and sampling are often criticised in the event of poor performance of docking programs. A study by [Greenidge et al. 2014] demonstrated that identification of the correct pose (docking power) can be improved by incorporating ligand strain into the scoring function or rescoring an ensemble of diverse docking poses with MM-GBSA in a post processing step.

2.11: Binding site prediction, Interaction hotspots and docking

Information on binding site (BS) on target protein plays an important role in obtaining accurate docking results. BS can be classified as following: (i) Lipophilic buried cavities such as COX-2 and estrogen receptor, (ii) binding sites of intermediate polarity with hydrogen bonding motifs common to the majority of inhibitors such as p38 MAP kinase, gyrase B and thrombin and (iii) which are very polar, solvent-exposed binding sites seen in neuraminidase and gelatinase A. [Schulz-Gasch et al. 2003] have described a set of guidelines for virtual screening/docking system based upon their results for the benefits of the users. Advances in technologies are also contributing towards our improved understanding of role of binding or interaction sites [Nero et al. 2014, Kahraman et al. 2013]. In a comprehensive cross-docking study, Lopes et al. (2003), docked over 300,000 conformations per protein pair for the set of 28,224 possible pairs (168 proteins of the Mintseris Benchmark 2.0) [Lopes et al. 2013]. From a docking point of view, Protein-Protein Interaction (PPIs) is in principle similar to traditional drug targets and was shown to be amenable to docking [Koes et al. 2012]. Therefore, docking methods are used in several stages during the design of PPI inhibitors as well as in finding interaction hotspots [Sable et al. 2015].

2.12: Inverse/Reverse docking systems

Chen and Zhi introduced this term in 2001 for finding potential protein targets of a small molecule by the computer automated docking search of a protein cavity database. Subsequently, reverse docking was used in a number of important investigations: (A) the virtual target screening method calibrating a set of small molecules against a protein library [Sung et al. 2012], and (B) the activity prediction of 656 marketed drugs on 73 unintended “side effect” targets [Lounkine et al. 2012], (**See Table 5**).

2.13: Ensemble based Docking

In a recent work, Kim et al. developed a new program named as ALIS-DOCK (Automated pLatform for Integrative Structure-based DOCKing) for automated structure based virtual screening (SBVS) to identify inhibitor against Heat shock protein 90 (Hsp90) [Kim et al. 2018]. Authors employed ensemble-based docking strategies in which multiple input receptor conformations are fed into docking programs followed by experimental verification studies. Ensemble-based methods are considered to be better than a single receptor conformation input [Sinko et al. 2013]. However, several drawbacks limit ensemble-based docking which includes the lack of a protocol to generate ensembles, in terms of both size and membership [Yuriev et al. 2011, Yuriev et al. 2013, Korb et al. 2012]. Rueda et al. (2012) dealt with this problem by introducing a method based on exhaustive combinatorial searching and individual addition of pockets, selecting only those that maximize the discrimination of known active compounds from decoys. To address these problems, Xu and Lill (2012) combined experimental knowledge with different computational methods to reduce the ensemble of protein structures to increase efficiency and enrichment quality. Apart from the above mentioned studies, several other techniques have been introduced to address the issues of receptor flexibility.

2.14: Fragment based docking

In order to design efficient drugs, fragment-based drug design (FBDD) was proposed in 1996. FBDD focuses to find molecules/fragments having low in molecular-weight and chemical complexity to target sub-pockets in active site. The approach is inspired from the divide and conquer algorithm and the fragments serve as starting points for “growing” the lead candidate. Though various computational methods have been developed for FBDD, molecular docking remains an attractive way to prioritize fragments from much larger commercially available data sets. Several factors such as non-optimised scoring functions, affect the accuracy of fragment-based docking results. Programs such as LUDI, GLIDE, LigBuilder, and S4MPLE are powerful enough to place fragments into the correct pocket of the active site. In 2016, Hao et al. developed a web-based server dedicated for FBDD [Hao et al. 2012]. Apart from these, FBDD and docking continue to be major strategy to discover new lead molecules and efforts are being made to improve FBDD. MM-PBSA rescoring; [Kawatkar et al. 2012, Zhu et al. 2013] a combination of structure-based and ligand-based screening; [Cortes-Cabrera et al. 2012], protein mapping with

FTMap; [Hall et al. 2012], templating of fragment ligands on known structures; [Tosh et al. 2012], and GPU-accelerated MD [Zhu et al. 2013].

2.15: Benchmarking datasets/studies and evaluation of docking tools

Evaluating docking tools is challenging since we are dealing with a system which is highly non-linear and multi-dimensional which treats proteins as a rigid entity. Further docking methods are strongly dependent on choices of input preparation that vary between different practitioners. The next screening process is also biased since it involves a highly skewed population of actives (very few) versus inactives (very many) coupled with an operational cost function that varies from user to user [Jain et al. 2008]. Moreover, problems in dataset sharing, bias in datasets, variations in sample sizes, enrichment issues and statistical measures for reporting- are major factors to be considered when designing any evaluation or benchmarking study.

The benchmarking of docking tools started as early as 1990 when one research group docked 103 ligands against chymotrypsin using the DOCK tool and found that the computational predictions matched the experimental data [Stewart et al. 1990]. In 2004, eight docking programs (DOCK, FLEXX, FRED, GLIDE, GOLD, SLIDE, SURFLEX, and QXP) were compared to recover the X-ray pose of 100 small-molecular-weight ligands, and for their capacity to discriminate known inhibitors of an enzyme (thymidine kinase) from randomly chosen “drug-like” molecules (Kellenberger). Continuing on similar lines, Huang et al. constructed the directory of useful decoys (DUD), with 2,950 ligands for 40 different targets leading to a database of 98,266 compounds [Huang et al. 2006]. This is an important resource for the evaluation of docking tools (http://blaster.docking.org/dud/). The same research group generated another resource, DUD-E, which includes more diverse targets such as GPCRs and ion channels, totalling 102 proteins with 22886 clustered ligands drawn fromChEMBL, each with 50 property-matched decoys drawn from ZINC. In 2010, Plewczynski et al. conducted first large-scale evaluation of seven popular docking tools on the extensive dataset composed of 1300 protein–ligands complexes from PDBbind 2007 database, where experimentally measured binding affinity values were also available. In another study, Bohari and Sastry (2012) evaluated the performances of five popular docking protocols, (Glide, Gold, FlexX, Cdocker and LigandFit) on 199 FDA approved drugs and declared Glide and Cdocker as top ranking tools [Plewczynski et al. 2010, Bohari et al. 2012]. One of the research teams evaluated a panel of 20 scoring

functions in terms of "scoring power" (binding affinity prediction), "ranking power" (relative ranking prediction), "docking power" (binding pose prediction), and "screening power" (discrimination of true binders from random molecules) [Li et al. 2014]. Wang et al. (2016) found that academic programs performed better than commercially available docking tools. We have compiled a list of benchmarking and evaluation studies for the benefits of the users of docking tools [Wang et al. 2016].

3. APPLICATIONS OF DOCKING

3.1: Drug repositioning (repurposing) using molecular docking

Drug repositioning is finding new uses for existing drugs and offers several advantages such as reducing time efforts, expenses and failures typically associated with the drug discovery process. Scientists have devised several strategies for repositioning which includes the use of transcriptional signatures [Lamb et al. 2006, Chang et al. 2010, Iskar et al. 2013], networks [Hu et al. 2012, Agarwal et al. 2009, Jin et al. 2012], ligand based approaches [Brown et al. 2017, Patel et al.2017, Shameer et al. 2017, Keiser et al. 2009, Liu et al. 2010, Vasudevan et al. 2012, Sawada et al. 2015], ligand based chemigenomics and machine learning approaches [Mestres et al. 2006, Bender et al. 2007, Gregori-Puigjané et al. 2008, Mestres et al. 2008, Bender et al. 2007]. [Unterthiner et al. 2014, Alaimo et al. 2016], structure-based approaches [Ehrt et al. 2016 ,Zhang et al. 2004, Jalencas et al. 2013, Mestres et al. 2013, Anighoro et al. 2015], and molecular docking [Kinnings et al. 2009, Li et al. 2011, Dakshanamurthy et al. 2012]. Li et al. (2011) used docking methods on drugs of the DrugBank database and 35 crystal structures of MAPK14. The study identified the chronic myeloid leukemia drug nilotinib as a potential anti-inflammatory drug with an in vitro IC50 of 40 nM [Li et al. 2011]. Dakshanamurthy et al. (2012) successfully tested an anti-parasitic drug as an anti-angiogenic Vascular Endothelial Growth Factor Receptor 2 (VEGFR2) inhibitor, and a new connection was discovered between previously untargeted Cadherin-11, implied in rheumatoid arthritis, and cyclooxygenase-2 (COX-2) inhibitor celecoxib. We have compiled several research studies which used molecular docking tools for repositioning purposes.

3.2: Side effect prediction using docking

The docking technique plays an important role in predicting effects; docking-based tools have predicted the efficacy of potential therapeutic compounds and have also helped in predicting the range of unintended and undesired interactions between the specific compound and the human proteome. Using docking studies in combination with pharmacophore modeling, novel benzodiazepine (binding site) agonists in GABA receptors were designed, examined and compared with existing agonists [Sieghart et al. 2006]. These analyses have been used for finding comparative side effects of individual drugs against the same disease. Docking studies and a subsequent analysis has enabled us to find the probable off-target receptors in certain pockets which had a higher affinity for one drug; this was demonstrated in a study where Sunitinib co-existed more frequently than Sorafenib with respect to the hypothyroidism events [Venkatapathy et al. 2004].

Moreover, docking methods on adverse reactions on enzymes have also been used for quite some time [Drwal et al. 2005, Malgorzata et al. 2005]. Using pharmacophore pre-alignment and QSAR models along with flexible docking techniques to quantify the binding affinity, adverse reactions were predicted for a certain drug [Devillers et al. 2010]. It was reported that SolB (Schisandrol B) has a protective effect against APAP overdose induced acute liver failure. While the same was checked in mice, docking studies confirmed the binding of SolB with the residue through inhibiting their activities [Jiang et al. 2014]. Drug modelling for gout also used docking to devise compounds which are expected to report fewer side effects than the previous drugs used [Moon Ho et al. 2012].

A study by LaBute et al. (2014) also depicted the use of molecular docking for high throughput screening of drug molecules and for prediction of ADRs. Based on the docking score of 506 compounds out of 906 small molecule drugs docked against 409 protein targets from DrugBank via Autodock (Vina LC), a logistic regression model predicted 85 side-effects. The validation of ADR prediction modes is based upon docking score and is carried out by comparing AUCs/area-under-the-receiver-operating-characteristic-curves (AUCs) with experimentally derived drug-protein interactions [Liu et al. 2010].

Additionally, inverse docking has also been believed to lead to the identification of the proteins which the specific molecule has a likelihood of acting on, leading to a predictive analysis of the potential ADRs the drug molecule could cause [Gfeller et al. 2014]. Grinter et al. (2011) used the

docking software MDock to perform an inverse docking study to identify potential targets of PRIMA-1, to investigate its ability to cause apoptosis in cancer cells [Grinter et al. 2011].

3.3: Docking and Experimental studies

Apart from drug repositioning and side effect prediction, docking has also been used as an intermediate step in the search for finding new drugs in conjunction with time-consuming experimental high-throughput screening. Due to the use of virtual screening and docking, researchers are able to save time and efforts for screening new drugs. Docking, being a part of virtual screening has been used as this initial step in a number of studies. In this section, we discuss studies where docking is integrated with experimental system (in-vivo or in-vitro) to confirm the predictions. These studies majorly focussed on discovery of new inhibitors for targets drawn from infectious agents which include *Mycobacterium tuberculosis*, *Bacillus anthracis*, *Vibrio harveyi*, HIV, vaccinia, variola and monkey-pox viruses. Apart from that, in a number of studies, docking was used in conjunction with wet-lab experiments for finding new drugs/treatment modalities for metabolic and non-communicable disorders such as diabetes, cancer, obesity and allergies (**Table 6**). Recently, structure-guided design [Cobb et al. 2015] and virtual screening [Chaudhary et al. 2014] were successfully applied in order to identify and evaluate new molecules with a potent inhibitory effect on *Plasmodium falciparum*.

3.4: Docking in Immunoinformatics

Zhang et al. (2013) used docking for epitope prediction methods in combination with 3D structural modeling of peptide-MHC-TCR complex to identify MHC class I restricted T-cell epitopes for use in epitope-based vaccines like HIV and human cancers [Zhang et al. 2013]. In another collaborative study by Indian-UK based researchers worked on Crimean–Congo hemorrhagic fever virus (CCHFV) to predict epitopes which can be helpful for vaccine designing [Papa et al. 2002]. Krawczyk et al. developed a new method which combines conformational matching of the antibody-antigen structures and a specific antibody-antigen score [Krawczyk et al. 2014, Konrad et al. 2014]. Recently in 2018, researchers described the use of an incremental meta-docking approach for structural prediction of pMHC complexes to overcome challenges faced by previous methods [Antunes et al. 2018]. This study is important since it addressed major limitations of docking approaches since docking methods are known to be much less reliable when applied to larger ligands (e.g., ligands with more than 10 internal DoFs) [Chang et

al. 2010, Michel et al. 2010]. For instance, peptides are known to be very flexible ligands [Devaurs et al. 2015]; binding mode prediction of even small peptides, composed of up to 5 amino acids (which means around 24 internal DoFs), can be particularly challenging for available docking method [Rentzsch et al. 2015, Wang et al. 2016].

In the vaccine design domain, docking is being increasingly used to find novel candidates. For example, Alam et al. (2007) docked two predicted epitopes to HLA-A*53:01 with Autodock and reported good predicted binding affinities for the peptides [Mirza et al. 2016]. In another study by Mirza et al. (2016) investigated the binding interactions of CTL epitopes with three class I major histocompatibility complex (MHC I) proteins after docking the peptides to the binding groove of the MHC I proteins.

Recently this approach is being used to target pathogens responsible for neglected tropical diseases (NTDs) in order to develop innovative "anti-poverty" vaccines [Hotez, 2018]. Studies by Khatoon et al. [2017, 2018], used an immunoinformatics approach to evaluate both membrane and secretory proteins of *Leishmania donovani* followed by molecular docking and dynamics to evaluate the binding affinity and stability of receptor (TLR-4) and ligand (vaccine protein) complex. Recently, our group has started working on a collaborative project to identify new vaccine candidates for Chagas Disease, a poverty related NTD in the Americas (Beaumier et al 2016, Jones et al 2018). This approach is focused on augmenting host immunity to improve on current chemotherapeutic approaches, and proposes combining text mining, machine learning, network sciences and immunoinformatics approaches to build multi-layered network of *Trypanosoma cruzi* and host to obtain comprehensive understanding of molecular pathophysiology of Chagas Disease (Jagannadham et al. 2016). In our platform, we shall use the docking systems for investigation of binding interactions of CTL epitopes with MHC proteins (https://sites.google.com/view/vaccinepipeline/). The hope is that this approach might accelerate the discovery, development and testing of NTD anti-poverty vaccines.

3.5: Use of Automation, Cloud, Parallel and Distributed Computing in Docking

Pharmaceutical companies value workflows and pipelines which integrate various steps of docking or virtual screening process. Taking these cues, Therrien et al. (2014) built a web enabled system for drug discovery system which implements steps such as ligand molecule processing, macromolecule preparation for docking, and docking with Flexibility Induced

682 through targeted Evolutionary Description (FITTED) method. Docking methods when used in
683 VS workflow suffer from bottleneck due to lack of computational capabilities. Advancements in
684 computational field particularly in cloud computing, parallel and distributed computing can
685 alleviate such problems [Yuriev et al. 2015, Dong et al. 2015]. Servers such as iSCREEN and
686 MTiOpenScreenv are also good example of cloud-based web implementation of docking tools.
687

688 **4. DISCUSSIONS AND FUTURE DIRECTIONS**

689 A cursory look across the wide range of studies we reviewed reveals that docking is a powerful
690 tool, engendering many success stories in drug discovery process as well as side effect
691 prediction. It complements the experimental approaches or can even be used to find novel
692 unknown targets. The field is quickly advancing and expanding its practical applications due to
693 the continuous increases in computational power. Making docking services available online, thus
694 letting external servers do the computing, and allowing the user to visualize and obtain the
695 docking results. However, there is still a necessity to resolve certain issues such as construction
696 of datasets of target structure, computational efficiency, the inclusion of receptor flexibility,
697 improved search algorithm and scoring function accuracy for explicit target identification. More
698 importantly, normalization of docking scores is necessary in order for it to be a truly successful
699 tool. A recent study suggested role of machine learning in combining multiple docking tools as
700 well as scoring functions to improve performance [Hsin et al. 2013]. There is lot of interest in the
701 application of machine learning techniques in virtual screening and computational docking as
702 evident by huge number of publications in recent years. The effort needs to be concentrated in
703 these areas so that more intriguing applications can be uncovered in the future.
704

705 **MATERIALS AND METHODS**

706 We have created a new technology assisted review system which incorporates support vector
707 machines, information retrieval programs, web based forms and programs built in Perl and
708 Python (Jagannadham et al. 2016; Cormack et al. 2015) (Figure 1). The system consists of
709 automated paper writing module and automated review module. We searched literature resources
710 such as PubMed and Google Scholar with queries such as “Molecular Docking”, “Docking”, and
711 “Docking tools” to retrieve abstracts & full length articles. The manual screening was conducted
712 by the three independent teams comprising trained researchers.

ACKNOWLEDGEMENTS

We thank Yogendar Thakur for his assistance. We thank Department of Biotechnology, Government of India, for their financial support (Grant Id: BT/PR17252/BID/7/708/2016). We also thank the Robert J. Kleberg, Jr. and Helen C. Kleberg Foundation, USA and Baylor College of Medicine, Houston USA for grant support. We also thank our volunteers (B.Tech/M. Tech) for the assistance provided for development of website associated with this manuscript as well as during data collection phase of this manuscript.

REFERENCES

1. Adolf-Bryfogle, J., and Dunbrack Jr, R. L. (2013). The PyRosetta Toolkit: a graphical user interface for the Rosetta software suite. *PloS one*, *8*(7), e66856.

2. Alaimo, S., et al. (2016). Recommendation Techniques for Drug–Target Interaction Prediction and Drug Repositioning. In *Data Mining Techniques for the Life Sciences* (pp. 441-462). Humana Press, New York, NY.

3. Alam, S. M., et al. (2007). The role of antibody polyspecificity and lipid reactivity in binding of broadly neutralizing anti-HIV-1 envelope human monoclonal antibodies 2F5 and 4E10 to glycoprotein 41 membrane proximal envelope epitopes. *The Journal of Immunology* 178.7:4424-4435.

4. Al-Khafaji, Z. M., and Al-Mulla, A. F. (2014). In Silico design of inhibitors for *Staphylococcus epidermidis* biofilm. *Int J Adv Pharm BiolChem*, *3*, 334-340.

5. Alonso, H., et al. (2006). Combining docking and molecular dynamic simulations in drug design. *Medicinal research reviews*, *26*(5), 531-568.

6. Aloy, P., et al. (2003). The relationship between sequence and interaction divergence in proteins. *Journal of molecular biology*, *332*(5), 989-998.

7. Alves, C. N., et al. (2007). Calculation of binding energy using BLYP/MM for the HIV-1 integrase complexed with the S-1360 and two analogues. *Bioorganic and medicinal chemistry*, *15*(11), 3818-3824.

740 8. Anighoro, A., et al. (2015). Computational polypharmacology analysis of the heat shock
741 protein 90 interactome. *Journal of chemical information and modeling*, *55*(3), 676-686.

742 9. Anishchenko, I., et al. (2014). Protein models: The Grand Challenge of protein docking.
743 *Proteins: Structure, Function, and Bioinformatics*, *82*(2), 278-287.

744 10. Annamala, M. K., et al. (2007). Docking of phosphonate and trehaloseanalog inhibitors into
745 *M. Tuberculosis* mycolyltransferase Ag85C: Comparison of the two scoring fitness
746 functions GoldScore and ChemScore, in the GOLD software. *Bioinformation*, *1*(9), 339.

747 11. Antunes, D. A., et al. (2018). General Prediction of Peptide-MHC Binding Modes Using
748 Incremental Docking: A Proof of Concept. *Scientific reports*, *8*(1), 4327.

749 12. Atkovska, K., et al. (2014). Multipose binding in molecular docking. *International journal*
750 *of molecular sciences*, *15*(2), 2622-2645.

751 13. Baig, M. H., et al. (2018). Computer aided drug design and its application to the
752 development of potential drugs for neurodegenerative disorders. *Current*
753 *neuropharmacology*, *16*(6), 740-748.

754 14. Beato, C., et al. (2013). Use of experimental design to optimize docking performance: The
755 case of ligendock, the docking module of ligen, a new de novo design program.

756 15. Beglov, D., et al. (2012). Minimal ensembles of side chain conformers for modeling
757 protein–protein interactions. *Proteins: Structure, Function, and Bioinformatics*, *80*(2), 591-
758 601.

759 16. Bender, A., et al. (2007). Analysis of pharmacology data and the prediction of adverse drug
760 reactions and off-target effects from chemical structure. *ChemMedChem: Chemistry*
761 *Enabling Drug Discovery*, *2*(6), 861-873.

762 17. Ben-Naim, A. (2002). Molecular recognition—viewed through the eyes of the solvent.
763 *Biophysical chemistry*, *101*, 309-319.

764 18. Beutler, T. C., et al. (1994). Avoiding singularities and numerical instabilities in free energy
765 calculations based on molecular simulations. *Chemical physics letters*, *222*(6), 529-539.

766 19. Billeter, M., Havel, T. F., and Kuntz, I. D. (1987). A new approach to the problem of
767 docking two molecules: the ellipsoid algorithm. *Biopolymers: Original Research on*
768 *Biomolecules*, *26*(6), 777-793.

769 20. Bissantz, C., et al. (2000). Protein-based virtual screening of chemical databases. 1.
770 Evaluation of different docking/scoring combinations. *Journal of medicinal chemistry*,
771 *43*(25), 4759-4767.

772 21. Block, P., et al. (2006). AffinDB: a freely accessible database of affinities for protein–ligand
773 complexes from the PDB. *Nucleic acids research*, *34*(suppl_1), D522-D526.

774 22. Blomme, E. A., and Will, Y. (2015). Toxicology strategies for drug discovery: present and
775 future. *Chemical research in toxicology*, *29*(4), 473-504.

776 23. Bohari, M. H., and Sastry, G. N. (2012). FDA approved drugs complexed to their targets:
777 evaluating pose prediction accuracy of docking protocols. *Journal of molecular modeling*,
778 *18*(9), 4263-4274.

779 24. Brown, A. S., and Patel, C. J. (2017). A standard database for drug repositioning. *Scientific*
780 *data*, *4*, 170029.

781 25. Brunsteiner, M., and Petukhov, P. A. (2012). Insights from comprehensive multiple receptor
782 docking to HDAC8. *Journal of molecular modeling*, *18*(8), 3927-3939.

783 26. Brylinski, M., and Skolnick, J. (2008). A threading-based method (FINDSITE) for ligand-
784 binding site prediction and functional annotation. *Proceedings of the National Academy of*
785 *sciences*, *105*(1), 129-134.

786 27. Bullock, et al. (2013). DockoMatic 2.0: high throughput inverse virtual screening and
787 homology modeling.

788 28. Bursulaya, B. D., et al. (2003). Comparative study of several algorithms for flexible ligand
789 docking. *Journal of computer-aided molecular design*, *17*(11), 755-763.

790 29. Buturak, B., et al. (2014). Designing of multi-targeted molecules using combination of
791 molecular screening and in silico drug cardiotoxicity prediction approaches. *Journal of*
792 *Molecular Graphics and Modelling*, *50*, 16-34.

30. Byrd, C. M., et al. (2002). The vaccinia virus I7L gene product is the core protein proteinase. *Journal of virology*, *76*(17), 8973-8976.

31. Cabrera, A. C., et al. (2011). VSDMIP 1.5: an automated structure-and ligand-based virtual screening platform with a PyMOL graphical user interface. *Journal of computer-aided molecular design*, *25*(9), 813.

32. Camacho, C. J., and Gatchell, D. W. (2003). Successful discrimination of protein interactions. *Proteins: Structure, Function, and Bioinformatics*, *52*(1), 92-97.

33. Camacho, C. J., and Vajda, S. (2002). Protein–protein association kinetics and protein docking. *Current opinion in structural biology*, *12*(1), 36-40.

34. Cavalli, A., et al. (2004). A computational study of the binding of propidium to the peripheral anionic site of human acetylcholinesterase. *Journal of medicinal chemistry*, *47*(16), 3991-3999.

35. Cerqueira, N. M. F. S. A., et al (2009). MADAMM: a multi-staged docking with an automated molecular modelling protocol. *Proteins: Structure, Function, and Bioinformatics*, *74*(1), 192-206.

36. Chan, A. H., et al. (2013). Discovery of *Staphylococcus aureus* Sortase-A Inhibitors Using Virtual Screening and the Relaxed Complex Scheme. *Chemical biology and drug design*, *82*(4), 418-428.

37. Chang, M. W., et al. (2010). Virtual screening for HIV protease inhibitors: a comparison of AutoDock 4 and Vina. *PloS one*, *5*(8), e11955.

38. Chang, M., et al. (2010). Evaluation of phenoxybenzamine in the CFA model of pain following gene expression studies and connectivity mapping. *Molecular pain*, *6*(1), 56.

39. Charifson, P. S., et al. (1999). Consensus scoring: A method for obtaining improved hit rates from docking databases of three-dimensional structures into proteins. *Journal of medicinal chemistry*, *42*(25), 5100-5109.

40. Chaudhary, K. K., and Prasad, C. S. (2014). Virtual Screening of compounds to 1-deoxy-D-xylulose 5-phosphate reductoisomerase (DXR) from *Plasmodium falciparum*. *Bioinformation*, *10*(6), 358.

41. Chen, F., et al. (2017). Application of reverse docking for target prediction of marine compounds with anti-tumor activity. *Journal of Molecular Graphics and Modelling*, *77*, 372-377.

42. Chen, J., et al. (2005). ChemDB: a public database of small molecules and related chemoinformatics resources. *Bioinformatics*, *21*(22), 4133-4139.

43. Chen, P. Y., et al. (2012). Computational analysis of novel drugs designed for use as acetylcholinesterase inhibitors and histamine H3 receptor antagonists for Alzheimer's disease by docking, scoring and de novo evolution. *Molecular medicine reports*, *5*(4), 1043-1048.

44. Chen, P. Y., et al. (2013). Old Drug for New Use: Searching for MEK1 (Mitogen-Activated Protein Kinase Kinase 1) Inhibitor by the Computer Aided Drug Design. *Journal of Life Sciences*, *7*(5), 453.

45. Chen, R., et al. (2003). ZDOCK: an initial-stage protein-docking algorithm. *Proteins: Structure, Function, and Bioinformatics*, *52*(1), 80-87.

46. Chen, Y. C. (2015). Beware of docking! *Trends in pharmacological sciences*, *36*(2), 78-95.

47. Chhabra, S., et al. (2007). RETRACTED: A computational docking study for prediction of binding mode of diospyrin and derivatives: Inhibitors of human and leishmanial DNA topoisomerase-I.

48. Choi, V. (2005). Yucca: An Efficient Algorithm for Small-Molecule Docking. *Chemistry and biodiversity*, *2*(11), 1517-1524.

49. Chung, J. Y., et al. (2011). A python-based docking program utilizing a receptor bound ligand shape: PythDock. *Archives of pharmacal research*, *34*(9), 1451.

50. Cobb, R. E., et al. (2015). Structure-guided design and biosynthesis of a novel FR-900098 analogue as a potent *Plasmodium falciparum* 1-deoxy-D-xylulose-5-phosphate reductoisomerase (Dxr) inhibitor. *Chemical Communications*, *51*(13), 2526-2528.

51. Comeau, S. R., et al. (2004). ClusPro: a fully automated algorithm for protein–protein docking. *Nucleic acids research*, *32*(suppl_2), W96-W99.

52. Comeau, S. R., et al. (2004). ClusPro: an automated docking and discrimination method for the prediction of protein complexes. *Bioinformatics*, *20*(1), 45-50.

53. Comeau, S. R., et al. (2005). Performance of the first protein docking server ClusPro in CAPRI rounds 3–5. *PROTEINS: Structure, Function, and Bioinformatics*, *60*(2), 239-244.

54. Comeau, S. R., et al. (2007). ClusPro: performance in CAPRI rounds 6–11 and the new server. *Proteins: Structure, Function, and Bioinformatics*, *69*(4), 781-785.

55. Compton, Corey L., et al. (2013). "Antibacterial activity of and resistance to small molecule inhibitors of the ClpP peptidase." *ACS chemical biology* 8.12: 2669-2677.

56. Cormack, G. V., & Grossman, M. R. (2015). Autonomy and reliability of continuous active learning for technology-assisted review. *arXiv preprint arXiv:1504.06868*.

57. Cortés-Cabrera, et al. (2012). A reverse combination of structure-based and ligand-based strategies for virtual screening. *Journal of computer-aided molecular design*, *26*(3), 319-327.

58. Cross, J. B., et al. (2009). Comparison of several molecular docking programs: pose prediction and virtual screening accuracy. *Journal of chemical information and modeling*, *49*(6), 1455-1474.

59. Cutullé, C., et al. (2009). Population structure of Australian isolates of the cattle tick Rhipicephalus (Boophilus) microplus. *Veterinary parasitology*, *161*(3-4), 283-291.

60. Dakshanamurthy, S., et al. (2012). Predicting new indications for approved drugs using a proteochemometric method. *Journal of medicinal chemistry*, *55*(15), 6832-6848.

61. Dallakyan, S., and Olson, A. J. (2015). Small-molecule library screening by docking with PyRx. In *Chemical Biology* (pp. 243-250). Humana Press, New York, NY.

62. Damm-Ganamet, K. L., et al. (2013). CSAR benchmark exercise 2011–2012: evaluation of results from docking and relative ranking of blinded congeneric series. *Journal of chemical information and modeling*, *53*(8), 1853-1870.

63. Dave, J. A., et al. (2002). Mycosin-1, a subtilisin-like serine protease of *Mycobacterium tuberculosis*, is cell wall-associated and expressed during infection of macrophages. *BMC microbiology*, *2*(1), 30.

64. de Ruyck, J., et al. (2016). Molecular docking as a popular tool in drug design, an in silico travel. *Advances and applications in bioinformatics and chemistry: AABC*, *9*, 1.

65. de Vries, S. J., et al. (2007). HADDOCK versus HADDOCK: new features and performance of HADDOCK 2.0 on the CAPR In targets. *Proteins: structure, function, and bioinformatics*, *69*(4), 726-733.

66. de Vries, S. J., et al. (2010). The HADDOCK web server for data-driven biomolecular docking. *Nature protocols*, *5*(5), 883.

67. de Vries, S. J., et al. (2015). A web interface for easy flexible protein-protein docking with ATTRACT. *Biophysical journal*, *108*(3), 462-465.

68. DeLuca, S., et al. (2015). Fully flexible docking of medium sized ligand libraries with RosettaLigand. *PLOS one*, *10*(7), e0132508.

69. DesJarlais, R. L., et al. (1986). Docking flexible ligands to macromolecular receptors by molecular shape. *Journal of medicinal chemistry*, *29*(11), 2149-2153.

70. Devaurs, D., et al. (2015). Characterizing energy landscapes of peptides using a combination of stochastic algorithms. *IEEE transactions on nanobioscience*, *14*(5), 545-552..

71. Devillers, J., and Mombelli, E. (2010). Evaluation of the OECD QSAR Application Toolbox and Toxtree for estimating the mutagenicity of chemicals. Part 1. Aromatic amines. *SAR and QSAR in Environmental Research*, *21*(7-8), 753-769.

72. Dhanik, A., McMurray, J. S., and Kavraki, L. E. (2013). DINC: a new AutoDock-based protocol for docking large ligands. *BMC structural biology*, *13*(1), S11.

73. Diller, D. J., and Merz Jr, K. M. (2001). High throughput docking for library design and library prioritization. *Proteins: Structure, Function, and Bioinformatics*, *43*(2), 113-124.

74. Do, M. H., et al. (2012). Physiochemical and quantitative analysis for the standardization of K3 preparation. *Journal of Medicinal Plants Research*, *6*(26), 4347-4353.

75. Dominguez, C., et al. (2003). HADDOCK: a protein− protein docking approach based on biochemical or biophysical information. *Journal of the American Chemical Society*, *125*(7), 1731-1737.

76. Dong, Y., et al. (2017). A Hadoop-based Molecular Docking System. In *Journal of Physics: Conference Series*, *910*(1), p. 012-025. IOP Publishing.

77. Douguet, D., et al. (2006). Dockground resource for studying protein–protein interfaces. *Bioinformatics*, *22*(21), 2612-2618.

78. Drwal, M. N., et al. (2014). Application of in silico methods for predictions of different toxicological endpoints.

79. Dunbar Jr., J. B., et al (2011). CSAR benchmark exercise of 2010: selection of the protein–ligand complexes. *Journal of chemical information and modeling*, *51*(9), 2036-2046.

80. Ehrt, C., et al. (2016). Impact of binding site comparisons on medicinal chemistry and rational molecular design. *Journal of medicinal chemistry*, *59*(9), 4121-4151.

81. Ekins, S., et al. (2007). In silico pharmacology for drug discovery: methods for virtual ligand screening and profiling. *British journal of pharmacology*, *152*(1), 9-20.

82. Elokely, K. M., and Doerksen, R. J. (2013). Docking challenge: protein sampling and molecular docking performance. *Journal of chemical information and modeling*, *53*(8), 1934-1945.

83. Engel, S., et al. (2008). A virtual screen for diverse ligands: discovery of selective G protein-coupled receptor antagonists. *Journal of the American Chemical Society*, *130*(15), 5115-5123.

84. Ewing, T. J., et al. (2001). DOCK 4.0: search strategies for automated molecular docking of flexible molecule databases. *Journal of computer-aided molecular design*, *15*(5), 411-428.

85. Feher, M. (2006). Consensus scoring for protein–ligand interactions. *Drug discovery today*, *11*(9-10), 421-428.

86. Feher, M., and Williams, C. I. (2012). Numerical errors and chaotic behavior in docking simulations. *Journal of chemical information and modeling*, *52*(3), 724-738.

87. Ferrari, A. M., et al. (2004). Soft docking and multiple receptor conformations in virtual screening. *Journal of medicinal chemistry*, *47*(21), 5076-5084.

88. Fogel, G. B., et al (2008). Modeling the inhibition of quadruple mutant *Plasmodium falciparum* dihydrofolate reductase by pyrimethamine derivatives. *Journal of computer-aided molecular design*, *22*(1), 29-38.

89. Forli, S., and Olson, A. J. (2012). A force field with discrete displaceable waters and desolvation entropy for hydrated ligand docking. *Journal of medicinal chemistry*, *55*(2), 623-638.

90. Francis, S. M., et al. (2008). Design of Benzene-1, 2-diamines as selective inducible nitric oxide synthase inhibitors: a combined de novo design and docking analysis. *Journal of molecular modeling*, *14*(3), 215-224.

91. Friesner, R. A., et al. (2006). Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein− ligand complexes. *Journal of medicinal chemistry*, *49*(21), 6177-6196.

92. Fu, Y., Wu, X., et al. (2015). A new approach for flexible molecular docking based on swarm intelligence. *Mathematical Problems in Engineering*, *2015*.

93. Gabb, H. A., et al. (1997). Modelling protein docking using shape complementarity, electrostatics and biochemical information. *Journal of molecular biology*, *272*(1), 106-120.

94. Gao, Y., et al. (2007). DOCKGROUND system of databases for protein recognition studies: Unbound structures for docking. *Proteins: Structure, Function, and Bioinformatics*, *69*(4), 845-851.

95. Gao, Z., et al. (2008). PDTD: a web-accessible protein database for drug target identification. *BMC bioinformatics*, *9*(1), 104.

96. Gfeller, D., et al. (2014). SwissTargetPrediction: a web server for target prediction of bioactive small molecules. *Nucleic acids research*, *42*(W1), W32-W38.

97. Ghoorah, A. W., et al. (2011). Spatial clustering of protein binding sites for template based protein docking. *Bioinformatics*, *27*(20), 2820-2827.

98. Gilpin, B. J., et al. (1998). A novel, secreted form of human ADAM 12 (meltrin α) provokes myogenesis in vivo. *Journal of Biological Chemistry*, *273*(1), 157-166.

99. Gohlke, H., et al. (2000). Knowledge-based scoring function to predict protein-ligand interactions. *Journal of molecular biology*, *295*(2), 337-356.

100. Gomez, M. S., et al. (1997). Substrate and cofactor specificity and selective inhibition of lactate dehydrogenase from the malarial parasite *P. falciparum*. *Molecular and biochemical parasitology*, *90*(1), 235-246.

101. Greenidge, P. A., et al. (2014). Improving docking results via reranking of ensembles of ligand poses in multiple X-ray protein conformations with MM-GBSA. *Journal of chemical information and modeling*, *54*(10), 2697-2717.

102. Gregori-Puigjané, E., and Mestres, J. (2008). A ligand-based approach to mining the chemogenomic space of drugs. *Combinatorial chemistry and high throughput screening*, *11*(8), 669-676.

103. Grinter, S. Z., et al. (2011). An inverse docking approach for identifying new potential anti-cancer targets. *Journal of Molecular Graphics and Modelling*, *29*(6), 795-799.

104. Grosdidier, A., et al. (2007). EADock: docking of small molecules into protein active sites with a multiobjective evolutionary optimization. *Proteins: Structure, Function, and Bioinformatics*, *67*(4), 1010-1025.

105. Grosdidier, A., et al. (2011). SwissDock, a protein-small molecule docking web service based on EADock DSS. *Nucleic acids research*, *39*(suppl_2), W270-W277.

106. Guerler, A., et al. (2013). Mapping monomeric threading to protein–protein structure prediction. *Journal of chemical information and modeling*, *53*(3), 717-725.

107. Günther, S., et al. (2007). Docking without docking: ISEARCH—prediction of interactions using known interfaces. *Proteins: Structure, Function, and Bioinformatics*, *69*(4), 839-844.

108. Hall, D. R., et al. (2011). Hot spot analysis for driving the development of hits into leads in fragment-based drug discovery. *Journal of chemical information and modeling*, *52*(1), 199-209.

109. Halperin, I., et al. (2002). Principles of docking: An overview of search algorithms and a guide to scoring functions. *Proteins: Structure, Function, and Bioinformatics*, *47*(4), 409-443.

110. Hartshorn, M. J., et al. (2007). Diverse, high-quality test set for the validation of protein–ligand docking performance. *Journal of medicinal chemistry*, *50*(4), 726-741.

111. Hoelder, S., et al. (2012). Discovery of small molecule cancer drugs: successes, challenges and opportunities. *Molecular oncology*, *6*(2), 155-176.

112. Hogg, T., et al. (2006). Structural and functional characterization of Falcipain-2, a hemoglobinase from the malarial parasite *Plasmodium falciparum*. *Journal of Biological Chemistry*, *281*(35), 25425-25437.

113. Hong, H. J., et al. (2012). Computational pharmaceutical analysis of anti-Alzheimer's Chinese medicine Coptidis Rhizoma alkaloids. *Molecular medicine reports*, *5*(1), 142-147.

114. Hotez, PJ (2018). The global fight to develop antipoverty vaccines in the anti-vaccine era. *Human vaccines & immunotherapeutics*; 14(9):2128-31. doi: 10.1080/21645515.2018.1430542. PubMed PMID: 29393710; PMCID: PMC6183138

115. Hsin, K. Y., et al. (2013). Combining machine learning systems and multiple docking simulation packages to improve docking prediction reliability for network pharmacology. *PloS one*, *8*(12), e83922.

116. Hu, G., and Agarwal, P. (2009). Human disease-drug network based on genomic expression profiles. *PloS one*, *4*(8), e6536.

117. Huang, N., et al. (2006). Benchmarking sets for molecular docking. *Journal of medicinal chemistry*, *49*(23), 6789-6801.

118. Huang, S. Y., and Zou, X. (2013). A nonredundant structure dataset for benchmarking protein-RNA computational docking. *Journal of computational chemistry*, *34*(4), 311-318.

119. Huang, S. Y., and Zou, X. (2014). A knowledge-based scoring function for protein-RNA interactions derived from a statistical mechanics-based iterative method. *Nucleic acids research*, *42*(7), e55-e55.

120. Huang, S. Y., et al. (2010). Scoring functions and their evaluation methods for protein–ligand docking: recent advances and future directions. *Physical Chemistry Chemical Physics*, *12*(40), 12899-12908.

1011 121. Hughes, J. P., et al. (2011). Principles of early drug discovery. *British journal of*
1012 *pharmacology*, *162*(6), 1239-1249.

1013 122. Hui-fang, L., et al. (2010). Evaluation of various inverse docking schemes in multiple targets
1014 identification. *Journal of Molecular Graphics and Modelling*, *29*(3), 326-330.

1015 123. Irwin, J. J., and Shoichet, B. K. (2005). ZINC− A free database of commercially available
1016 compounds for virtual screening. *Journal of chemical information and modeling*, *45*(1), 177-
1017 182.

1018 124. Iskar, M., et al. (2013). Characterization of drug-induced transcriptional modules: towards
1019 drug repositioning and functional understanding. *Molecular systems biology*, *9*(1), 662.

1020 125. Jablonowski, J. A., et al. (2003). The first potent and selective non-imidazole human
1021 histamine H4 receptor antagonists. *Journal of medicinal chemistry*, *46*(19), 3957-3960.

1022 126. Jain, A. N. (2007). Surflex-Dock 2.1: robust performance from ligand energetic modeling,
1023 ring flexibility, and knowledge-based search. *Journal of computer-aided molecular*
1024 *design*, *21*(5), 281-306.

1025 127. Jain, A. N. (2008). Bias, reporting, and sharing: computational evaluations of docking
1026 methods. *Journal of computer-aided molecular design*, *22*(3-4), 201-212.

1027 128. Jalencas, X., and Mestres, J. (2013). Identification of similar binding sites to detect distant
1028 polypharmacology. *Molecular informatics*, *32*(11-12), 976-990.

1029 129. Janin, J. (2013). The targets of CAPRI rounds 20–27. *Proteins: Structure, Function, and*
1030 *Bioinformatics*, *81*(12), 2075-2081.

1031 130. Jiang, Y., et al. (2014). Schisandrol B protects against acetaminophen-induced
1032 hepatotoxicity by inhibition of CYP-mediated bioactivation and regulation of liver
1033 regeneration. *Toxicological Sciences*, *143*(1), 107-115.

1034 131. Jiménez-García, B., et al. (2015). pyDockSAXS: protein–protein complex structure by
1035 SAXS and computational docking. *Nucleic acids research*, *43*(W1), W356-W361.

1036 132. Jin, G., et al. (2012). A novel method of transcriptional response analysis to facilitate drug
1037 repositioning for cancer therapy. *Cancer research*, *72*(1), 33-44.

1038 133. Jones, G., et al. (1997). Development and validation of a genetic algorithm for flexible
1039 docking. *Journal of molecular biology*, *267*(3), 727-748.

1040 134. Jones, Nathaniel G., et al (2018). Genetically validated drug targets in Leishmania: current
1041 knowledge and future prospects. *ACS infectious diseases* 4.4: 467-477.

1042 135. Kahraman, A., et al. (2013). Cross-link guided molecular modeling with ROSETTA. *PLoS*
1043 *One*, *8*(9), e73411.

1044 136. Kar, G., et al. (2012). Human proteome-scale structural modeling of E2–E3 interactions
1045 exploiting interface motifs. *Journal of proteome research*, *11*(2), 1196-1207.

1046 137. Karha, J., and Topol, E. J. (2004). The sad story of Vioxx, and what we should learn from
1047 it. *Cleveland Clinic journal of medicine*, *71*(12), 933-4.

1048 138. Katchalski-Katzir, E., et al. (1992). Molecular surface recognition: determination of
1049 geometric fit between proteins and their ligands by correlation techniques. *Proceedings of*
1050 *the National Academy of Sciences*, *89*(6), 2195-2199.

1051 139. Kawatkar, S., et al. (2012). Virtual fragment screening: exploration of MM-PBSA re-
1052 scoring. *Journal of computer-aided molecular design*, *26*(8), 921-934.

1053 140. Keiser, M. J., et al. (2009). Predicting new molecular targets for known
1054 drugs. *Nature*, *462*(7270), 175.

1055 141. Kellenberger, E., et al. (2004). Comparative evaluation of eight docking tools for docking
1056 and virtual screening accuracy. *Proteins: Structure, Function, and Bioinformatics*, *57*(2),
1057 225-242.

1058 142. Keserű, G. M. (2001). A virtual high throughput screen for high affinity cytochrome
1059 P450cam substrates. Implications for in silico prediction of drug metabolism. *Journal of*
1060 *computer-aided molecular design*, *15*(7), 649-657.

1061 143. Khan, K. M., et al. (2010). Identification of potent urease inhibitors via ligand-and structure-
1062 based virtual screening and in vitro assays. *Journal of Molecular Graphics and*
1063 *Modelling*, *28*(8), 792-798.

144. Khatoon, Nazia, et al. (2018). "Exploratory algorithm to devise multi-epitope subunit vaccine by investigating Leishmania donovani membrane proteins." *Journal of Biomolecular Structure and Dynamics* just-accepted: 1-42.

145. Kim, D. S., et al. (2011). BetaDock: shape-priority docking method based on beta-complex. *Journal of Biomolecular Structure and Dynamics*, *29*(1), 219-242.

146. Kinnings, S. L., et al. (2009). Drug discovery using chemical systems biology: repositioning the safe medicine Comtan to treat multi-drug and extensively drug resistant tuberculosis. *PLoS computational biology*, *5*(7), e1000423.

147. Kirys, T., et al. (2012). Rotamer libraries and probabilities of transition between rotamers for the side chains in protein–protein binding. *Proteins: Structure, Function, and Bioinformatics*, *80*(8), 2089-2098.

148. Kitchen, D. B., et al. (2004). Docking and scoring in virtual screening for drug discovery: methods and applications. *Nature reviews Drug discovery*, *3*(11), 935.

149. Klimpel, K. R., et al. (1994). Anthrax toxin lethal factor contains a zinc metalloprotease consensus sequence which is required for lethal toxin activity. *Molecular microbiology*, *13*(6), 1093-1100.

150. Koes, D., et al. (2012). Enabling large-scale design, synthesis and validation of small molecule protein-protein antagonists. *PloS one*, *7*(3), e32839.

151. Kontoyianni, M., et al. (2004). Evaluation of docking performance: comparative data on docking algorithms. *Journal of medicinal chemistry*, *47*(3), 558-565.

152. Korb, O., et al. (2009). Empirical scoring functions for advanced protein− ligand docking with PLANTS. *Journal of chemical information and modeling*, *49*(1), 84-96.

153. Korb, O., et al. (2012). Potential and limitations of ensemble docking. *Journal of chemical information and modeling*, *52*(5), 1262-1274.

154. Kramer, B., et al. (1999). Evaluation of the FLEXX incremental construction algorithm for protein–ligand docking. *Proteins: Structure, Function, and Bioinformatics*, *37*(2), 228-241.

155. Krawczyk, K., et al. (2014). Improving B-cell epitope prediction and its application to global antibody-antigen docking. *Bioinformatics*, *30*(16), 2288-2294.

1092 156. Kruiskamp, D. L. W. (1995). DARWIN: CMOS opamp synthesis by means of a genetic
1093 algorithm. In *Design Automation, 1995. DAC'95. 32nd Conference on* (pp. 433-438). IEEE.

1094 157. Kuhl, F. S., et al. (1984). A combinatorial algorithm for calculating ligand binding. *Journal*
1095 *of Computational Chemistry*, *5*(1), 24-34.

1096 158. Kumar, A., and Zhang, K. Y. (2013). Investigation on the effect of key water molecules on
1097 docking performance in CSARdock exercise. *Journal of chemical information and*
1098 *modeling*, *53*(8), 1880-1892.

1099 159. Kumar, A., et al. (2017). Docking Methodologies and Recent Advances. In *Oncology:*
1100 *Breakthroughs in Research and Practice* (pp. 804-828). IGI Global.

1101 160. Kundrotas, P. J., and Vakser, I. A. (2010). Accuracy of protein-protein binding sites in high-
1102 throughput template-based modeling. *PLoS computational biology*, *6*(4), e1000727.

1103 161. Kundrotas, P. J., and Vakser, I. A. (2013). Global and local structural similarity in protein–
1104 protein complexes: Implications for template-based docking. *Proteins: Structure, Function,*
1105 *and Bioinformatics*, *81*(12), 2137-2142.

1106 162. Kundrotas, P. J., et al. (2008). Homology-based modeling of 3D structures of protein–
1107 protein complexes using alignments of modified sequence profiles. *International journal of*
1108 *biological macromolecules*, *43*(2), 198-208.

1109 163. Kundrotas, P. J., et al. (2012). Templates are available to model nearly all complexes of
1110 structurally characterized proteins. *Proceedings of the National Academy of*
1111 *Sciences*, *109*(24), 9438-9441.

1112 164. Labute, M.X. et al. (2014). Adverse Drug Reaction Prediction Using Scores Produced by
1113 Large-Scale Drug-Protein Target Docking on High-Performance Computing Machines. *PloS*
1114 *one*, *9*(9), e106298.

1115 165. Lafleur, K., et al. (2009). Structure-based optimization of potent and selective inhibitors of
1116 the tyrosine kinase erythropoietin producing human hepatocellular carcinoma receptor B4
1117 (EphB4). *Journal of medicinal chemistry*, *52*(20), 6433-6446.

1118 166. Lamb, J., et al. (2006). The Connectivity Map: using gene-expression signatures to connect
1119 small molecules, genes, and disease. S*cience*, *313*(5795), 1929-1935.

1120 167. Leaver-Fay, A., et al. (2011). ROSETTA3: an object-oriented software suite for the
1121 simulation and design of macromolecules. In *Methods in enzymology* (Vol. 487, pp. 545-
1122 574). Academic Press.

1123 168. Lee, H. S., and Zhang, Y. (2012). BSP-SLIM: A blind low-resolution ligand-protein docking
1124 approach using predicted protein structures. *Proteins: Structure, Function, and*
1125 *Bioinformatics*, *80*(1), 93-110.

1126 169. Lee, K., et al. (2009). Structure-based virtual screening of Src kinase inhibitors. *Bioorganic*
1127 *and medicinal chemistry*, *17*(8), 3152-3161.

1128 170. Lemmon, G., and Meiler, J. (2012). Rosetta Ligand docking with flexible XML protocols.
1129 In *Computational Drug Discovery and Design* (pp. 143-155). Springer, New York, NY.

1130 171. Lesk, V. I., and Sternberg, M. J. (2008). 3D-Garden: a system for modelling protein–protein
1131 complexes based on conformational refinement of ensembles generated with the marching
1132 cubes algorithm. *Bioinformatics*, *24*(9), 1137-1144.

1133 172. Li, H., et al. (2014). Istar: A web platform for large-scale protein-ligand docking. *PLoS*
1134 *One*, *9*(1), e85678.

1135 173. Li, J. (2007). *Systematic investigation of quorum sensing in Escherichia coli* (Doctoral
1136 dissertation).

1137 174. Li, X., et al. (2010). Evaluation of the performance of four molecular docking programs on a
1138 diverse set of protein-ligand complexes. *Journal of Computational Chemistry*, *31*(11), 2109-
1139 2125.

1140 175. Li, Y. Y., Ann, J., and Jones, S. J. (2011). A computational approach to finding novel targets
1141 for existing drugs. *PLoS computational biology*, *7*(9), e1002139.

1142 176. Liu, J., and Wang, R. (2015). Classification of current scoring functions. *Journal of*
1143 *chemical information and modeling*, *55*(3), 475-482.

1144 177. Liu, X., et al. (2010). PharmMapper server: a web server for potential drug target
1145 identification using pharmacophore mapping approach. *Nucleic acids research*, *38*(suppl_2),
1146 W609-W614.

178. Lopes, A., et al. (2013). Protein-protein interactions in a crowded environment: an analysis via cross-docking simulations and evolutionary information. *PLoS computational biology*, *9*(12), e1003369.

179. Lounkine, E., et al. (2012). Large-scale prediction and testing of drug activity on side-effect targets. *Nature*, *486*(7403), 361.

180. Lu, L., et al. (2002). MULTIPROSPECTOR: an algorithm for the prediction of protein–protein interactions by multimeric threading. *Proteins: Structure, Function, and Bioinformatics*, *49*(3), 350-364.

181. Lyskov, S., and Gray, J. J. (2008). The RosettaDock server for local protein–protein docking. *Nucleic acids research*, *36*(suppl_2), W233-W238.

182. Macindoe, G., et al. (2010). HexServer: an FFT-based protein docking server powered by graphics processors. *Nucleic acids research*, *38*(suppl_2), W445-W449.

183. Maeda, S., et al. (1999). Diabetic nephropathy is not associated with the dinucleotide repeat polymorphism upstream of the aldose reductase (ALR2) gene but with erythrocyte aldose reductase content in Japanese subjects with type 2 diabetes. *Diabetes*, *48*(2), 420-421.

184. McGann, M. (2012). FRED and HYBRID docking performance on standardized datasets. *Journal of computer-aided molecular design*, *26*(8), 897-906.

185. McGaughey, G. B., et al. (2007). Comparison of topological, shape, and docking methods in virtual screening. *Journal of chemical information and modeling*, *47*(4), 1504-1519.

186. Meng, X. Y., et al. (2011). Molecular docking: a powerful approach for structure-based drug discovery. *Current computer-aided drug design*, *7*(2), 146-157.

187. Mestres, J., et al. (2006). Ligand-based approach to in-silico pharmacology: nuclear receptor profiling. *Journal of chemical information and modeling*, *46*(6), 2725-2736.

188. Michel, J., and Essex, J. W. (2010). Prediction of protein–ligand binding affinity by free energy simulations: assumptions, pitfalls and expectations. *Journal of computer-aided molecular design*, *24*(8), 639-658.

1173 189. Miles, R. W., et al. (1998). One-third-the-sites transition-state inhibitors for purine
1174 nucleoside phosphorylase. *Biochemistry*, *37*(24), 8615-8621.

1175 190. Miyauchi, S., et al. (2010). New frontiers in gut nutrient sensor research: free fatty acid
1176 sensing in the gastrointestinal tract. *Journal of pharmacological sciences*, *112*(1), 19-24.

1177 191. Mladenović, M., et al. (2015). Newly discovered chroman-2, 4-diones neutralize the in vivo
1178 DNA damage induced by alkylation through the inhibition of Topoisomerase IIα: A story
1179 behind the molecular modeling approach. *Biochemical pharmacology*, *98*(1), 243-266.

1180 192. Montes, M., et al. (2008). Receptor-based virtual ligand screening for the identification of
1181 novel CDC25 phosphatase inhibitors. *Journal of chemical information and modeling*, *48*(1),
1182 157-165.

1183 193. Morris, G. M., et al. (2009). AutoDock4 and AutoDockTools4: Automated docking with
1184 selective receptor flexibility. *Journal of computational chemistry*, *30*(16), 2785-2791.

1185 194. Mosca, R., et al. (2009). Pushing structural information into the yeast interactome by high-
1186 throughput protein docking experiments. *PLoS computational biology*, *5*(8), e1000490.

1187 195. Mosca, R., et al. (2013). Towards a detailed atlas of protein–protein interactions. *Current*
1188 *opinion in structural biology*, *23*(6), 929-940.

1189 196. Nabuurs, S. B., et al. (2007). A flexible approach to induced fit docking. *Journal of*
1190 *medicinal chemistry*, *50*(26), 6507-6518.

1191 197. Namasivayam, V., and Günther, R. (2007). PSO@ AUTODOCK: A fast flexible molecular
1192 docking program based on swarm intelligence. *Chemical biology and drug design*, *70*(6),
1193 475-484.

1194 198. Natesan, S., et al. (2012). Binding affinity prediction for ligands and receptors forming
1195 tautomers and ionization species: inhibition of mitogen-activated protein kinase-activated
1196 protein kinase 2 (MK2). *Journal of medicinal chemistry*, *55*(5), 2035-2047.

1197 199. Nero, T. L., et al. (2014). Oncogenic protein interfaces: small molecules, big challenges.
1198 *Nature Reviews Cancer*, *14*(4), 248.

1199 200. Nissen, S. (2010). Erratum: The rise and fall of rosiglizazone (European Heart Journal 31: 7
1200 (773)). *European Heart Journal*, *31*(10), 1281.

1201 201. Nissink, J. W. M., et al. (2002). A new test set for validating predictions of protein–ligand
1202 interaction. *Proteins: Structure, Function, and Bioinformatics*, *49*(4), 457-471.

1203 202. Osguthorpe, D. J., et al. (2012). Exploring protein flexibility: incorporating structural
1204 ensembles from crystal structures and simulation into virtual screening protocols. *The*
1205 *journal of physical chemistry B*, *116*(23), 6952-6959.

1206 203. Österberg, F., et al. (2002). Automated docking to multiple target structures: incorporation
1207 of protein mobility and structural water heterogeneity in AutoDock. *Proteins: Structure,*
1208 *Function, and Bioinformatics*, *46*(1), 34-40.

1209 204. Ostrov, D. A., et al. (2007). Discovery of novel DNA gyrase inhibitors by high-throughput
1210 virtual screening. *Antimicrobial agents and chemotherapy*, *51*(10), 3688-3698.

1211 205. O'Toole, N., and Vakser, I. A. (2008). Large-scale characteristics of the energy landscape in
1212 protein–protein interactions. *Proteins: Structure, Function, and Bioinformatics*, *71*(1), 144-
1213 152.

1214 206. Pagadala, N. S., et al. (2017). Software for molecular docking: a review. *Biophysical*
1215 *reviews*, *9*(2), 91-102.

1216 207. Pang, S. S., et al. (2002). Crystal structure of yeast acetohydroxyacid synthase: a target for
1217 herbicidal inhibitors. *Journal of molecular biology*, *317*(2), 249-262.

1218 208. Pang, Y. P., et al. (2001). EUDOC: a computer program for identification of drug interaction
1219 sites in macromolecules and drug leads from chemical databases. *Journal of computational*
1220 *chemistry*, *22*(15), 1750-1771.

1221 209. Papa, A., et al. (2002). Genetic characterization of the m-RNA segment of Crimean Congo
1222 hemorrhagic fever virus strains, China. *Emerging infectious diseases*, *8*(1), 50.

1223 210. Park, H., et al. (2004). Loop flexibility and solvent dynamics as determinants for the
1224 selective inhibition of cyclin-dependent kinase 4: comparative molecular dynamics
1225 simulation studies of CDK2 and CDK4. *Chembiochem*, *5*(12), 1662-1672.

1226 211. Paul, N., and Rognan, D. (2002). ConsDock: A new program for the consensus analysis of
1227 protein–ligand interactions. *Proteins: Structure, Function, and Bioinformatics*, *47*(4), 521-
1228 533.

1229 212. Pei, J., et al. (2006). PSI-DOCK: Towards highly efficient and accurate flexible ligand
1230 docking. *Proteins: Structure, Function, and Bioinformatics*, *62*(4), 934-946.

1231 213. Pérez, C., and Ortiz, A. R. (2001). Evaluation of docking functions for protein− ligand
1232 docking. *Journal of medicinal chemistry*, *44*(23), 3768-3785.

1233 214. Perez-Cano, et al. (2010). Structural prediction of protein-RNA interaction by computational
1234 docking with propensity-based statistical potentials. In *Biocomputing 2010* (pp. 293-301).

1235 215. Pierce, A. C., et al. (2008). Docking study yields four novel inhibitors of the protooncogene
1236 Pim-1 kinase. *Journal of medicinal chemistry*, *51*(6), 1972-1975.

1237 216. Pierce, B. G., et al. (2014). ZDOCK server: interactive docking prediction of protein–protein
1238 complexes and symmetric multimers. *Bioinformatics*, *30*(12), 1771-1773.

1239 217. Plewczynski, D., et al. (2011). Can we trust docking results? Evaluation of seven commonly
1240 used programs on PDBbind database. *Journal of computational chemistry*, *32*(4), 742-755.

1241 218. Plewczynski, D., et al. (2011). VoteDock: consensus docking method for prediction of
1242 protein–ligand interactions. *Journal of Computational Chemistry*, *32*(4), 568-581.

1243 219. Pradelles, P., et al. (1985). Enzyme immunoassays of eicosanoids using acetylcholine
1244 esterase as label: an alternative to radioimmunoassay. *Analytical Chemistry*, *57*(7), 1170-
1245 1173.

1246 220. Prakhov, N. D., et al. (2010). VSDocker: a tool for parallel high-throughput virtual
1247 screening using AutoDock on Windows-based computer clusters. *Bioinformatics*, *26*(10),
1248 1374-1375.

1249 221. Premont, R. T., et al. (1995). Protein kinases that phosphorylate activated G protein-coupled
1250 receptors. *The FASEB Journal*, *9*(2), 175-182.

1251 222. Puton, T., et al. (2012). Computational methods for prediction of protein–RNA interactions.
1252 *Journal of structural biology*, *179*(3), 261-268.

1253 223. Puvanendrampillai, D., and Mitchell, J. B. (2003). Protein Ligand Database (PLD):
1254 additional understanding of the nature and specificity of protein–ligand complexes.
1255 *Bioinformatics*, *19*(14), 1856-1857.

1256 224. Raha, K., and Merz, K. M. (2004). A quantum mechanics-based scoring function: study of
1257 zinc ion-mediated ligand binding. *Journal of the American Chemical Society*, *126*(4), 1020-
1258 1021.

1259 225. Rajkhowa, S., and Deka, R. C. (2017). Protein-Ligand Docking Methodologies and Its
1260 Application in Drug Discovery. In *Oncology: Breakthroughs in Research and Practice* (pp.
1261 891-914). IGI Global.

1262 226. Rashad, A. A., and Keller, P. A. (2013). Structure based design towards the identification of
1263 novel binding sites and inhibitors for the chikungunya virus envelope proteins. *Journal of*
1264 *Molecular Graphics and Modelling*, *44*, 241-252.

1265 227. Rawal, R. K., et al (2005). 2-(Aryl)-3-furan-2-ylmethyl-thiazolidin-4-ones as selective HIV-
1266 RT inhibitors. *Bioorganic and medicinal chemistry*, *13*(24), 6771-6776.

1267 228. Rentzsch, R., and Renard, B. Y. (2015). Docking small peptides remains a great challenge:
1268 an assessment using AutoDock Vina. *Briefings in bioinformatics*, *16*(6), 1045-1056.

1269 229. Ricote, M., et al. (1998). The peroxisome proliferator-activated receptor-γ is a negative
1270 regulator of macrophage activation. *Nature*, *391*(6662), 79.

1271 230. Rodrigues, J. P. G. L. M., et al. (2013). Defining the limits of homology modeling in
1272 information-driven protein docking. *Proteins: Structure, Function, and Bioinformatics*,
1273 *81*(12), 2119-2128.

1274 231. Rueda, M., et al. (2012). ALiBERO: evolving a team of complementary pocket
1275 conformations rather than a single leader. *Journal of chemical information and modeling*,
1276 *52*(10), 2705-2714.

1277 232. Ruiz-Carmona, S., et al. (2014). rDock: a fast, versatile and open source program for
1278 docking ligands to proteins and nucleic acids. *PLoS computational biology*, *10*(4),
1279 e1003571.

1280 233. Ruvinsky, A. M., and Vakser, I. A. (2008). Chasing funnels on protein-protein energy
1281 landscapes at different resolutions. *Biophysical journal*, *95*(5), 2150-2159.

1282 234. Sable, R., and Jois, S. (2015). Surfing the protein-protein interaction surface using docking
1283 methods: application to the design of PPI inhibitors. *Molecules*, *20*(6), 11569-11603.

1284 235. Santiago, D. N., et al. (2012). Virtual target screening: validation using kinase inhibitors.
1285 *Journal of chemical information and modeling*, *52*(8), 2192-2203.

1286 236. Sauton, N., et al. (2008). MS-DOCK: accurate multiple conformation generator and rigid
1287 docking protocol for multi-step virtual ligand screening. *BMC bioinformatics*,*9*(1), 184.

1288 237. Sawada, R., et al. (2015). Target-based drug repositioning using large-scale chemical–
1289 protein interactome data. *Journal of chemical information and modeling*, *55*(12), 2717-2730.

1290 238. Schneidman-Duhovny, D., et al. (2005). PatchDock and SymmDock: servers for rigid and
1291 symmetric docking. *Nucleic acids research*, *33*(suppl_2), W363-W367.

1292 239. Schreyer, A., and Blundell, T. (2009). CREDO: a protein–ligand interaction database for
1293 drug discovery. *Chemical biology and drug design*, *73*(2), 157-167.

1294 240. Schulz-Gasch, T., and Stahl, M. (2003). Binding site characteristics in structure-based
1295 virtual screening: evaluation of current docking tools. *Journal of molecular modeling*, *9*(1),
1296 47-57.

1297 241. Schutz, C. N., and Warshel, A. (2001). What are the dielectric "constants" of proteins and
1298 how to validate electrostatic models? *Proteins: Structure, Function, and Bioinformatics*,
1299 *44*(4), 400-417.

1300 242. Shameer, K., et al. (2017). Systematic analyses of drugs and disease indications in
1301 RepurposeDB reveal pharmacological, biological and epidemiological factors influencing
1302 drug repositioning. *Briefings in bioinformatics*.

1303 243. Sherman, W., et al. (2006). Novel procedure for modeling ligand/receptor induced fit
1304 effects. *Journal of medicinal chemistry*, *49*(2), 534-553.

1305 244. Shin, W. H., et al. (2011). LigDockCSA: protein–ligand docking using conformational space
1306 annealing. *Journal of computational chemistry*, *32*(15), 3226-3232.

1307 245. Shoichet, B. K., and Kobilka, B. K. (2012). Structure-based drug screening for G-protein-
1308 coupled receptors. *Trends in pharmacological sciences*, *33*(5), 268-272.

1309 246. Sieghart, W. (2006). Structure, pharmacology, and function of GABAA receptor subtypes.
1310 *Advances in pharmacology*, *54*, 231-263.

1311 247. Singh, S., et al. (2017). Evaluation of protein-ligand docking methods on peptide-ligand
1312 complexes for docking small ligands to peptides. *bioRxiv*, 212514.

1313 248. Singh, T., et al. (2011). AADS-An automated active site identification, docking, and scoring
1314 protocol for protein targets based on physicochemical descriptors. *Journal of chemical*
1315 *information and modeling*, *51*(10), 2515-2527.

1316 249. Sinha, R., et al. (2010). Docking by structural similarity at protein-protein interfaces.
1317 *Proteins: Structure, Function, and Bioinformatics*, *78*(15), 3235-3241.

1318 250. Sinko, W., et al. (2013). Accounting for receptor flexibility and enhanced sampling methods
1319 in computer-aided drug design. *Chemical biology and drug design*, *81*(1), 41-49.

1320 251. Smith, G. R., and Sternberg, M. J. (2002). Prediction of protein–protein interactions by
1321 docking methods. *Current opinion in structural biology*, *12*(1), 28-35.

1322 252. Sobolev, V., et al. (1997). CASP2 molecular docking predictions with the LIGIN software.
1323 *Proteins: Structure, Function, and Bioinformatics*, *29*(S1), 210-214.

1324 253. Sousa, S. F., et al. (2013). Protein-ligand docking in the new millennium–a retrospective of
1325 10 years in the field. *Current medicinal chemistry*, *20*(18), 2296-2314.

1326 254. Sridhar, A., et al. (2017). Waterdock 2.0: Water placement prediction for Holo-structures
1327 with a pymol plugin. *PloS one*, *12*(2), e0172743.

1328 255. Sriram, D., et al. (2006). Gatifloxacin derivatives: synthesis, antimycobacterial activities,
1329 and inhibition of *Mycobacterium tuberculosis* DNA gyrase. *Bioorganic and medicinal*
1330 *chemistry letters*, *16*(11), 2982-2985.

1331 256. Stewart, K. D., et al. (1990). DOCKing ligands into receptors: The test case of α-
1332 chymotrypsin. *Tetrahedron Computer Methodology*, *3*(6), 713-722.

1333 257. Stroganov, O. V., et al. (2008). Lead finder: an approach to improve accuracy of protein–
1334 ligand docking, binding energy estimation, and virtual screening. *Journal of Chemical*
1335 *Information and Modeling*, *48*(12), 2371-2385.

1336 258. Sudha, S., et al. (2018). Protein Docking using Constrained Self-adaptive Differential
1337 Evolution Algorithm. *bioRxiv*, 312801.

1338 259. Suthram, S., et al. (2010). Network-based elucidation of human disease similarities reveals
1339 common functional modules enriched for pluripotent drug targets. *PLoS computational*
1340 *biology*, *6*(2), e1000662.

1341 260. Szilagyi, A., and Zhang, Y. (2014). Template-based structure modeling of protein–protein
1342 interactions. *Current opinion in structural biology*, *24*, 10-23.

1343 261. Taylor, R. D., et al. (2002). A review of protein-small molecule docking methods. *Journal of*
1344 *computer-aided molecular design*, *16*(3), 151-166.

1345 262. Taylor, R. D., et al. (2003). FDS: flexible ligand and receptor docking with a continuum
1346 solvent model and softcore energy function. *Journal of computational chemistry*, *24*(13),
1347 1637-1656.

1348 263. Therrien, E et al. (2014). Docking Ligands into Flexible and Solvated Macromolecules. 7.
1349 Impact of Protein Flexibility and Water Molecules on Docking-Based Virtual Screening
1350 Accuracy. *J. Chem. Inf. Model.*, *54* (11), pp 3198–3210.

1351 264. Thomsen, R., and Christensen, M. H. (2006). MolDock: a new technique for high-accuracy
1352 molecular docking. *Journal of medicinal chemistry*, *49*(11), 3315-3321.

1353 265. Torchala, M., et al. (2013). SwarmDock: a server for flexible protein–protein docking.
1354 *Bioinformatics*, *29*(6), 807-809.

1355 266. Tosh, D. K., et al. (2012). Optimization of adenosine 5′-carboxamide derivatives as
1356 adenosine receptor agonists using structure-based ligand design and fragment screening.
1357 *Journal of medicinal chemistry*, *55*(9), 4297-4308.

1358 267. Tovchigrechko, A., and Vakser, I. A. (2006). GRAMM-X public web server for protein–
1359 protein docking. *Nucleic acids research*, *34*(suppl_2), W310-W314.

1360 268. Tovchigrechko, A., et al. (2002). Docking of protein models. *Protein Science*, *11*(8), 1888-
1361 1896.

1362 269. Trott, O., and Olson, A. J. (2010). AutoDock Vina: improving the speed and accuracy of
1363 docking with a new scoring function, efficient optimization and multithreading. *Journal of*
1364 *computational chemistry*, *31*(2), 455-461.

1365 270. Tuncbag, N., et al. (2012). Fast and accurate modeling of protein–protein interactions by
1366 combining template-interface-based docking with flexible refinement. *Proteins: Structure,*
1367 *Function, and Bioinformatics*, *80*(4), 1239-1249.

1368 271. Tuszynska, I., et al. (2015). NPDock: a web server for protein–nucleic acid docking. *Nucleic*
1369 *acids research*, *43*(W1), W425-W430.

1370 272. Urquhart, B. L., Tirona, R. G., & Kim, R. B. (2007). Nuclear receptors and the regulation of
1371 drug-metabolizing enzymes and drug transporters: implications for interindividual
1372 variability in response to drugs. *The Journal of Clinical Pharmacology*, *47*(5), 566-578.

1373 273. Vakser, I. A. (1997). Evaluation of GRAMM low-resolution docking methodology on the
1374 hemagglutinin-antibody complex. *Proteins: Structure, Function, and Bioinformatics*,
1375 *29*(S1), 226-230.

1376 274. Vakser, I. A. (2013). Low-resolution structural modeling of protein interactome. *Current*
1377 *opinion in structural biology*, *23*(2), 198-205.

1378 275. Van De Laar, F. A., et al. (2005). α-Glucosidase inhibitors for patients with type 2 diabetes:
1379 results from a Cochrane systematic review and meta-analysis. *Diabetes care*, *28*(1), 154-
1380 163.

1381 276. Vasudevan, S. R., et al. (2012). Shape-based reprofiling of FDA-approved drugs for the H1
1382 histamine receptor. *Journal of medicinal chemistry*, *55*(16), 7054-7060.

1383 277. Venkatapathy, R., et al. (2004). Assessment of the oral rat chronic lowest observed adverse
1384 effect level model in TOPKAT, a QSAR software package for toxicity prediction. *Journal of*
1385 *chemical information and computer sciences*, *44*(5), 1623-1629.

1386 278. Verdonk, M. L., et al. (2003). Improved protein–ligand docking using GOLD. *Proteins:*
1387 *Structure, Function, and Bioinformatics*, *52*(4), 609-623.

1388 279. Verdonk, M. L., et al. (2005). Modeling water molecules in protein− ligand docking using
1389 GOLD. *Journal of medicinal chemistry*, *48*(20), 6504-6515.

1390 280. Vitale, G., et al. (2000). Susceptibility of mitogen-activated protein kinase kinase family
1391 members to proteolysis by anthrax lethal factor. *Biochemical Journal*, *352*(3), 739-745.

1392 281. Wang, C., et al. (2005). Improved side-chain modeling for protein–protein docking. *Protein*
1393 *Science*, *14*(5), 1328-1339.

1394 282. WANG, S. Q., et al. (2009). Optimization Design of Engine Piston Profile Basing on Glide
1395 Software [J]. *Internal Combustion Engine and Powerplant*, *6*, 008.

1396 283. Wang, Y., et al. (2009). PubChem: a public information system for analyzing bioactivities of
1397 small molecules. *Nucleic acids research*, *37*(suppl_2), W623-W633.

1398 284. Wang, Z., et al. (2016). Comprehensive evaluation of ten docking programs on a diverse set
1399 of protein–ligand complexes: the prediction accuracy of sampling power and scoring power.
1400 *Physical Chemistry Chemical Physics*, *18*(18), 12964-12975.

1401 285. Wass, M. N., et al. (2011). Challenges for the prediction of macromolecular interactions.
1402 *Current opinion in structural biology*, *21*(3), 382-390.

1403 286. Winkler, G (2012). Image analysis, random fields and Markov chain Monte Carlo methods:
1404 a mathematical introduction. Vol. 27. *Springer Science and Business Media*.

1405 287. Wodak, S. J., et al. (2013). Protein–protein interaction networks: the puzzling riches.
1406 *Current opinion in structural biology*, *23*(6), 941-953.

1407 288. Wu, G., et al. (2003). Detailed analysis of grid-based molecular docking: A case study of
1408 CDOCKER—A CHARMm-based MD docking algorithm. *Journal of computational*
1409 *chemistry*, *24*(13), 1549-1562.

1410 289. Xu, M., and Lill, M. A. (2011). Utilizing experimental data for reducing ensemble size in
1411 flexible-protein docking. *Journal of chemical information and modeling*, *52*(1), 187-198.

1412 290. Yang, J. M., and Chen, C. C. (2004). GEMDOCK: a generic evolutionary method for
1413 molecular docking. *Proteins: Structure, Function, and Bioinformatics*, *55*(2), 288-304.

1414 291. Yu, J., et al. (2016). InterEvDock: a docking server to predict the structure of protein–
1415 protein interactions using evolutionary information. *Nucleic acids research*, *44*(W1), W542-
1416 W549.

1417 292. Yuriev, E., and Ramsland, P. A. (2013). Latest developments in molecular docking: 2010–
1418 2011 in review. *Journal of Molecular Recognition*, *26*(5), 215-239.

1419 293. Yuriev, E., et al. (2011). Challenges and advances in computational docking: 2009 in
1420 review. *Journal of Molecular Recognition*, *24*(2), 149-164.

1421 294. Yuriev, E., et al. (2015). Improvements, trends, and new ideas in molecular docking: 2012–
1422 2013 in review. *Journal of Molecular Recognition*, *28*(10), 581-604.

1423 295. Yusuf, D., et al. (2008). An alternative method for the evaluation of docking performance:
1424 RSR vs RMSD. *Journal of Chemical Information and Modeling*, *48*(7), 1411-1422.

1425 296. Zeng, J. (2000). Mini-review: computational structure-based design of inhibitors that target
1426 protein surfaces. *Combinatorial Chemistry and High Throughput Screening*, *3*(5), 355-362.

297. Zhang, J., et al. (2017). Inhibition of PHD3 by salidroside promotes neovascularization through cell–cell communications mediated by muscle-secreted angiogenic factors. *Scientific reports*, *7*, 43935.

298. Zhang, Q. C., et al. (2012). Structure-based prediction of protein–protein interactions on a genome-wide scale. *Nature*, *490*(7421), 556.

299. Zhang, X. W. (2013). A combination of epitope prediction and molecular docking allows for good identification of MHC class I restricted T-cell epitopes. *Computational biology and chemistry*, *45*, 30-35.

300. Zhang, X., et al. (2000). Crystal structure of the conserved core of protein arginine methyltransferase PRMT3. *The EMBO journal*, *19*(14), 3509-3519.

301. Zhang, Y. M., et al. (2004). Synthesis and SAR of potent EGFR/erbB2 dual inhibitors. *Bioorganic and medicinal chemistry letters*, *14*(1), 111-114.

302. Zhang, Y., and Skolnick, J. (2005). TM-align: a protein structure alignment algorithm based on the TM-score. *Nucleic acids research*, *33*(7), 2302-2309.

303. Zheng, S., et al. (2007). A knowledge-based potential function predicts the specificity and relative binding energy of RNA-binding proteins. *The FEBS journal*, *274*(24), 6378-6391.

304. Zhu, T., et al. (2013). Fragment-based drug discovery using a multidomain, parallel MD-MM/PBSA screening protocol. *Journal of chemical information and modeling*, *53*(3), 560-572.

305. Zhu, Z., et al. (2008). Large-scale structural modeling of protein complexes at low resolution. *Journal of bioinformatics and computational biology*, *6*(04), 789-810.

Table 1: Examples of docking tools used in rigid and flexible docking.

RIGID DOCKING	FLEXIBLE DOCKING
• ZDOCK	• AutoDock
• RDOCK	• FLIPDock
• MEGADOCK	• HADDOCK
• SOFTDOCK	• FTDOCK
• BiGGER	
• SKE-DOCK	

1483 **Table 2: Docking algorithms: Features.**
1484

Algorithms	Features	Disadvantages	Softwares
• Evolutionary Programming (EP)	• It uses a heuristic search algorithm that combines differential evolution with a cavity prediction algorithm. • The docking accuracy has been evaluated by docking flexible ligands to 77 protein targets. MolDock was able to identify the correct binding mode of 87% of the complexes.	• No explicit operator is used to maintain the spreading of solutions in the obtained non-dominated set.	• MolDock [Thomsen et al. 2006]
• Fast Fourier Transform algorithm	• Converts a signal from its original domain to a representation in the frequency domain and vice versa. • Enables a systematic global docking search on a 3D framework.	• Very specific • Provides a limited view of frequency in the context of signal processing. • Frequencies are sensitive to noise.	• ZDOCK server [Chen et al. 2003]

• Genetic algorithms	• A stochastic genetic algorithm in which the variables to be optimized are referred to as genes and the string containing the genes known as chromosomes. • Ability to handle a diverse and large set of variables.	• Protein motion is not modeled. • Large-scale protein conformational changes prediction is difficult. • Good at finding the regions where extremes are located but difficult to find the precise location.	• AutoDock [Morris et al. 1998] • GOLD [Jonnes et al. 1997] • DIVALI [Clark et al. 1995] • DARWIN [Kruiskamp et al. 1995]
• Guided Differential Evolution	• Uses the knowledge of cavities present in the target protein to restrict the search space. • It starts with the initial set of candidate solutions and the poses are then evaluated using a scoring function.	• Unstable convergence. • The number of computations used for finding the minimum energy conformation is more since it is an iterative process.	• MolDock [Sudha et al. 2018]
• Incremental Construction	• The method fragments the ligand and docks them separately on the receptor site.	• Ineffective if ligands have greater than 17 rotatable bonds.	• DOCK 4.0 [Pagadala et al. 2017] • FlexX [Rajkhowa et al. 2017] • eHiTS [Zsoldos et al. 2007]
• LUDI	• Can be used to search large databases of three-dimensional structures for putative ligands	• It may be difficult to find a template that connects the fragments in a stereo-chemically and a synthetically	• FlexX [Kramer et al. 1999] • Pagadala et al. 2017

	of proteins with known 3D structure. • The method holds promise to retrieve protein ligands from a 3D database automatically if the 3D structure of the target protein is known. • Utilizes the hydrogen bond formed between the ligand and proteins at the binding site.	feasible way.	
• Multiple Copy Simultaneous Search (MCSS)	• Makes thousands of copies of a ligand functional group and places them in the receptor binding site. • It then obtains favorable ligand functional conformations by subjecting it to energy minimization.	• Applicable to rigid receptors and not to flexible receptors.	• HOOK [Eisen et al. 1994] • FlexX [Zeng et al. 2000]
• Matching Algorithm	• Chemical information and shape features are used to map a ligand into the active site of a protein.	• Depends on the pharmacophoric pattern i.e., the geometric pattern of atoms responsible for the observed activity.	• EUDOC [Pang et al. 2001]

	• Represents proteins and ligands as pharmacophores.		
• Molecular Dynamics	• Each atom can be separately moved in the field of atoms in a state of rest. • The flexibility of protein and ligand is more effectively represented. • Local optimization can be done efficiently.	• Progresses in very small steps and thus have difficulties in stepping over high energy conformational barriers.	• AUTODOCK VINA [Trott et al. 2010]
• Monte Carlo	• Leads to the class of stochastic methods. • Bond rotation, rigid-body translation or rotation of the ligand is done to generate multiple poses.	• Parameters for optimization have to be pre-defined. • Valid for small molecule conformations on receptor sites not for large molecules.	• MCDOCK [Liu et al. 1999] • ICM (Iterated Conditional Modes) [Winkler et al. 2012]
• Simulated Annealing	• Every docking conformation is simulated; in each cycle of simulation, the temperature gradually decreases in a fixed interval of time. • Considers the	• The temperature keeps the algorithm from getting stuck by permitting uphill moves. • Needs to be combined with MC, GA, and LGA to give higher accuracy results.	• MolDock [Thomsen and Christensen et al. 2006] • AutoDock4 [Morris et al. 2009] • ROSETTA3 [Leaver-Fay et al. 2011] • AutoDock Vina [Trott et al. 2010]

	flexibility and conformational state of both the ligand and the protein.		
• Tabu Search	• It is a MetaHeuristic algorithm. • Uses a Tabu list that prevents revisiting of the previously considered solutions and enables the search for new solutions.	• Regression model needs to be solved every time any of the first m(n+1) weights are changed in order to calculate the mean squared error.	• Pro_leads [Fogel et al. 2008] • SFDock [Fogel et al. 2008]

Table 3: Examples of scoring functions (Also See Table 8).

Force-Field	Empirical	Knowledge-Based
• D-Score	• LUDI	• PMF
• AutoDock	• F-Score	• Drug Score
• DOCK	• Chem Score	• Smog
• GOLD	• X-SCORE	

Table 4: Comparison of Rigid and Flexible Docking.

Parameter	Rigid	Flexible
• Computational Hardware	• Can work on standard systems.	• High-end computational power is needed in terms of RAM, processors etc.
• When the number of ligands is more i.e. screening of thousands of compounds from a single database.	• Performs well even if the numbers of ligands are large.	• Not a good choice
• Change in binding pocket shape during binding.	• Does not work in situations where shape changes during the docking process.	• Works without any problem.

1560 **Table 5: Widely used docking programs and their applications in large-scale**
1561 **docking studies.**
1562

DOCKING TOOLS & VERSIONS	FEATURES	PERFORMANCE	PROTEIN DATASET	BENCHMARKS	REFERENCES
AADS (Automated Active Site Identification)	• Can be accessed free on the internet. • The program detects a total of 10 possible binding sites within a target protein taking into consideration the physicochemical properties of the amino acid side chains around the possible protein cavities.	85%	• 1A4K - Diels alderase catalytic antibody	Performs rigid docking of an input ligand/candidate molecule at the 10 predicted binding sites using an all-atom energy based Monte Carlo method. Based on BappI.	[Singh et al. 2011]
Autodock Vina Autodock 1 Autodock 2.4 Autodock 3 Autodock 4 & 4.2	• Can be accessed free on the internet and is based on flexible ligand and flexible protein side chains docking. • It is used for docking of the ligand to a set of grids classifying the target proteins.	>=70 %	• Set of HIV-1 Protease complex	• Used eight protein-peptide complexes(PPeCs), with peptides up to four residue and 12 rotatable bonds (RBs), introducing 'blind' docking.	[Chang et al. 2010]

BetaDock	• Can be accessed free on the internet. • Based on the use of Voronoi diagrams. • Uses theory of complexes, shape complementarity between a receptor and a ligand.	Not available	• Astex Diverse set of protein database	• It was tested against AutoDock 4 (ligand flexibility turned off) where 85 protein-ligand complexes were taken from the Astex Diverse set database, gave better results, both in terms of the structural quality of the solutions obtained and also in terms of speed.	[Kim et al. 2010]
CDOCKER	• A molecular dynamics (MD) simulated-annealing-based algorithm. • Used to compare the relative performance and accuracy of various grid-based approximations to explicit all-atom force field calculations	74%	• Penicillin binding protein 4	• For calculating the all-atom force field explicitly of various grid-based estimation to compare the relative performance and accuracy. • In these docking studies, the proteins are rigid while the ligands are considered fully flexible and a terminating	[Wu et al. 2002]

				minimization step is used to refine the docked poses.	
Dock Dock4 Dock3	• Can be accessed freely on the internet. • Search strategies used: - Incremental construction and random conformation search. It utilizes the Coulombic and Lennard-Jones-grid based scoring function.	30%	• In silico mutagenesis and docking in *Ralstonia solanacearum* lectin (RSL).	• 15 crystallographic test cases, created from 12 unique complexes whose ligands vary in size and flexibility. For all test cases, at least one docked position is generated within 2 Å of the crystallographic position. • For 7 of 15 test cases, the top scoring position is also within 2 Å of the crystallographic position.	[Ewing and Todd et al. 2001]
DockoMatic DockoMatic 10.0.4.2145	• It is a free open source application and a Linux-based HTVS program, which uses a collaboration of front- end-back-end processing tools for file	76%	• Conotoxins binding with neuronal nicotinic acetylcholine receptors (nAchRs).	• Ligands that accommodate an open access domain NMR solution structure, PDB file was analyzed in the bound state in the crystal structure, the peptide was	[Jacob et al. 2013]

	preparation, parsing, and data analysis. DockoMatic can dock secondary ligands and may be used to assist inverse virtual screening.			dissociated from the ligand binding domain, and hence it was used to redock the peptides.	
EADock SWISS DOCK S3DB (Simple Sloppy Semantic Database)	• Free open source software and a graphical user nterface application that is pre-determination of Auto Dock Jobs (creation & management, repose & automate) for high-throughput screening of receptor/ligand interactions.	77-86%	• The RGD cyclic pent peptide on the αVβ3 integrin.	• It has the ability to generate a good solution through sampling and recognize this solution as the correct one by its scoring function. 37 test cases using a realistic seeding ranging from 3 to 10 A˚ RMSD to the crystal structure was performed.	[Grosdidier et al. 2011]
FDS	• The docking of flexible small molecule ligands to large flexible protein targets is addressed using a two-stage simulation-based method. • It is a hybrid approach where	Not available	• Docking procedure is optimized for single complex of arabinose binding protein	• 14 complexes were examined for fully flexible ligand, both with or without protein side chain flexibility • 11 for the flexible docking, due to the presence of	[Taylor et al. 2003]

	the first component is docking of the ligand to the protein binding site. It is based on sets of simultaneously satisfied intermolecular hydrogen bonds using graph theory and a recursive distance geometry algorithm.			clusters of low energy structures which shows the possibility of more than one binding conformation during ligand optimization.	
FINDSITE-LHM	• Freely available to the academic community. FINDSITE-LHM is a hybrid evolutionary docking algorithm with two fitness functions. • Based on binding-site similarity across groups of weakly homologous template structures identified from threading.	67-71%	• G protein-coupled receptors (GPCRs)	• Ensures chemical diversity of ligands and maintains the physicochemical similarity between ligands and decoys. It also makes the decoys dissimilar in chemical topology to all ligands to avoid false negatives, and maximize spatial random distribution.	[Brylinski et al. 2008]

FLOG	• This docking program heavily prunes the matching search tree using a minimum-residue search heuristic. • Examines all possible nodes (pairing of atom on particular site).	79%	• Receptor structures from PDB are obtained. • DNA Dodecamer and Netrospin. • Purine nucleoside phosphorylase with guanine.	• Study for single docking graph representation.	[Meng et al. 2012]
FTDOCK	• Fourier-transform rigid body docking. • To speed up the surface complementarity calculations, Fourier transform is used to predict the correct binding geometry.	Not available	• Enzyme-inhibitor • CHI, Human pancreatic trypsin inhibitor • Bovine pancreatic trypsin inhibitor • Antibody antigen system • FDL-D44.1 Fab • MLC-D1.3 Fv	• 5 enzyme-inhibitor and 2 antibody-antigen complexes studied, where the antibody was from the bound crystallographic complex.	[Smith et al. 2002]
GASP (Genetic Algorithm Similarity Program) GASP LAB7	• It is freely available for researchers up to 30 trials and is used to measure a ligand conformation	92.5%	• On post-operative implants	• User-specified one-sided communication operations into hardware-level communication operations.	[Khan et al. 2015]

	and orientation relative to the active site of the target protein.				
GEMDOCK GemDock-windows GemDock cent OS5 GemDock linux 9	• It is based on Flexible ligand docking and is a GA-based docking program, which is freely available. It has a partial flexibility for protein.	79%	• Dihydrofolate reductase and trypsin	• 10 ligand-receptor complexes were taken to evaluate GEMDOCK on a problem in which a protein structure is in small motion during docking processing. Experimental results indicate that GEMDOCK is robust and the empirical scoring function is simple and fast to recognize compounds.	[Yang and Chen et al. 2004]
Glide Glide 1.8 Glide 2 Glide 2.5	• A license purchase is required. • It is a homology modeling approach to dock ligands flexibly. By Examining the substructures of repeated	82%	• Adenosine A_{2A} receptor	• It is a special target-specific pose classifier trained to discriminate native-like from decoy poses.	[Kawatkar et al. 2009]

	molecules domain in similarity-based ligand binding pose prediction. It also ranks the docked compounds freely using a simple scoring function.				
GOLD (Genetic optimization for Ligand Docking) The version on CSC's Servers Taito-shell: 2018 Discovery Studio 2018 server	• It is an exhaustive search-based docking program which is efficiently protein flexible. It also provides free ligand docking.	59.8%	• Mycolyl transferase enzyme, Ag85C of M. tuberculosis	• Docking of phosphonate and trehalose analog inhibitors into the three-dimensional structure of Mycolyl transferase enzyme, Ag85C of *M. tuberculosis* was done by the use of GOLD software.	[Annamala et al. 2007] [Dautin et al. 2017]
ICM-Dock ICM 2.8	• Internal coordinate mechanics • Based on Monte Carlo methods • User can select whether protein is modeled rigidly on a grid or flexibly. • Ligand is	76%	• Validation carried out on Astex and CCDC (protein coding) dataset.	• Study on virtual ligand docking. • Selection of a surface model suggests a conformational search strategy, which then implies how to rank ligands to be pursued for further study.	[Bursulaya et al. 2003] [de Graaf et al. 2005]

	explicitly modeled in torsion space.				
iGEMDOCK GemDock 1.0 GemDock 2.0 GemDock 2.1	• It is freely available and generates pharmacophores that utilize a genetic algorithm.	79%	• HER2 of oral cancer	• Docking studies were performed for natural compounds (ligands) from the plant *Limonia acidissima* with HER2 of oral cancer by using iGEMDOCK suite.	[Hsu et al. 2011] [Raj and Krishna 2014] [Glaab 2015]
Leadfinder Leadfinder 2.3 3.8	• It is freely available on the internet but no. of leads is restricted after ten times. • It also provides correct energy-ranking of docked ligands poses, accurate binding energy predictions, and correct rank-ordering of active and inactive compounds in virtual screening experiments.	80-96%	• PPI inhibitor complex	• Binding energies for 330 diverse protein−ligand complexes yielding rmsd of 1.50 kcal/mol. The accuracy of ligand docking was assessed on a set of 407 structures.	[Strogano v et al. 2008] [Smith et al. 2011]

Ledock Ledock Mac/Windows/Linux Leprosy-Mac Lewater-Mac	• It is freely accessible and provides a graphic environment for virtual screening, docking, and post-screening analysis. • LeDock is flexible small-molecule docking software, which performs an exhaustive search of position, orientation and conformation of a ligand in the active site of a protein.	80.8 %	• Anti-tumor protein	• 2002 protein-ligand complexes with high-resolution crystal structures and experimental binding affinity data were selected from the refined set of PDB bind.	[Wang et al. 2016] [Chen et al. 2017] [Li et al. 2018]
LigDockCSA	• Freely available combines a highly efficient search method - Conformational Space Annealing (CSA) - with a scoring function based on the AutoDock energy function with a piecewise	89.4%	• Astex diverse set	• The performance of LigDockCSA was tested on the Astex diverse set which consist 85 protein-ligand complexes. Comparative study shows that LigDockCSA finds best scoring poses	[Shin et al. 2011]

	linear potential (PLP) torsional energy. Conformational space annealing is designed to search over broad ranges of conformational space, generating numerous local minima before arriving at the global minimum free energy conformation.			for native structure at 84.7% where Autodock and gold has 81.7% and 80.5% respectively.	
Ligin	• This program uses surface complementarity approach for predicting the structure of ligand-receptor complex.	Not available	• Docking • Methyl a-D-Arabinofuranoside to Concanavalin A (T0013) • Pentamidine to Trypsin (T0033) • SBB Inhibitor to Pancreatic Elastase (T0036) • Protein residue-Tyr12 Asn14 Gly98	• CASP2 tested for predicting the binding pocket location, ligand orientation and major interactions stabilizing the ligand-receptor complex.	[Sobolev et al. 1997] [Sable and Jois 2015]

MADAMM (MADAMM 8	• MADAMM considered protein flexibility initially using rotamer libraries to produce several combinations of conformers involving the most important • Allows flexibilization of both the receptor and the ligand during a multi-staged docking with an automated molecular modeling protocol. • Amino acid residues at the active-site.	90%	• Binding and recognition of polysaccharides to the carbohydrate-binding modules (CBMs) also known as cellulose. • Protein residues- Asp99 Arg126 Asp128 Asp146	• 1000 target structures, implicitly accounting for protein flexibility. • The program then automatically docks the ligand against each of these target structures using a standard docking program that treats the ligand as flexible, with the current version using GOLD.	[Cerqueira et al. 2009]
MolDock	• Based on search algorithm that combines differential evolution with a cavity prediction algorithm.	87%	• HIV-1 reverse transcriptase with phytochemicals	• Used 77 complexes for checking docking accuracy	[Thomsen and Christensen 2006]

MS-DOCK	• Multi conformation rigid body docking approach. • This program can be used as the first step of a multi-step docking/scoring protocol.	75-90%	• Seven target proteins- Ribonuclease A (RNAs), Coagulation factor X (FX), Estrogen receptor (ER), CDK2 (CDK), Thymidine kinase (TK), Carboxypeptidase A (CBXpe) and Neuraminidase (NA) • With different binding site properties for its ability to retrieve 65 known inhibitors in a library of 37970 drug-like compounds.	• The performance of MS-DOCK was additionally validated through a comparison with the commercial program OMEGA for multi-conformer generation and the program FRED for rigid-body docking (i.e., in this study we used FRED as a shape complementarity filter not for a full screening procedure).	[Sauton et al. 2008] [Pagadala et al. 2017] [Singh et al. 2011]
PhDock	• Based on multiple copy simultaneous search(MCSS) • To determine target-based theoretical pharmacophore	87%	• 1(HIV1) Protease structure is used with PhDock to dock a set of HIV-1 protease	• Study on MCSS2SPTS to reproduce the three-dimensional pharmacophoric features of ligands from	[Cross et al. 2009] [Sastry et al. 2013] [Li et al.

	s.		ligands • The docked poses are compared to the correspondi ng complex structures of the ligands.	known ligand–protein complex structures.	2014]
PLANTS (Protein-Ligand ANT System)	• PLANTS is available free of charge for academic users. This program is based in Ant Colony Optimization (ACO), a methodological approach that protein-ligand Docking in the new millennium is based on the behavior of real ants on finding the shortest path between their nest and a food source.	84%	• Study Astex diverse set	• The program has been used to generate 87% of astex diverse set complexes while 77% shown CCDC/Astex with RMSD deviations of less than 2 angstrom with respect to the experimentally determined structures.	[Korb et al. 2009] [Elokely and Doerksen 2013]
PSI-DOCK (Pose-Sensitive Inclined)	• The program uses a tabu-enhanced genetic algorithm (TEGA) with a shape complementary scoring	74%	• 21 different conformatio ns of HIV-1 protease	• The program was also shown to be able to reproduce the binding energy of a training set of 200 protein–ligand complexes with	[Pei et al. 2006] [Guedes et al. 2014] [Li et al. 2011]

	function to explore in a first step the potential binding poses of the ligand. • The predicted binding poses are then optimized through a competition genetic algorithm and evaluated through a specifically developed improved scoring function (SCORE) to determine the binding pose with the lowest docking energy.			a correlation coefficient of 0.788 and a standard error of 8.13 kJ/mol, while in a test set of 64 complexes a correlation coefficient of 0.777 and standard error of 7.96 kJ/mol were obtained. • All protein hydrogen atoms and the flexibility of the terminal protein atoms are intrinsically taken into account in PSI-DOCK.	
PSO@AUTODOCK (Particle Swarm Optimization)	• Fast, efficient protein ligand docking program. • This program based on swarm optimisation. • It is designed for analysis of highly flexible ligands. • Particle Swarm	66%	• Study 21 different conformations of the HIV-1 protease	• 10-fold decrease in the number of steps required for identification of the local minimum in comparison with SODOCK, and a 60-fold decrease when comparing with AutoDock 3.	[Namasivayam et al. 2007] [Bello et al. 2013] [Lin 2011]

	Optimization (PSO) algorithms varCPSO and varCPSO-ls are suited for rapid docking of highly flexible ligands.			• These results make PSO@AUTODOCK a very promising alternative for flexible ligand docking, and enable the inclusion of ligand flexibility in virtual screening campaigns of reasonably sized libraries comprising several thousands of compounds.	
PyMOL PyMOL 2.1.1 PyMOL 1.4.1	• It is freely available on the internet but no. of leads are restricted after the input is done ten times. • PyMOL plugins give a GUI application incorporating individual academic package designed for protein preparation (Reduce and	64.4 %	• Not available	• 5 InhA inhibitors were taken whose bioactive conformations are known, sequentially docked in the substrate cavity of each protein.	[Seeliger and de-Groot 2011] [Zhang et al. 2013] [Wilson and Lill 2011]

	AMBER Package), leading molecular mechanics applications (AMBER package), and docking and scoring (SLIDE and AutoDock Vina).				
PyRX PyRX 0.9.6	• It is a free open source and is based on SBVS. PyRx includes an embedded Python Molecular Viewer (ePMV) for visual analysis of results, as well as a built-in SQLite database for result storage.	Not available	• Aromatase inhibitor	• For studies output of PyRX compared to X-ray structures to examine the binding mode prediction.	[Dallakyan and Olson 2015] [Saeed et al. 2017] [Prieto-Martinez et al. 2018]
PythDock	• Uses Python programming language with a simple scoring function and a population based search engine. Function includes electrostatic	Not available	• MECL-1 binding with luteolin	• Exploring the potential of herbal ligands toward multidrug-resistant bacteria pathogens by computational drug discovery.	[Chung et al. 2011] [Pettinari et al. 2006]

	and dispersion/repulsion terms only, together with a search algorithm based on the particle swarm optimization method. The program is a rigid protein-ligand docking program, in the sense that treats ligands and proteins with fixed conformations.				
Q-Dock	• Low-resolution flexible ligand docking program with pocket-specific threading restraints models. Q-Dock describes both the ligand and the protein in a reduced representation mode.	Not available	• Study 23 protein ligand complexes for computation. • 1aaq-psi • 1apt-iva • 1epo-mor • 1apu-iva	• Ligand flexibility is accounted for through an ensemble docking of pre-calculated discrete ligand conformations with Replica Exchange Monte Carlo (REMC). • A database of 206 X-ray structures used for the experimentation on self-docking approach	[Brylinski et al. 2008]

				commonly used for the standardization of protein-ligand docking approach.	
rDock rDock 3.0	• Can be accessed free on the internet but license purchase is required for the full version. It is computationally efficient and achieves optimal performance initially for RNA (nucleic acids) targets now for protein targets as well.	78%	• Viral structural proteins	• The CCDC-Astex diverse Set of 85 complexes of protein-ligand specify for comparative study on binding mode prediction.	[Ruiz-Carmona et al. 2014] [Li et al. 2003
RosettaLigand	• It provides free accessibility but license purchase is required. • This tool leverages the Rosetta energy function and side chain repacking algorithm to account for flexibility of all	64%	• Membrane protein CASPIII	• It has been shown to successfully fold only small, soluble proteins (fewer than 150 amino acids), and it performs best if the proteins are mainly composed of secondary structural	[Meiler and Baker 2006] [Davis and Baker 2009] [Combs et al. 2013]

	side chains in the binding site.			elements (α-helices and β-strands). • Structures of helical membrane proteins between 51 and 145 residues were predicted to within 4 Å of the native structure, but only very small proteins (up to 80 residues) have been predicted to atomic-detail accuracy.	
SANDOCK	• Uses point complementary method. • Based on shape and chemical complementarity between interacting molecules. • For shape recognition uses FFT algorithm. • Guided matching algorithm.	74%	• X-ray structure of thrombin-ligand complex predicted	• Newly developed docking program can efficiently screen very large databases in a reasonable time and has been used to successfully identify novel ligands like the binding of a ligand to thrombin show RMSD of 0.7Å.	[Burkhard et al. 1999] [Detering and Varani 2004]

SLIM SLIM21* SlimDrivers 2.3.1.0	• It is freely accessible software and is used to predict binding poses in protein-small molecule complexes. It combines rotational and translational adjustments in a single step.	Not available	• Not available	• 40 proteins/ligands.	[Lee et al. 2012]
SOFT Docking	• A new approach that combines an ab-initio docking calculation and the mapping of an interaction site using chemical shift variation analysis.	72 %	• Cytochrome c553-ferrodin complex structural model is used for experimental studies.	• Study on T4 lysozyme and aldose reductase for identifying conformation changes on ligand binding. • Soft docking calculation, were tested experimentally for enzyme inhibition and four of these six inhibited the enzyme, the best with an IC_{50} of 8 μM.	[Ferrari et al. 2004]
Surflex Dock	• Surflex-Dock increases its robustness, particularly with respect to screening	85-95 % (Jain and Vertex	• Receptor protein-Deoxycytidine kinase ligand – gemcitabine	• A diverse set of 85 protein-ligand complexes and virtual screening	[Jain et al. 2007]

	effectiveness. • Surflex-Dock allows sensitive control of the use of the placed molecular fragment.	bench mark) ;	.	performance is reported on the DUD (Directory of Useful Decoys) set of 40 protein targets.	
SYMMDOCK	• Available free to academics. • Used in the prediction of cyclically symmetric Homo multimers.	85%	• C-5 symmetric Shiga toxin	• On a non-redundant docking benchmark of 213 Cn targets and 35 Dn targets.	[Schneidman-Duhovny et al. 2005] [Yan et al. 2018]
VoteDock	• It is a consensus docking method for prediction of protein-ligand interaction.	Not available	• Lymphoid specific tyrosine phosphatase inhibitors using multiple crystal structure.	• Extensive benchmark dataset of 1300 protein–ligands pairs were taken and compare its ability of scoring and posing.	[Plewczynski et al. 2011]
VSDocker	• VSDocker provides automation of all virtual screening steps, as well as ligand and receptor preparation, docking and analysis of results.	Not available	• Not available	• Not available	[Prakhov et al. 2010]

	• VSDocker works both on multiprocessor computing clusters as well as multiprocessor workstations operated by Windows, thus makes execution of virtual screening tasks even on a single high-performance multicore desktop that may be found nearly in each laboratory.				
YUCCA	• Based on an efficient heuristic for local search, for rigid protein–small-molecule docking.	76%	• Not available	• 100-complex benchmark, using the conformer generator OMEGA to generate a set of low-energy conformers.	[Choi et al. 2005]

Table 6: Different experimental studies (In vivo/In vitro) using docking strategies to find the potent inhibitor (virtual screening).

OBJECTIVE OF THE STUDY	TARGET PROTEIN	DOCKING APPROACH	RESULTS	IN VITRO VALIDATION	REFERENCES
To find novel inhibitors of mycosin protease-1, involved in the virulence of drug resistant *Mycobacterium tuberculosis*.	Mycosin protease 1 of *Mycobacterium tuberculosis.*	485,000 ligands were subjected to LBVS (Ligand-Based Virtual Screening) and SBVS (Structure-Based Virtual Screening).	2 compounds were found to inhibit the activity of the enzyme by more than 40%.	Yes	[Hamza et al. 2014]
To find an inhibitor which could stop anthrax?	LF (lethal factor) protein of a *Bacillus anthracis* exotoxin.	25,595 screened compounds from DrugBank and vendor databases were docked using Surflex-Dock.	5 compounds were found with IC50 values less than 100 microM.	Yes, by LF FRET assay	[Vitale et al. 2000]
To find the inhibitory activity of AchE, involved in nerve impulse transmission.	Electric eel Acetylcholinesterase (AchE) derived from *Electrophorous electricus*	Virtually screened with 157,000 compounds using the docking	35 compounds showed inhibitory activities with IC50	Yes	[Pradelles et al. 1985]

	(Electric eel).	algorithm, ADAM&EVE.	values of less than 100 microM.		
To find proteins that could block the binding site of AI-2.	Auto-inducer2 (AI-2) of *Vibrio harveyi*, which binds to a receptor protein, LuxP.	7 million compounds were virtually screened (through docking) using DOCK 5.4	5 compounds were found to show IC50 values at micromolar concentrations.	Yes, using Quorum Sensing Assay	[Li et al. 2007]
To find inhibitors of PNP in order to combat lymphoproliferative disorders, as well as to counter autoimmunity.	Purine Nucleoside Phosphorylase (PNP) of calf spleen.	30,000 compounds. From Astex database were docked using GOLD	6 compounds showed inhibitory activity against the protein PNP.	Yes	[Miles et al. 1998]
To find AHAS inhibitors, involved in the biosynthesis of amino acids valine, leucine and isoleucine.	Acetohydroxyacid synthase (AHAS) of plant and yeast.	164,000 compounds were docked using DOCK 4.0	3 compounds were found to show inhibitory activity.	Yes, using AHAS Assay	[Pang et al. 2002]
To find inhibitors of ADAM 12, involved in cardiovascular disease.	ADAM 12 of humans.	Virtual screening of compounds from a database of 67, 062 molecules.	4 molecules showed IC50 values less than 50 nm.	Yes, through a cell-based activity assay.	[Gilpin et al. 1998]
To find out the inhibitors of histamine H4 receptor.	Human histamine H4 receptor.	8.7 million 3D structures of ligands were docked.	16 of them were found to possess significant activity, expressed in the term of ‘displacemen	Yes, using binding assay tests.	[Jablonowski et al. 2003]

			t' values.		
To find to activators of hPXR, involved in the upregulation of drug-metabolizing enzymes.	Human pregnane X receptor (hPXR)	496 compounds were from ChemBridge database were docked using Surflex v1.27.	4 molecules were found to be activators of the target protein.	Yes	[Urquhart et al. 2007]
To find molecules that can bind to Human PIM-1, a potential anticancer target.	Human PIM-1 (Proviral Integration site for MuLV (murine leukemia virus)) kinase.	Docking of 7, 00,000 compounds using GLIDE.	4 compounds were found to show noticeable activity.	Yes	[Pierce et al. 2008]
To find inhibitors of enzymes involved in protein arginine methylation.	Protein arginine methyltransferases (PRMTs).	Using GOLD, 6,232 molecules were docked into the binding pocket of hPRMT1.	9 compounds showed IC50 values < 50 microM, while 4 showed values less than 16 microM.	Yes	[Zhang et al. 2000]
To find active compounds, that have the ability to bind to FFAR-1, a receptor for medium and long chain frees fatty acids and may be involved in the metabolic regulation of insulin secretion.	Free Fatty Acid Receptor-1 (FFAR-1)	70,477 compounds with FFAR-1, using GLIDE.	6 were found to be active compounds.	Yes	[Miyauchi et al. 2010]
To find molecules that can bind to CDC25 phosphatases, a	CDC25 phosphatases	Using FRED, Surflex, and LigandFit, the target protein	99 compounds were able to inhibit the	Yes	[Montes et al. 2008]

proposed target in anticancer treatments.		was virtually screened with 310,000 molecules.	CDC25B activity at 100 microM.		
To find out the inhibitor of, Src kinase, this is associated with tumor growth and development.	Src kinase	61,000 molecules were docked using LIGANDFIT software.	4 molecules were found to satisfy the condition at 10 microM concentration .	Yes, using Src inhibition assay	[Lee et al.,2009]
To find out the inhibitor of LDH, an important enzyme in the parasite's glycolytic machinery.	Lactate dehydrogenase (LDH)	50 analogs were docked with Molegro Virtual Docker.	3 compounds with the best binding energies showed IC50 values ranging from 13 microM to 2.6 microM.	Yes, using ELISA	[Gomez et al. 1997]
To find out the inhibitor of FP-2.	*Plasmodium falciparum* falcipain-2 (FP-2)	Docked 80 000 compounds in the SPECS database using GLIDE and GAsDock.	28 were found to have IC50 values ranging from 2.4 to 54.2 microM.	Yes	[Hogg et al. 2006].
To find out the inhibitors of Alpha-glucosidase, an effective inhibition target in the case of Type 2 diabetes mellitus	Alpha-glucosidase	Docked 40 natural compounds	3 were found to perform as effective inhibitors, each with IC50 values less than 100 microM.	Yes	[van de Laar et al. 2005]
Inhibition study on ALR2, having a role in diabetes mellitus.	Aldose reductase (ALR2)	1261 compounds were docked using FlexX.	9 compounds were selected for further characterizati	Yes	[Maeda et al. 1999]

			on.		
To find the inhibitor of 17L core proteinase, involved in the replication of vaccinia, variola and monkeypox viruses.	17L core proteinase	Using ICM docking algorithm, around 230,000 thio-acyl intermediates were docked.	6 were found to show IC50 values less than 50 microM.	Yes, through a cleavage assay	[Byrd et al. 2002]
To find the anti-HIV-1 RT inhibitor.	HIV-RT	2800 compounds were filtered using FILTER (version 2), = OMEGA (2.1.0) generated a minimum of 23 conformers per molecule, which were then docked to the target protein.	Out of the top 20 binding poses, only 6 were available.	Yes	[Ravindra et al. 2005]
To find new structural targets of DNA gyrase, involved in bacterial DNA replication.	DNA gyrase	139,644 compounds were docked using DOCK 5.1.0.	3 diverse compounds showed activity against the enzyme.	Yes	[Ostrov et al. 2007]
To find molecules that can bind to EphB4.	Human hepatocellular carcinoma receptor tyrosine kinase B4 (EphB4)	728,202 compounds were subjected to flexible ligand docking.	2 compounds were found to be effective.	Yes	[Lafleur et al. 2009]
To find the inhibitor of EF.	Anthrax edema factor (EF).	10,000 compounds from the ZINC	3 compounds were found to show inhibition of	Yes	[Klimpel et al. 1994]

		database were docked using AutoDock 3.	the protein in the ranges of 1.7-9 microM (IC50 values).		
To find the binding ligands of GPCRs, used as a drug target.	G-Protein Coupled Receptors (GPCRs)	The ICM docking software was used to dock 187,084 compounds.	6 were found to show activity.	Yes	[Shoichet et al. 2012]
To find the inhibitor of TRH-R1	TRH-R1 (Thyrotropin Releasing Hormone receptor, isotype 1)	10,000 compounds were docked using FlexE.	1 molecule was found to be the most potent inhibitor of TRH-R1 at a Ki of 0.29 microM.	Yes	[Engel et al. 2008].
To find the inhibitors of DNA gyrase	DNA gyrase of *Mycobacterium tuberculosis*	Gatifloxacin analogs were docked using Molegro Virtual Docker.	One compound was found to perform the best among the 8 studied		[Sriram et al. 2006]
To find out ligands, that could bind to PPAR-γ, an important drug target for regulating glucose metabolism.	Peroxisome proliferator activated receptor- γ (PPAR-γ)	Used 2,4-thiazolidinediones (TZD) and chromone conjugates, a total 19 of them, and docked them with the PPAR- γ target using Schrodinger Glide software.	7 compounds, of the total 19, showed the most promising docking scores.	Yes	[Ricote et al. 1998]

To search for inhibitors for these kinases, which regulate GPCRs.	Kinases such as cAPK (cAMP-dependent Protein Kinase) and GRK (G-protein coupled receptor kinases).	The search was performed using DOCK 3.5, and a database of 13,028 compounds.	With respect to GRK2 inhibitors, 5 had IC50 values below 100 microM while cAPK inhibitors had IC50 values of less than 100 microM.	Yes	[Sugden et al.1995]

Table 7: List of molecular docking programs and their salient features.

TOOLS	FEATURES	CITATIONS	REFERENCES
AutoDock Vina	• AutoDock Vina is a program for molecular docking and virtual screening. AutoDock Vina achieves degree roughly two orders of magnitude speed-up compared with the Autodock-4. • AutoDock Vina utilizes an iterated local search global optimizer. It is free for academic use. The tool is maintained by the Molecular Graphics Laboratory, The Scripps Research Institute, La Jolla. **Benchmark:** • Top ranking tool: GOLD and LeDock had the best sampling power (GOLD: 59.8% accuracy for the top scored poses; LeDock: 80.8% accuracy for the best poses) and AutoDock Vina had the best scoring stats (rp/rs of 0.564/0.580 and 0.569/0.584 for the top scored poses and best poses).	1187565	[Wang Z et al. 2016] [Trott et al. 2010]
AutoDockTools	• AutoDockTools facilitates formatting of input molecule files, with a set of methods that guide the user through protonation, calculating charges, and specifying rotatable bonds in the ligand and the protein. • To change the planning and preparation of	6899	[Morris et al. 2009]

	docking experiments, it permits the user to identify the active site and determine visually the volume of space searched within the docking simulation. • It includes a range of novel ways for clustering, displaying, and analyzing the results of docking experiments. **Benchmark:** • Top ranking tools: Vina performs notably well within the docking power evaluation, that measures the ability of a scoring function to differentiate decoys from the native pose.		
PatchDock	• PatchDock performs structure prediction of protein–protein and protein–small molecule complexes. • The inputs given to the servers were either protein PDB codes or uploaded protein structures. The services are available at http://bioinfo3d.cs.tau.ac.il . • The strategies behind the servers were very efficient, allowing large-scale docking tests. • PatchDock is an efficient rigid docking method that maximizes geometric shape complementarity.	1447	[Schneidman-Duhovny et al. 2005]
SymmDock	• The SymmDock technique predicts the structure of a homomultimer with cyclic symmetry given the structure of the monomeric unit. • The methods behind the server were very efficient which allows large-scale docking experiments.	1447	[Schneidman-Duhovny et al. 2005]

	• The user interface of SymmDock (http://bioinfo3d.cs.tau.ac.il/SymmDock/) is even easier than that of PatchDock, since the input here consists of only one molecule and also the symmetry order. • If the arrangement of the input monomers in its native complex follows a different kind of symmetry, then SymmDock would not be appropriate for such a prediction.		
MolDock	• MolDock is based on heuristic search algorithm that combines differential evolution with a cavity prediction algorithm. • Docking scoring function of MolDock is an extension of the piecewise linear potential (PLP) which include hydrogen bonds and electrostatic bonds. • To further improve docking accuracy, a re-ranking scoring function was introduced, which identified the most promising docking solutions from the information provided by the docking algorithm. **Benchmark:** • MolDock has greater accuracy than surflex, glide, flexX and GOLD. • Dataset: the author utilized flexible ligands of 77 protein targets.	1339	[Thomsen et al. 2006]
AutoDock	• AutoDock is a suite of C programs designed to predict the bound conformations of a small, flexible ligand to a macromolecular target of known structure.	9042	[Goodsell et al. 1998]

	• The technique combines Monte Carlo simulated annealing, a traditional genetic algorithm and the Lamarckian Genetic Algorithm for conformation searching with rapid grid-based methods of energy evaluation. • The AutoDock scoring function is a subset of the AMBER force fields that evaluates molecules using the United Atom model. **Benchmark:** • EADock and ICM were better than AutoDock for information set of thirty seven crystallized protein–ligand complexes that features 11 different proteins.		
ZDOCK	• ZDOCK is a Fast Fourier Transform based docking algorithm • This tool performs a full rigid-body scanning of docking orientations between two proteins. The version, 3.0.2, includes performance optimization and a novel pair wise statistical energy potential. • Since its initial implementation, the ZDOCK Server has experienced major changes to improve its docking performance, functionality and user interface. These include upgrading the docking algorithm from ZDOCK 2.3 to ZDOCK 3.0.2, resulting in more prominent accuracy and highly efficient searching. **Benchmark:** • ZDOCK 3.0 showed vast upgrades in its predictive ability versus the previous version when tested on a protein-protein	1143 448	[Chen et al. 2003] [Brian G. Piercem 2014]

	docking benchmark. • ZDOCK 2.3 has a faster (almost thrice) average running time on the docking benchmark versus ZDOCK 3.0; ZDOCK 2.3.2 was twice as fast as ZDOCK 3.0.2 during the evaluation tests.		
ClusPro	• ClusPro is a fast algorithm for filtering docked conformations with good surface complementarity, and rank them based on their clustering properties. • The free energy filters select complexes with minimal desolvation and electrostatic energies. • Clustering has been used to smooth the local minima and to choose the ones with the broadest energy wells—a property related with the free energy at the binding site. **Benchmark:** • SwarmDock demonstrated better performance than that of ClusPro.	791	[Comeau et al. 2004]
EADock	• EADock DSS engine is combined with setup scripts for curating common problems and for preparing the target protein and the ligand input files. • EADock was able to identify binding modes with high accuracy. The accuracy is necessary to compute the binding free energy of the ligand. **Benchmark:** • Dataset used: 37 crystallized protein–ligand complexes featuring 11 different proteins	547	[Grosdidier et al. 2011]

	• The average RMSD between the best clusters was predicted by EADock and crystal structures was 0.75 A˚. This was significantly better than what was reported for ICM (1.04 A˚), AutoDock (2.46 A˚), GOLD (3.31 A˚), FlexX (3.85 A˚), and DOCK.		
SwissDock	• SwissDock is a web-server program dedicated to the docking of small molecules on target proteins. • It is based on the EADock DSS engine, combined with setup scripts for curating common issues and for preparing both the target protein and the ligand input files. • The structure of the target protein, as well as that of the ligand, could be automatically prepared for docking using SwissDock **Benchmark:** • SwissDock shows higher performance than AutoDock4 and has a greater binding affinity.	550	[Grosdidier et al. 2011]
GEMDOCK	• GEMDOCK utilizes a Generic Evolutionary Method for molecular docking and an empirical scoring function. The former combined both discrete and continuous global search strategies with local search strategies to speed up convergence, whereas the latter result in rapid recognition of potential ligands. • GEMDOCK was experimented on a diverse dataset of 100 protein–ligand complexes from the Protein Data Bank. • GEMDOCK had been a useful tool for molecular recognition and may be used to	392	[Yang et al. 2004]

	systematically assess and thus improve scoring functions. **Benchmark:** ● Average RMSD value ranged from 4.74Å to 12.63Å. GEMDOCK displayed better performance than GOLD.		
RosettaDock	● The RosettaDock server identifies low-energy conformations of a protein–protein interaction near a given starting configuration by optimizing rigid-body orientation and side-chain conformations. ● It can generates 1000 independent structures, and the server returns pictures, coordinate files and detailed scoring information for the 10 top-scoring models. ● RosettaDock was approved on the docking benchmark set and through the Critical Assessment of Predicted Interactions blind prediction challenge. **Benchmark:** ● The benchmark consisted of a diverse set of 116 docking targets including 22 antibody-antigen complexes, 33 enzyme-inhibitor complexes, and 60 'other' complexes. ● The tool performed better in comparison to Docking Benchmark 3.0.	363	[Lyskov et al. 2008]
FireDock	● The FireDock web server is used for flexible refinement and scoring of protein–protein docking solutions. It includes optimization of side-chain conformations, rigid-body orientation and permits a high-throughput refinement. ● The server provides a user-friendly	335	[Mashiach et al. 2008]

	interface and a 3D visualization of the results. A docking protocol comprise of a global search by PatchDock and a refinement by FireDock was extensively tested. ● The protocol was successful in screening and scoring docking solution candidates for cases taken from docking benchmarks. ● They provide an alternate for using this protocol by automatic redirection of PatchDock candidate solutions to the FireDock web server for refinement. **Benchmark:** ● It permits a high-throughput refinement of up to 1000 solution candidates. The technique simultaneously targets the problem of flexibility and scoring of solutions produced by fast rigid-body docking algorithms. ● FireDock succeeded in positioning a near-native solution in the top 15 predictions for 83% of the 30 enzyme–inhibitor test cases and for 78% of the 18 semi-unbound antibody–antigen test cases.		
INVDOCK	● Inverse-docking approach (INVDOCK) can be used for finding potential protein targets of a small molecule by the computer-automated docking search of a protein cavity database. ● Results on two therapeutic agents, 4H-tamoxifen and vitamin E, demonstrated that 50% of the computer-identified potential protein targets were confirmed by experiments.	318	[Chen et al. 2001]

	• The application of this methodology may facilitate the prediction of unknown and secondary therapeutic target proteins and those related to the side effects and toxicity of a drug or drug candidate. • INVDOCK have been developed as a tool for searching putative protein and nucleic acid targets of a drug • Results for a number of therapeutic drugs demonstrated the applicability of INVDOCK. • INVDOCK has potential application in probing molecular mechanism of bioactive Chinese natural products(CNP) as well as in facilitating the prediction of unknown therapeutic and side effect and toxicity targets of drugs and drug candidates protein targets of several active CNPs was used.		
RDOCK	• The main component of RDOCK is a three-stage energy minimization scheme, followed by the assessment of electrostatic and desolvation energies. • Ionic side chains were kept neutral in the initial two stages of minimization, and reverted to their full charge states in the last stage of brief minimization. • Without side chain conformational search or filtering/clustering of resulting structures, RDOCK represents the simplest methodology toward refining unbound docking predictions. • RDOCK is a molecular docking program that was developed at Vernalis for high-throughput VS (HTVS) applications. This	288 114	[Li et al. 2003] [Ruiz-Carmona et al. 2014]

	program was evolved from RiboDock and could be used against proteins and nucleic acids. It was designed to be computationally very efficient and allows the user to incorporate additional constraints and information as a bias to guide docking.		
TarFisDock	• TarFisDock is a web-based tool for automating the procedure of searching for small molecule–protein interactions over a large collection of protein structures. • It offered PDTD (Potential Drug Target Database), a target database containing 698 protein structures covering 15 therapeutic zones and a reverse ligand–protein docking program. • It is a useful tool for target identification, mechanism study of old drugs and probes discovered from natural products. • TarFisDock is a web server that identifies drug targets using a reverse docking strategy to seek all possible binding proteins for a given small molecule. • TarFisDock identifies potential targets for a compound with known biological activity, a newly isolated natural product or an existing drug whose pharmacological mechanism was unclear. In addition, this platform was also able to find potential targets that could be responsible for the toxicity and side effects of a drug, which could allow for the prediction of the side effects of a drug candidate. **Benchmark:** • The effectiveness of different docking strategies in multiple targets identification	280	[Li et al. 2006]

	is unclear. ● Five inverse docking schemes were evaluated to find out the most effective method in multiple targets identification. A target database containing a highly qualified dataset that composed of 1714 entries from 1594 known drug targets covering 18 biochemical functions was gathered as a testing pool for inverse docking. ● The inverse docking engines including GOLD, FlexX, Tarfisdock and two in-house target search schemes TarSearch-X and TarSearch-M were assessed by eight multiple target systems in the dataset. ● Their resulted demonstrated that TarSearch-X was the most effective method in multiple targets identification and validation in a given situation.		
pyDock	● pyDOCK is a program which was implemented in order to check the scoring of rigid-body docking poses. ● The scheme is based on Coulomb electrostatics with distance dependent dielectric constant, and implicit desolvation energy with atomic solvation parameters previously adjusted for rigid-body protein–protein docking. This scoring function was not highly dependent on specific geometry of the docking poses and therefore could be used in rigid-body docking sets generated by a variety of method. ● pyDockWEB server is a web application for the use of the protein–protein docking and scoring program pyDock. Users can	87	[Jiménez-García et al. 2013]

	easily send pyDock jobs to be executed in a five-step process via a user friendly front-end.		
FlexPepDock	• Using the Rosetta fragments library and a coarse-grained structural representation of the peptide and the receptor, FlexPepDock *ab-initio* samples efficiently and simultaneously the space of possible peptide backbone conformations and rigid-body orientations over the receptor surface of a given binding site. • The subsequent all-atom refinement of the coarse-grained models includes full side-chain modeling of both the receptor and the peptide, resulting in high-resolution models in which key side-chain interactions were recapitulated. **Benchmark:** • The validation on a representative benchmark set of crystallographically solved high-resolution peptide-protein complexes demonstrates significantly improved performance over all existing docking protocols. This opened up the way to the modeling of many more peptide-protein interactions, and to a more detailed study of peptide-protein association in general. • Rosetta FlexPepDock web server provides an interface to a high-resolution peptide docking (refinement) protocol for the modeling of peptide–protein complexes, implemented within the Rosetta framework. • Given a protein receptor structure and an approximate, possibly inaccurate model of the peptide within the receptor binding	212	[London et al. 2011]

	site, the FlexPepDock server refines the peptide to high resolution, allowing full flexibility to the peptide backbone and to all side chains.		
FlexDock	• The flexible docking algorithm, FlexDock, is unique in its ability to handle any number of hinges in the flexible molecule, without degradation in run-time performance, as compared to rigid docking. • The algorithm for reconstruction of cyclically symmetric complexes successfully assembles multimolecular complexes satisfying C_n symmetry for any n in a matter of minutes on a desktop PC.	190	[Schneidman-Duhovny et al. 2005]
DOCK Blaster	• The method requires a PDB code, sometimes with a ligand structure, and from that alone could launch a full screen of large libraries. • A critical feature of this program was self-assessment, which estimated the anticipated reliability of the automated screening results using pose fidelity and enrichment.	199	[Irwin et al. 2009]
Misdocked	• Proteins are misdocked because water molecules or ions are not included in the receptor model. Uncertainty in the ionization state of the ligand or the receptor, due to receptor-induced (ligand-induced) pKa changes in the ligand (receptor). They are also misdocked because of insufficient sampling or they are docked correctly, but they do not score properly because of failures in the scoring function. The first two reasons are related	230	[Verkhivker et al. 2000]

	to rearrangements of the binding pocket upon ligand binding. ● Misdocked predictions in ligand-protein docking were classified as `soft' and `hard' failures. While a soft failure arises when the search algorithm is unable to find the global energy minimum corresponding to the crystal structure, a hard failure resulted from a flaw of the energy function to qualify the crystal structure as the predicted lowest energy conformation in docking simulations.		
MCDOCK	● MCDOCK was developed to carry out the molecular docking operation automatically. ● The particular version of the MCDOCK program (version 1.0) allows for the full flexibility of ligands in the docking calculations. ● The scoring function used in MCDOCK is the sum of the interaction energy between the ligand and its receptor, and the conformational energy of the ligand. ● MCDOCK can be used to predict the precise binding mode of ligands in lead optimization and to discover novel lead compounds through structure-based database searching. ● MCDOCK applies a multiple stage strategy to dock a flexible ligand to a rigid receptor.	266	[Liu et al. 1999] [R. D. Taylor et al. 2002]
FiberDock	● FiberDock models backbone flexibility by an unlimited number of normal modes ● The method iteratively minimizes the	147	[Mashiach, Efrat et al. 2010] [Alper

	structure of the flexible protein along the most relevant modes. The relevance of a mode was calculated according to the correlation between the chemical forces, applied on each atom, and the translation vector of each atom, according to the normal mode. ● The FiberDock server can refine up to 100 rigid-docking solution candidates. The user can upload PDB (Protein Data Bank) files, receptor and ligand, and provides a list of up to 100 transformations. ● For side-chain flexibility, Fiberdock uses a rotamer library and finds optimum combination of rotamers with the lowest total energy. **Benchmark:** ● FiberDock calculates several binding energy scores, including attractive and repulsive van der Waals forces, the atomic contact energy, partial electrostatics, hydrogen and disulfide bonds, π stacking, and aliphatic interactions. These scores were used as a feature vector to train a Random Forest Classifier (RFC) returning a single probabilistic score to assess whether two interacting proteins are biologically relevant. ● eRankPPI rearranged dimer models. In addition, FiberDock also produced accurate results. Further refinement procedure used by FiberDock yielded improvements for eRankPPI as compared to ZDOCK.		Baspinar et al. 2014]

PRODOCK	• PRODOCK is used for protein modeling and flexible docking. • It was based on a residue data dictionary that makes the programming easier and the definition of molecular flexibility more straight forward.	135	[Trosset et al. 1999]
iGemdock	• For post-screening analysis, *i*GEMDOCK provides biological insights by deriving the pharmacological interactions from screening compounds without relying on the experimental data of active compounds. • The pharmacological interactions represent conserved interacting residues, which often form binding pockets with specific physico-chemical properties, to play the essential functions of a target protein. The experimental results show that the pharmacological interactions derived by *i*GEMDOCK are often hotspots involving in the biological functions. **Benchmark:** Parameters such as Population size: 200, Number of generations: 70 and Number of solutions: 3 were selected. The anti-tumor compounds were sorted at the end of docking process based on their interaction energies and fitness values produced by the docking via iGemdock software. • Total 29 plant anti-tumor compounds were screened against the structure of FAT10 protein via iGemdock.	134	[Hsu et al. 2011]
LibDock	• LibDock had been applied to the GSK validation data set. LibDock is based on the algorithm developed by Diller and Merzand. It is one of the commercially	126	[Rao et al. 2007] [Miriam Sgobba et

	available docking programs that use protein binding site features to guide docking. ● The LibDock methodology was originally developed to handle the rapid docking of combinatorial libraries of compounds with the goal of prioritizing the selection of libraries rather than rank ordering the compounds themselves. ● The algorithm has four functional aspects: conformation generation of the ligands, creating a binding site image (hot spot identification), matching the binding site image and the ligand, and a final optimization stage and scoring. The binding site image consists of lists of polar and non-polar hot spots. These were generated by laying a grid in the binding site volume and then scoring a non-polar and polar probe at each grid point. **Benchmark:** ● Evaluated the performance of MM-PBSA and MM-GBSA scoring functions, implemented in post-docking procedure BEAR, in rescoring docking solutions. For the first time, the performance of this post-docking procedure was evaluated on six different biological targets (namely estrogen receptor, thymidine kinase, factor Xa, adenosine deaminase, aldose reductase, and enoyl ACP reductase) by using i) both a single and a multiple protein conformation approach, and ii) two different software, namely AutoDock and LibDock. The assessment was carried out on two of the most important criteria for the evaluation of docking methods, *i.e.*, the ability of known ligands to enrich the top		al. 2012]

	positions of a ranked database with respect to molecular decoys, and the consistency of the docking poses with crystallographic binding modes. It was found, in many cases, that MM-PBSA and MM-GBSA were able to yield higher enrichment factors compared to those obtained with the docking scoring functions alone. However, for only a minority of the cases, the enrichment factors obtained by using multiple protein conformations were higher than those obtained by using only one protein conformation.		
ASEDock	• ASEDock is a docking program based on a shape similarity evaluation between a concave portion (i.e., concavity) on a protein and the ligand. • Two concepts were introduced into ASEDock. One was an ASE model, which was characterized by the combination of alpha spheres produced at a concavity in a protein and the excluded volumes around the concavity. The other was an ASE score, which assesses the shape similarity between the ligand and the ASE model. • The ASE score chooses and refines the initial pose by maximizing the overlap between the alpha spheres and the ligand, and minimizing the overlap between the excluded volume and the ligand. • ASE score makes great utilization of the Gaussian-type function for assessing and optimizing the overlap between the ligand and the site model, it can pose a ligand onto the docking site moderately quicker and more effectively than using potential energy functions. The posing stage using	128	[Goto et al. 2008]

	the ASE score was followed by full atomistic energy minimization. ● The posing algorithm of ASEDock is free from any predisposition with the exception of shape; it is an exceptionally powerful docking technique. **Benchmark:** ● **Datasets used**: ● A validation study has demonstrated that ASEDock can faithfully reproduce experimentally determined docking modes of various drug-like molecules in their target proteins. ● Almost 80% of the structures were reconstructed within the estimated experimental error. The success rate of approximately 98% was accomplished based on the docking criterion of the root-mean-square deviation (RMSD) of non-hydrogen atoms (< or = 2.0 A). ● The uniquely high success of ASEDock in redocking experiments plainly demonstrated that the most important factor governing the docking process was shape complementarity.		
ConsDock	● ConsDock is a consensus docking approach that takes advantage of three widely used docking tools (Dock, FlexX, and Gold). ● The consensus analysis of all possible poses was generated by several docking tools was performed consecutively in four steps: (i) hierarchical clustering of all poses produced by a docking tool into	125	[Paul et al. 2002]

	families represented by a leading molecule (leaders); (ii) definition of all consensus pairs from leaders generated by various docking programs; (iii) clustering of consensus pairs into classes, represented by a mean structure; and (iv) positioning the different means beginning from the most populated class of consensus pairs. **Benchmark:** ● When applied to a test set of 100 protein–ligand complexes from the Protein Data Bank, ConsDock altogether outperformed single docking with respect to the docking accuracy of the top-ranked pose. ● In 60% of the cases, ConsDock was able to rank as top solution a pose within 2 Å RMSD of the X-ray structure. ● It can be applied as a post processing filter to either single- or multiple-docking programs to prioritize three-dimensional guided lead optimization from the most likely docking solution. ● Three different database docking programs (Dock, FlexX, Gold) have been utilized in combination with seven scoring functions (Chemscore, Dock, FlexX, Fresno, Gold, Pmf, Score) to survey the accuracy of virtual screening methods against two protein targets (thymidine kinase, estrogen receptor) of known 3-D structures. For both targets, it was generally possible to separate about 7 out of 10 true hits from a random database of 990 ligands. The use of consensus lists common to two or three scoring capacities clearly enhances hit rates among the top 5% scorers from 10% (single scoring) to 25-40% (double		

	scoring) and up to 65-70% (triple scoring). However, in all tested cases, no clear connections could be found between docking and ranking accuracies. Moreover, predicting the absolute binding free energy of true hits was impractical whatever docking accuracy was achieved and scoring function used.		
SODOCK	• SODOCK was developed to improve efficiency and robustness of particle swarm optimization (PSO). • This tool works as an optimization algorithm based on particle swarm optimization (PSO) for solving flexible protein-ligand docking problems. • PSO is a population-based search algorithm. It is very simple and efficient. SODOCK works cooperatively with the environment of AutoDock 3.05 as per the reports. **Benchmark:** • Benchmarking studies' outcomes revealed that SODOCK was superior to the Lamarckian genetic algorithm (LGA) of AutoDock, in terms of convergence performance, power, and obtained energy, especially for highly flexible ligands. • The outcomes also revealed that PSO was more suitable than the conventional GA in dealing with flexible docking problems with high correlations among parameters. • This experimentation also compared SODOCK with four state-of-the-art docking methods, namely GOLD 1.2,	132	[Chen et al. 2007]

	DOCK 4.0, FlexX 1.8, and LGA of AutoDock 3.05. It was revealed that SODOCK obtained the least RMSD value in 19 of 37 cases. The average (2.29 Å) of the 37 RMSD values of SODOCK, was reported to be better than those of other docking programs, which were all above 3.0 Å.		
DynaDock	• DynaDock is a docking tool that was used for docking peptides into flexible receptors. • For this purpose a two step procedure was created: first, the protein–peptide conformational space was scanned and approximate ligand poses were identified and second, the identified ligand poses were refined by a molecular dynamics based strategy: optimized potential molecular dynamics (OPMD). • The OPMD approach utilized soft-core possibilities for the protein–peptide interactions and applied an optimization scheme to the soft-core potential. • Comparison with refinement results obtained by conventional molecular dynamics and a soft-core scaling approach demonstrated significant upgrades in the sampling capability for the OPMD method. • The DynaDock method uses a soft-core molecular dynamics-based refinement. **Benchmark:** • Eight docking programs (DOCK, FLEXX, FRED, GLIDE, GOLD, SLIDE, SURFLEX, and QXP) that can be utilized for either single-ligand docking or	102	[Antes et al. 2010]

	database screening have been compared for their propensity to recover the X-ray pose of 100 small-molecular-weight ligands, and for their capacity to differentiate known inhibitors of an enzyme (thymidine kinase) from randomly chosen "drug-like" molecules. Interestingly, both properties were found to be correlated, since the tools showing the best docking accuracy (GLIDE, GOLD, and SURFLEX) are considered to be the most successful in positioning known inhibitors in a virtual screening experiment.		
RiboDock	• RiboDock® is a virtual screening system for automated flexible docking. Building on well-known protein-ligand scoring function establishments, features were added to describe the interactions of common RNA-binding functional groups that were not taken care adequately by conventional terms, to disfavour non-reciprocal polar contacts, and to control non-specific charged interactions. • rDock is a fast, versatile and open source program for docking ligands to proteins and nucleic acids. • It was intended for High Throughput Virtual Screening (HTVS) campaigns and binding mode prediction studies. **Benchmark:** • RiboDock found solution with RMSD<=3 in 5 cases out of 7, the original DrugScore RNA potential in 7 out of 9 cases, and MORDOR generated near-native poses in 11 out of 12 cases.	101	[Morley et al. 2004]
SwarmDock	• This server was validated in the CAPRI	113	[Torchała et

	blind docking experiment, against the last docking benchmark, and against the ClusPro docking server, the highest performing server which was available at that time. ● Subsequent to uploading PDB files of the binding partners, the server produces low energy conformations and returns a ranked list of clustered docking poses and their corresponding structures. ● The user can perform full global docking, or focus on particular residues that were involved in binding. ● The authors reported the mathematical model which intends to reduce the total cost of operations subjected to a set of constraints. Due to high complexity of model, the problem was solved by utilizing a variation of Particle Swarm Optimization (PSO) with a Self-Learning strategy, namely SLPSO. **Benchmark:** ● The previously published docking and affinity structural benchmarks were updated, increasing the number of cases by 31% and 24%, respectively. An updated and integrated version of their widely utilized protein–protein docking was presented and binding affinity benchmarks. Fifty-five new complexes were added to the docking benchmark, out of which 35 have experimentally measured binding affinities. These updated docking and affinity benchmarks contain 230 and 179 entries. Considering only the top 10 docking predictions per benchmark case, a prediction accuracy of 38% was achieved		al. 2013]

	on all 55 cases and up to 50% for the 32 rigid-body cases only.		
FLIPDock	• FLIPDock (Flexible LIgand–Protein Docking) is protein-ligand docking software which allows the automated docking of flexible ligand molecules into active sites of flexible receptor molecules. • In FLIPDock, conformational spaces of molecules are encoded using a data structure that was developed recently known as the Flexibility Tree (FT). • Conformational changes of biological macromolecules when binding with ligands have long been observed and remain a challenge for automated docking techniques. • While the Flexibility Tree can represent fully flexible ligands, it was initially designed as a hierarchical and multi-resolution data structure for the selective encoding of conformational subspaces of large biological macromolecules. • These conformational sub-spaces can be built to span a range of conformations that are important for the biological activity of a protein.	97	[Zhao et al. 2007]
FRODOCK	• FRODOCK (Fast ROtational DOCKing) is a novel docking methodology based on FRM (Fast Rotational Method) to perform protein-protein docking. • In contrast to other approaches, FRODOCK has the advantage of combining the capability to express the interaction terms into 3D grid-based potentials with the efficiency of a	92	[Garzon et al. 2009]

	Spherical Harmonics-based rotational search. • The binding energy upon complex formation was approximated by a sum of three types of possibilities: van der Waals, electrostatics and desolvation, each of which can be composed as a correlation function. • A parallel version of FRODOCK can perform the docking search in a short time period, and the competitive docking accuracy achieved on standard protein–protein benchmarks demonstrates its applicability and robustness.		
MEDock	• The MEDock web server incorporates a worldwide search strategy that exploits the maximum entropy property of the Gaussian probability distribution in the context of information theory. • As a result of the global search strategy, the optimization algorithm incorporated in MEDock was found to be significantly superior when dealing with very harsh energy landscapes, which usually have insurmountable obstructions.	89	[Chang et al. 2005]
DOCKovalent	• DOCKovalent is used for screening large virtual libraries of electrophilic small molecules. • It can discover reversible covalent fragments that target distinct protein nucleophiles, including the catalytic serine of AmpC β-lactamase and non-catalytic cysteines in RSK2, MSK1 and JAK3 kinases.	81	[London et al. 2014]
TreeDock	• TreeDock is a docking tool that is able to	75	[Fahmy et

	explore all clash-free orientations at very fine resolution in a reasonable amount of time. Due to the speed of the program, many contact pairs can be mobility of the docking surfaces and structural rearrangements upon interaction. A novel algorithm, TreeDock, which addresses the enumeration problem in a rigid-body docking search. ● By representing molecules as multi-dimensional binary search trees and by investing an adequate number of docking orientations such that two chosen atoms, one from each molecule, are always in contact, TreeDock was able to explore all clash-free orientations at very fine resolution in a short amount of time.		al. 2002]
SnugDock	● SnugDock had been used to predict high-resolution antibody-antigen complex structures by simultaneously structurally optimizing the antibody-antigen rigid-body positions, the relative orientation of the antibody light and heavy chains, and the conformations of the six complementarity determining region loops. ● The approach is especially useful when the crystal structure of the antibody is not available. ● Local docking using SnugDock has shown to produce more accurate predictions than standard rigid-body docking.	75	[Sircar et al. 2010]
pyDockWEB	● pyDockWEB is a web server for the rigid-body docking forecast of protein–protein complex structures using an updated version of the pyDock scoring algorithm.	82	[Jiménez-García et al. 2013]

	• If the 3D coordinates of two interacting proteins were given, pyDockWEB returned the best docking orientations as scored mainly by electrostatics and desolvation energy.		
SANDOCK	• SANDOCK is a docking tool that is primarily developed for the automated docking of small ligands to a target protein. • It utilizes a guided matching algorithm to fit ligand atoms into the protein binding pocket. The protein was described by a modified Lee-Richard's dotted surface with each dot coded by chemical property and availability. • Orientations of the ligand in the active site are generated such that a chemical and a shape complementary between the ligand and the active site cavity must be satisfied.	71	[Burkhard et al. 1998]
DOCKGROUND	• DOCKGROUND is a comprehensive database of co-crystallized (bound) protein–protein complexes in a relational database of annotated structures. • This database contained comprehensive sets of complexes suitable for large scale benchmarking of docking algorithms. • The authors reported the important features to the set of bound structures, such as regularly updated downloadable datasets: automatically generated non-redundant set, built according to most common criteria, and a manually curated set that includes only biological non-obligate complexes along with a number of additional useful characteristics. Complexes from the bound	70 251	[Gao et al. 2007] [Hwang et al. 2008]

	dataset were utilized to identify the crystallized unbound analogs. If such analogs were nowhere to be found, the unbound structures were simulated by rotamer library optimization. **Benchmark-** ● The version was reported by the authors was 3.0. This included 40 experimental cases, representing a 48% increase from Benchmark 2.0. For majority of the new cases, the crystal structures of both binding partners were available. As with Benchmark 2.0, Structural Classification of Proteins (Murzin et al., J Mol Biol 1995; 247:536–540) was utilized to expel redundant test cases. The 124 unbound-unbound test cases in Benchmark 3.0 were characterized into 88 rigid-body cases, 19 medium-difficulty cases, and 17 difficult cases, based on the degree of conformational change at the interface upon complex formation. In addition to providing the community with more test cases for evaluating docking methods, the expansion of Benchmark 3.0 would facilitate the advancement of updated algorithms that would require a large number of training examples.		
DockDE	● DockDE program was compared to the Lamarckian GA (LGA) provided with AutoDock, and the DockEA previously found to outperform the LGA. ● The comparison was performed on a suite of six commonly used docking problems. DockDE outperformed the other algorithms on all problems. ● Further, the DockDE demonstrated	64	[Thomsen et al. 2003]

	remarkable performance in terms of convergence speed and robustness regarding the found solution. **Benchmark-** ● DockDE outperformed the DockEA and the LGA algorithm on all test cases regarding the test and mean energy values obtained. The convergence graphs and the low standard deviations also indicated that the DockDE was fast and robust (in terms of reproducing the docking results). Moreover, the two termination criteria introduced makes the search performance of the DockDE even better by lowering the number of evaluations needed by a factor of 2-40 without losing much accuracy. These findings are important and show great promise for applying the DockDE to virtual screening applications, i.e., searching huge ligand databases for promising drug candidates.		
CovalentDock	● CovalentDock is a computational algorithm built on the top of the source code of Autodock to model the phenomenon of chemical bonding and extended it to the server, known as the CovalentDock Cloud to make it accessible directly online without any local installation and configuration. ● It is an empirical model of free energy change estimation for covalent linkage formation, which is compatible with existing scoring functions used in docking, while handling the molecular geometry constraints of the covalent linkage with special atom types and directional grid maps.	68	[Ouyang et al. 2012]

	• Integrated preparation scripts were also mentioned for the automation of the whole covalent docking workflow. • This tool accepts the structures of both the ligand and the receptor uploaded by the user or retrieved from online databases with valid access id. It identifies the potential covalent binding patterns, carries out the covalent docking experiments and provides visualization of the outcomes for user analysis. **Benchmark** : • The prediction and quantification of covalent linkage formation in molecular docking is of great interest and potential to the systematic discovery of covalent drugs. The CovalentDock Cloud gives a user-friendly portal to carry out covalent docking experiments and to examine the outcomes online via web browser. With the powerful backend docking engine, it was believed that CovalentDock web server would offer a more accessible way for simulation and prediction of more accurate covalent docking. • The website was fine tuned for better user-experience and to continuously develop and enhance the Covalent Dock package. It gave more responses adopting covalent binding mechanism available and enabled users to specify the covalent linking pattern by themselves.		
GAsDock	• GAsDock is a fast flexible docking program which is based on an improved multi-population genetic algorithm. • It is an accurate and remarkably faster docking program in comparison with other	58	[Li et al. 2004]

	docking programs, which is advantageous in the application of virtual screening. **Benchmark-** ● In comparison with the optimization algorithms of other docking methods, information entropy was employed in the genetic algorithm of GAsDock and contracted space was used as the convergence criterion, which effectively controls the convergence of the algorithm, ensuring that GAsDock could converge rapidly and steadily. That is why GAsDock could bring better results in accuracy and higher speed than other programs.		
BDOCK	● BDOCK is an FFT-based docking algorithm system which includes specific scoring functions for different types of complexes. ● BDOCK uses family-based residue interface propensities as a scoring function and obtains improvement factors of 4-30 for enzyme-inhibitor and 4-11 for antibody-antigen complexes in two specific SCOP families. **Benchmark-** ● The Meta method improves the prediction success rates of individual prediction approaches. The tightness of fit scoring function based on these correctly predicted interface residues effectively discriminates between near-native complex structures and non-native ones. This approach was implemented in BDOCK and was applicable to all types of complexes. ● Adding further background for special classes of complexes, such as enzyme–inhibitor complexes, these results could be	56	[Huang et al. 2008]

	improvised.		
NPDock	• NPDock (Nucleic acid–Protein Docking) is a web server for predicting complexes of protein–nucleic acid structures which implements a computational workflow that includes docking, scoring of poses, clustering of the best-scored models and sorting of the most promising solutions. • The NPDock server provides a user-friendly interface and 3D visualization of the outcomes. The smallest set of input data consists of a protein structure and a nucleic acid structure (DNA or RNA) in PDB format. **Benchmark-** • NPDock is a web server developed for protein–nucleic acid docking that utilizes specific protein–nucleic acid statistical possibilities for scoring and selection of modeled complexes. NPDock implements a unique workflow based on a combination of computational strategies that have been published and offers a user-compatible web interface to enter PDB structures and view their results. • The automation of the entire procedure makes the protein–nucleic acid docking accessible to users who would otherwise become tripped up installing many complex programs locally and then carrying out numerous manual advances; each requiring an assortment of manual format conversions that are highly prone to human error. Therefore, it can help users save even more than ten times the time required to run diverse strategies separately and sequentially.	73	[Tuszynska et al. 2015]

ScoreDock	• An empirical protein-ligand binding affinity estimation technique, SCORE, was incorporated into a popular docking program, DOCK4. The consolidated program was named as ScoreDock. • It had been used to reconstruct the 200 protein-ligand complex structures and found to give good results for the complexes with high binding affinities. **Benchmark:** • Using existing drugs for new indications (drug repurposing) is a compelling technique not only to reduce drug development time and costs but also to develop treatments for new disease including those that were rare. In order to discover novel indications, potential target identification is an essential step. One broadly utilized method to identify potential targets was through molecule docking. • It requires no prior data except structure inputs from both the drug and the target, and can identify potential targets for a given drug, or recognize potential drugs for a specific target. Despite the fact that molecular docking is popular for drug development and repurposing, challenges remain for the method. In order to improve the prediction accuracy, optimizing the target conformation, considering the solvents and adding co-binders to the system are conceivable arrangements.	56	[Luo et al. 2016]
SDOCKER	• The primary objective of SDOCKER is docking accuracy improvement. In this paradigm, simulated annealing molecular	55	[Wu et al. 2004]

	dynamics was utilized for conformational sampling and optimization and an additional similarity force is applied on the basis of the positions of ligands from X-ray information that focus the sampling on relevant regions of the active site. **Benchmark:** ● Genomic pipelines comprise of several pieces of third party software and, because of their experimental nature, frequent changes and updates were commonly necessary thus raising serious deployment and reproducibility issues. Docker containers are emerging as a possible solution for a large number of these issues, as they allow the packaging of pipelines in an isolated and self-contained manner. ● This makes it simple to distribute and execute pipelines in a portable manner across a wide range of computing platforms. Thus, the question that arises is to what degree the utilization of Docker containers might affect the performance of these pipelines.		
pyDockRST	● pyDockRST software uses the percentage of satisfied distance restraints, together with the electrostatics and desolvation binding energy, to identify correct docking orientations. ● This technique drastically improved the docking results when compared to the use of energy criteria alone, and was able to find the correct orientation within the top 20 docking solutions in 80% of the cases. **Benchmark-** ● pyDockWEB is a web server for the rigid-body docking prediction of protein–protein	48	[Chelliah et al. 2006]

	complex structures utilizing another version of the pyDock scoring algorithm. A custom parallel FTDock implementation was used, with adjusted grid size for optimal FFT calculations, and an updated version of pyDock, which dramatically speeds up calculations while keeping the same predictive accuracy. ● Given the 3D coordinates of two interacting proteins, pyDockWEB returns the best docking orientations as scored fundamentally by electrostatics and desolvation energy.		
GalaxyPepDock	● GalaxyPepDock web server, which is freely accessible at http://galaxy.seoklab.org/pepdock, performs similarity-based docking by finding templates from the database of experimentally determined structures and building models using energy-based optimization that allows for structural flexibility. ● The server can therefore effectively create the structural differences between the template and target protein–peptide complexes. ● The performance of GalaxyPepDock is better than those of the other available web servers when tested on the PeptiDB set and on several complex structures. ● When tested on the CAPRI target 67, GalaxyPepDock generates models that are more precise than the best server models submitted during the CAPRI blind prediction experiment. ● GalaxyPepDock is a similarity-based	58	[Lee et al. 2015]

	protein–peptide docking web-server that performs additional flexible-structure energy-based optimization. The effective combination of database search and physics-based optimization allows for a superior performance compared with the existing methods when complexes involving similar proteins could be found in the database. • **GalaxyWEB provides the following web services:** • **Protein Structure Prediction** • GalaxyTBM: Protein structure prediction from sequence by template-based modeling • GalaxyLoop: Modeling of loop and/or terminus regions specified by user • GalaxyDom: Protein modeling unit detection for protein structure predictions • **Protein Structure Refinement** • GalaxyRefine: Refinement of model structure provided by user • GalaxyRefineComplex: Refinement of protein-protein complex model structure provided by user • **Protein Interaction Prediction** • GalaxySite: Ligand binding site prediction from a given protein structure (experimental or model) • GalaxyPepDock: Protein-peptide docking based on interaction similarity • GalaxyHomomer: Protein homo-oligomer structure prediction from a monomer		

	sequence or structure • GalaxyGemini: Protein homomer structure prediction from a given protein monomer structure based on similarity • GalaxyTongDock: Symmetric and asymmetric protein-protein docking. • **GPCR Applications** • Galaxy7TM: Flexible GPCR-ligand docking by structure refinement with a GPCR and a ligand structure provided by user. • GalaxyGPCRloop: Structure prediction of the second extracellular loop of GPCR.		
CombDock	• CombDock is a combinatorial docking algorithm for the structural units' assembly problem which also gives a heuristic solution to a computationally hard problem (NPC). • CombDock is also used for the automated assembly of protein substructures. This application can anticipate near-native assemblies for various examples of both domains and to build blocks with different levels of distortion. It can also be utilized in protein structure prediction if the local structural units are given and assisted in obtaining a structural model. • It is used for protein-ligand binding. • CombDock can be operated only on Linux Operating System.	46	[Inbar et al. 2005]
FastDock	• FastDock engine which uses a Lamarckian genetic algorithm (LGA) so that	46	[Yadav et al. 2010]

	individuals adapt to the surrounding environment. The best fits are continued through analyzing the PMF scores of each chromosome and assigning more reproductive opportunities to the chromosomes having lower scores. This procedure rehashes for almost 3,000 generations with 500 individuals and 100,000 energy evaluations. Other parameters were left to their default values. • Structure-based screening includes docking of candidate ligands into protein targets, followed by applying a PMF scoring function to assess the probability that the ligand will bind to the protein with high affinity or not. • Now called as SWISSDOCK TOOL. • Targets can be uploaded by the user and also through PDB id. It is used for protein ligand interaction.		
GlamDock	• GlamDock tool is based on a Monte-Carlo with minimization search in a hybrid interaction matching or an internal coordinate search space. • The main features of the method are (1) the energy function, which is a continuously differentiable empirical potential and (2) the definition of the search space, which combines internal coordinates for the adaptation of the ligand, with a mapping-based portrayal of the rigid body translation and rotation. **Benchmark:** A set of 100 protein-ligand complexes, which enables comparative evaluation to	47	[Tietze et al. 2011]

	existing docking tools. The outcomes on the given benchmark demonstrated that GlamDock is at least comparable in efficiency and accuracy to the best existing docking tools. The primary focal point of this work was the validation on the scPDB database of protein-ligand complexes. The size of this dataset allowed a thorough analysis of the dependencies of docking accuracy on features of the protein-ligand system. Specifically, it allowed a two-dimensional analysis of the outcomes, which identifies a number of interesting dependencies that are generally lost or even misinterpreted in the one-dimensional approach. The overall outcome was that GlamDock correctly predicted the complex structure in practically half of the cases in the scPDB is critical not just to screen ligands against a specific protein but even more so for opposite screening, i.e., the identification of the correct targets for a particular ligand.		
PostDOCK	• PostDOCK distinguishes true binding ligand−protein complexes from docking artifacts (that were created by DOCK 4.0.1). • It is a pattern recognition system that relies on (1) a database of complexes, (2) biochemical descriptors of those complexes, and (3) machine learning tools. Protein databank (PDB) was used as the structural database of complexes and create diverse training and validation sets from it based on the "families of structurally similar proteins" (FSSP) hierarchy. • Allows analyzing and comparing molecular docking results. PostDock helps	45	[Springer et al. 2005]

	in visualization of docking results. It displays an interactive pseudo-3D snapshot of multiple docked ligand poses such that both the docking poses and docking scores are encoded visually for rapid assessment. The software possesses a full complement of display options to tailor the visual examination of any molecular design task. It is proficient for the rapid visual examination of molecular docking results.		
ParaDockS	• ParaDockS is software designed to hold different optimization algorithms and objective functions. • The functions of ParaDockS that were available are as follows (i) the empirical objective function p-Score and (ii) an adapted version of the knowledge-based potential PMF04. • Accurate prediction of protein–DNA complexes could give a critical stepping stone towards an exhaustive appreciation of vital intracellular processes. • ParaDock is an *ab initio* protein–DNA docking algorithm that combines short DNA fragments, which have been rigidly docked to the protein based on geometric complementarity, to create bent planar DNA molecules of arbitrary sequence.	44	[Meier et al. 2010]
AUDocker LE	• AUDocker LE was structured with a plan to develop a software tool as a front end graphical interface with C- language to perform docking experiments in Windows based PCs. It encourages users to perform automated continuous docking of expansive ligand databases into a set of predefined protein targets.	44	[Sandeep et al. 2011]

	• It would likewise assist the user with analyzing the results to select promising lead molecules.		
GAPDOCK	• GAPDOCK is a genetic algorithm based docking tool used to predict the structure of two protein-protein complexes in combination with surface complementarity, buried surface area, biochemical information, and human intervention. **Benchmark:** • Among the five models submitted for target 1, HP phospho carrier protein (*B. subtilis*) and the hexameric HPr kinase (*L. lactis*), the best accurately predicts 17 of 52 interprotein contacts, though for target 2, bovine rotavirus VP6 protein-monoclonal antibody, the best model predicts 27 of 52 correct contacts.	37	[Gardiner et al. 2003]
FIPSDock	• FIPSDock is a docking tool which implements a variant of the Fully Informed Particle Swarm (FIPS) optimization method and adopts the highly developed energy function of AutoDock 4.2 suite for solving flexible protein–ligand docking problems. • The search capacity and docking precision of FIPSDock were first assessed by multiple docking tests. In a benchmarking test for 77 protein/ligand complex structures generated from GOLD benchmark set, FIPSDock has acquired a successful predicting rate of 93.5% and outperformed a few docking programs. • FIPSDock is based on a variant of Particle	43	[Liu et al. 2013]

	Swarm Optimization (PSO) known as Fully Informed Particle Swarm (FIPS) and the semi-empirical free energy force field in AutoDock 4.0, an updated approach to flexible docking. FIPSDock is much better than AutoDock and SODOCK which was also proposed by improving AutoDock with PSO in term of obtaining a lower binding energy, a better docked conformation, convergence speed and robustness. Compared with the four currently widely used methods, i.e., GOLD, DOCK, FlexX and AutoDock, FIPSDock is more accurate. Thus, FIPSDock is an efficient and accurate docking method and its promising prospects can be relied upon in the application to virtual screening.		
GriDock	• GriDock is a parallelized tool based on the AutoDock4.0 engine which can perform efficient and easy virtual screening analyses of large molecular databases exploiting multi-core architectures.	36	[Vistoli et al. 2010]
RPDOCK	• RPDOCK is a novel docking procedure specific to RNA-protein complexes. • RPDOCK incorporates the features specific to RNA-protein interfaces (including looser atom packing at interface, preference of positively charged amino acid residues at RNA-protein interfaces and stacking interactions between the bases of nucleotides and aromatic rings of charged amino acids). • RPDOCK is an FFT-based algorithm that takes into account of RNA–protein interactions into consideration, and RPRANK is a knowledge-based potential using root-mean-square deviation as a	56	[Huang et al. 2008]

	measure.		
pDOCK	• pDOCK is a computational technique for rapid and accurate docking of flexible peptides to MHC receptors and primarily apply it on a non-redundant dataset of 186 pMHC (MHC-I and MHC-II) complexes with X-ray crystal structures.	34	[Khan et al. 2010]
MedusaDock	• In the molecular docking suite MedusaDock, both ligand and receptor side chain flexibilities were modeled simultaneously with sets of discrete rotamers, where the ligand rotamer library was generated "on the fly" in a stochastic manner. • Backbone flexibility was introduced into MedusaDock by implementing ensemble docking in a sequential manner for a set of distinct receptor backbone conformations.	39	[Ding et al. 2012]
LigDockCSA	• LigDockCSA was developed by using a powerful global optimization technique, conformational space annealing (CSA), and a scoring function that combines the AutoDock energy and the piecewise linear potential (PLP) torsion energy. • It was found that the CSA search method can discover the lower energy binding poses than the Lamarckian genetic algorithm of AutoDock. • LigDockCSA finds the best scoring poses within 2 Å root-mean-square deviations (RMSD) from the native structures for 84.7% of the test cases, compared to 81.7% for AutoDock and 80.5% for GOLD. • Scoring function of LigDockCSA is a	32	[Shin et al. 2011]

	modification of AutoDock3 scoring function (Morris et al., 1999) with adding torsion part of piecewise linear potential (PLP) (Gehlharr et al., 1995). • When CSA searches ligand binding mode with AutoDock3 scoring function, it can find lower energy conformation when compared to Lamarkian Genetic Algorithm (LGA).		
pyDockTET	• pyDockTET is a tethered-docking program which uses rigid-body docking system to generate domain-domain poses that are further scored by binding energy and a pseudo-energy term based on restraints derived from linker end-to-end distances. • The method had been benchmarked on a set of 77 non-repetitive pairs of domains with accessible X-ray structure. • *pyDockTET*, an advanced scoring function incorporated in *pyDock*, to model specifically the conformation of domain-domain assemblies.	29	[Cheng et al. 2008]
SDOCK	• SDOCK approach performs global docking based on force-field potentials; one of its advantages is that it provides global binding free energy surface profiles for further analysis. The efficiency of the program is also comparable with that of other FFT based protein-protein docking programs. • It suggests the robustness of FFT-based docking sampling algorithm along with the importance of electrostatics.	31	[Zhang et al. 2011]
ASPDock	• ASPDock is an FFT-based algorithm which is used to calculate Atomic	35	[Li et al. 2011]

	Solvation Parameters (ASP) scores of protein complexes. • As compared to other state-of-the-art docking algorithms, it was found that the ASP score produced a higher success rate than the pure shape complementarity score of FTDock but lower success rate than ZDOCK 3.0. • The ASP-based docking method performs well in CAPRI rounds 18 and 19. • The softly restricting method (SRM) is based on the ASPDock algorithm, which uses atomic solvation parameters (ASP) rather than geometric complementary.		
MEGADOCK	• MEGADOCK is a Protein-protein docking software package which samples an extremely large number of protein dockings at high speed. • MEGADOCK decreased the calculation time required for docking by using multiple techniques, one of which was a scoring function called the real Pairwise Shape Complementarity (rPSC) score. • It is capable of exhaustive PPI screening by completing docking calculations 7.5 times faster than the conventional docking software, ZDOCK, while maintaining an acceptable level of accuracy. • MEGADOCK can be applied to a large-scale protein-protein interaction-screening issue with accuracy superior to arbitrary.	34	[Ohue et al. 2014]
BetaDock	• BetaDock is a docking tool based on the theory of the β-complex. If the Voronoi diagram of the receptor, whose topology is	28	[Kim et al. 2011]

	stored in the quasi-triangulation, is given, the β-complex corresponding to water molecule is calculated. At that point, the boundary of the β-complex characterizes the β-shape which has the complete proximity data among all atoms on the receptor boundary. ● The performance of the algorithm was tested through a benchmark test and it was found that BetaDock is better than the popular docking software AutoDock 4.		
DOCKTITE	● DOCKTITE is a highly versatile workflow for covalent docking in the Molecular Operating Environment (MOE) combining automated warhead screening, nucleophilic side chain attachment, pharmacophore-based docking, and a novel consensus scoring approach which combines the knowledge-based scoring function drug score extended (DSX) and the empirical scoring functions implemented in MOE. ● DOCKTITE software can differentiate binders from non binders and rank active compounds regarding their experimentally determined binding affinity values in a congeneric series of ligands.	41	[Scholz et al. 2015]
MTiOpenScreen	● MTiOpenScreen is dedicated to docking of small molecules and also for virtual screening purposes. There are two services which are available, namely- MTiAutoDock and MTiOpenScreen. This tool uses AutoDock 4.2 and AutoDock Vina for processing. ● There is a valuable resource known as MTiOpenScreen which provide drug-like chemical libraries containing 150000 PubChem compounds: the Diverse-lib	36	[Labbé et al. 2015]

	containing diverse molecules and the iPPI-lib enriched in molecules likely to inhibit protein–protein interactions.		
TCRFlexDock	• TCR FlexDock improved predictive success over the fixed backbone protocol, leading to near-native predictions for 80% of the TCR/pMHC cases among the top 10 models, and 100% of the cases in the top 30 models • Flexible docking simulations can give precise models and atomic-level insights into TCR acknowledgement of MHC-like molecules displaying lipid and other small molecule antigens.	27	[Pierce et al. 2013]
mtsslDock	• mtsslDock is a docking tool which is used for translation of experimental distance distributions into structural information. • It is based on the mtsslWizard program for *in silico* spin labeling. • It has improved docking performances and also includes additional types of spin labels and contains applications for the trilateration of paramagnetic centres in biomolecules for rigid-body docking of sub-domains of macromolecular complexes.	26	[Hagelueken et al. 2013]
AnchorDock	• AnchorDock is a peptide docking approach which naturally focuses on the docking pursuit to the most applicable parts of the conformational space. • This is performed by pre-computing the free peptides structure and by automatically identifying anchoring spots on the protein surface. After that, a free peptide conformation undergoes anchor-driven simulated annealing molecular	27	[Ben-Shimon et al. 2015]

	dynamics simulations around the predicted anchoring spots. • AnchorDock produced exceptionally good results (backbone root-mean-square deviation ≤ 2.2Å, rank ≤15) in the challenging task of a completely blind docking test, for 10 of 13 unbound cases tested. pepATTRACT is a docking protocol that is fully blind, i.e. it does not require any information about the binding site. • Its performance was either similar to or better than state-of-the-art local docking protocols that do require binding site data. • Since it is fully blind, the short running time makes the pepATTRACT web server suitable for large-scale in silico protein–peptide docking experiments, and the performances in the identification of the receptor interacting residues can provide a useful starting point to justify the design of further experiments in the wet lab.		
AutoDockFR	• AutoDock for Flexible Receptors (AutoDockFR) is based on AutoDock4 scoring function. • It addresses challenges such as exponential growth of the search space and false positive results. • AutoDockFR reports more correctly cross-docked ligands than AutoDock Vina on both datasets with solutions found for 70.6% vs. 35.3% systems on SEQ17, and 76.9% vs. 61.5% on CDK2. • *AutoDockFR*, simulates partial receptor flexibility by allowing a large number of	36	[Ravindranath et al. 2015]

	explicitly specified receptor side-chains to explore their conformational space, while searching for energetically favorable binding poses for a given ligand. • Previous approaches have so far been limited to a small number of flexible protein side-chains (2–5), thus requiring prior knowledge of receptor side-chains undergoing conformational change upon binding of a given ligand. The demonstrated capability of AutoDockFR in identifying right answers for issues with up to 14 flexible receptor side-chains reduces this requirement.		
*e*SimDock	• *e*SimDock is an approach to ligand docking and binding affinity prediction. • It employs nonlinear machine learning-based scoring functions to improve the accuracy of ligand ranking and similarity-based binding pose prediction, and to increase the tolerance to structural imperfections in the target structures. • The performance of eSimDock is greatly unaffected by the deformations of ligand binding sites, thus it represents a practical technique for across-proteome virtual screening using protein models. eSimDock uses non-linear statistical model.	25	[Brylinski et al. 2013]
PharmDock	• PharmDock is a pharmacophore-based docking program that combines pose sampling and ranking based on optimized protein-based pharmacophore models with local optimization using an experimental scoring function. • Protein-based pharmacophore models were improved with the data of potential	27	[Hu et al. 2014]

	interactions between ligands and the protein target. ● A protein pharmacophore-based docking program, PharmDock, was made accessible with a PyMOL module. PharmDock and the PyMOL module are freely available at http://people.pharmacy.purdue.edu/~mlill/programming/pharmdock.		
ArgusLab	● ArgusLab is free docking software used for virtual screening and calculating the weight of van der Waals interactions unimportant for binding free energy calculations. ● The main advantage of this software is in terms of accuracy and short computational time as compared to other systems.	19	[Oda et al. 2009]
DockingApp	● DockingApp is a freely accessible, easy to use, platform-independent application for performing docking simulations and virtual screening tasks using AutoDock Vina. ● DockingApp sports a natural graphical user interface which greatly encourages both the input phase and result analysis, which can be visualized in graphical form using the embedded JMol applet. ● DockingApp is a user-friendly software application meant to allow a variety of differently-skilled users to perform docking simulations, with high confidence on the results produced and minimal effort for setup and configuration. ● AutoDock Vina, which is the "engine" used by DockingApp to carry out the actual docking simulation.	17	[Di Muzio et al. 2017]
KinDock	● KinDOCK is a web server for the analysis	18	[Martin et

	of ATP-binding sites of protein kinases. This characterization depends on the docking of ligands already co-crystallized with other protein kinases. ● A structural library of protein kinase–ligand complexes was extracted from the Protein Data Bank (PDB). This library can give both potential ligands and their putative binding orientation for a given protein kinase. ● The server and its documentation are freely accessible at http://abcis.cbs.cnrs.fr/kindock/. ● It combines structural comparisons, immediate transfer of known ligands from the template structure into the target structure, visualization of the deduced protein–ligand complexes and evaluation of protein–ligand interactions.		al. 2006]
OptiDock	● The OptiDock strategy portrayed in this involves choosing a different but representative subset of compounds that span the structural space incorporated by the full library. These compounds were docked individually using the FlexX program.	19	[Sprous et al. 2004]
ParaDock	● ParaDock is an *ab initio* protein–DNA docking algorithm. which joins short DNA fragments, which have been rigidly docked to the protein based on geometric complementarity, to make bent planar DNA molecules of discretionary sequence. ● The algorithm was tested on the bound and unbound targets of a protein– DNA benchmark consisting of 47 complexes. In terms of benchmarking, CAPRI acceptable solutions were obtained among the 10 top	21	[Banitt et al. 2011]

	ranked structures in 83% of the bound complexes and 70% of the unbound complexes. Without requiring prior information of DNA length and sequence and within less than 2h per target on a standard 2.0 GHz single processor CPU, ParaDock gives a fast *ab initio* docking solution. ● ParaDock algorithm is independent of DNA sequence and length.		
DockRank	● DockRank is an approach for scoring docked conformations based on the degree to which the interface residues of the docked conformation match a set of predicted interface residues. ● DockRank utilizes interface residues predicted by partner-specific sequence homology-based protein– protein interface indicator (PS-HomPPI), which predicts the interface deposits of an inquiry protein with a particular association accomplice. ● Variations of DockRank that use predicted interface residues obtained from a few protein interface predictors that don't consider the binding partner in making interface predictions. ● DockRank is accessible as a server at http://einstein.cs.iastate.edu/DockRank/.	21	[Xue et al. 2014]
ASPDock	● Atomic Solvation Parameters (ASP) model had turned out to be an exceptionally successful technique for calculating the binding free energy of protein complexes. This recommends incorporating it into docking algorithms so that the prediction accuracy gets improved. In this paper an FFT-based calculation was proposed to figure ASP scores of protein complexes	30	[Li et al. 2011]

	and build up an ASP-based protein-protein docking strategy (ASPDock). ● ASPDock is a docking algorithm based on FFT method. Traditional FFT docking methods consider the shape complementarity as a crucial criterion to rank the predicted complex structures whereas ASPDock implements atomic solvation parameters in traditional FFT method to rank the predicted complex structures. ● ASPDock performs better than the shape complementarity docking method on benchmark 3.0.		
DockBench	● DockBench 1.0 is a freely accessible platform. ● It automates the entire procedure, from docking benchmark to Virtual Screening (VS) setup. ● It offers the possibility to test up to seventeen distinct protocols. ● DockBench 1.0 handles seven docking programming bundles and offers the likelihood to test up to seventeen unique conventions. ● All functionalities were embedded in a graphical user interface (GUI) and are composed into five main tabs, corresponding to the tasks required to do a complete pipeline, from docking benchmark studies to VS tests: (1) Input Settings; (2) Docking Protocols Settings; (3) Results Visualization; (4) Plots Visualization; (5) Virtual Screening Settings.	20	[Cuzzolin et al. 2015]
InterEvDock	● InterEvDock is a server for protein	24	[Yu et al.

	docking based on a free rigid body docking procedure. An orderly rigid body docking search was performed utilizing the FRODOCK program and the resulting models were re-scored with InterEvScore and SOAP-PP statistical potentials. ● The InterEvScore potential was specifically designed to integrate co-evolutionary data in the docking procedure. ● InterEvDock web server is the free docking server allowing to directly predicting the structure of protein–protein interactions using co-evolutionary information.		2016]
MDockPeP	● MDockPeP docks the all-molecule, flexible peptide onto the entire protein. ● It requires only the peptide sequence and the protein structure. ● MDockPeP achieves significantly better performance than other existing docking methods and is suitable for large-scale applications	23	[Yan et al. 2016]
SOFTDOCK	● SOFTDOCK is one of the first molecular docking methods developed for protein–protein docking. ● It has the ability to represent the molecular surface with different shapes and properties and to dock a variety of molecular complexes with certain conformational changes. ● The SOFTDOCK package utilizes a coarse-grained docking strategy to sample all possible conformations of complexes. SOFTDOCK utilizes Voronoi molecular surface and figures several grid-based scores. It was shown by the leave-one-out test that three geometry scores and an	17	[Jiang et al. 2002]

	FTDOCK-like electrostatics score contribute the most to the discrimination of near-native conformations.		
HybridDock	• HybridDock is a general hybrid docking protocol that utilizes both the protein structures and known ligands by combining the molecular docking program MDock and the ligand-based similarity search method SHAFTS • Hybrid docking protocol significantly improves the performance in both binding affinity and binding mode predictions, compared to the sole MDock program. • It can act as an alternative docking approach for modern drug design/discovery. • Hybrid docking protocol significantly enhanced the performance in both binding affinity and binding mode predictions, compared with the sole MDock program.	21	[Huang et al 2015]
MpSDockZn	• MpSDockZn automatically extrapolate the binding poses, i.e., Best Dock (BD), Best Cluster (BC) and Best Fit (BF) poses as well as to perform consistent cluster and docking accuracy analyses.	16	[Ballante et al 2016]
EpiDOCK	• EpiDOCK is the structure-based server for MHC class II binding prediction. EpiDOCK predicts binding to the 23 most successive human MHC class II proteins. • It had identified 90% of true binders and 76% of non-true binders, with a total accuracy of 83%. • EpiDOCK is freely accessible at http://epidock.ddg-pharmfac.net. • EpiDOCK converts the input sequence into a collection of overlapping nonamers, because the peptide binding core consists of nine contiguous residues. Every	19	[Atanasova et al. 2013]

	nonamer is evaluated by a docking score-based quantitative matrix (DS-QM) derived for the selected HLA class II protein and assigned a score.		
pyDockCG	• pyDockCG is another coarse-grained potential for protein– protein docking scoring and refinement, in view of the known UNRES model for polypeptide chains. • The main feature was the inclusion of two terms accounting for the Coulomb electrostatics and the solvation energy. • It is suitable for the treatment of flexibility amid docking. • The coarse-grained potential yielded highly similar values to the full-atom scoring function pyDock when connected to the rigid body docking sets, however at much lower computational cost.	15	[Solernou et al. 2011]
bhDock	• The bhDock technique uses two-step algorithm. First, a comprehensive arrangement of low-resolution binding sites is determined by analyzing whole protein surface and ranked by a simple score function. Second, ligand position is determined by means of a molecular dynamics-based method of global optimization beginning from a small set of high ranked low-resolution binding sites. • Appraisal of the bhDock strategy on the set of 37 protein– ligand complexes has shown the success rate of forecasts of 78%, which is superior to the rate reported for the most cited docking techniques, for example, AutoDock, DOCK, GOLD, and FlexX, on similar sets of complexes. • The main developments in docking in this	17	[Vorobjev et al. 2010]

	period, covered in this review, are receptor flexibility, solvation, fragment docking, post-processing, docking into homology models, and docking comparisons.		
DockTrina	• DockTrina is a protein docking technique for demonstrating the 3D structures of non-symmetrical triangular trimers. • The strategy takes as input pair-wise contact predictions from a rigid body docking program. It then scans and scores all possible combinations of pairs of monomers utilizing a very fast root mean square deviation test. • It ranks the predictions by the use of scoring functions which combines triples of pair-wise contact terms and a geometric clash penalty term. • The method takes under 2 min for each complex on a modern desktop computer. • The method was tested and approved utilizing a benchmark set of 220 bound and seven unbound protein trimer structures.	13	[Popov et al. 2013]
MacDOCK	• MacDOCK is a similarity-driven docking program based on DOCK 4.0. • It is able to generate orientations consistent with the known covalent binding mode of the complexes, with a performance similar to that of other docking programs. • It can be used efficiently for the virtual screening of covalently bound ligands. • Various molecular docking techniques have been maximally exploiting all accessible structural and chemical information that can be obtained from proteins, from ligands, and from protein-	62	[Fradera et al. 2004]

	ligand complexes. In this regard, the term 'guided docking' was introduced to refer to docking approaches that incorporate some degree of chemical information to actively guide the orientation of the ligand into the binding site. ● Accelerating the drug discovery process requires predictive computational protocols fit for reducing or simplifying the synthetic as well as combinatorial challenge. Docking-based virtual screening strategies have been developed and successfully applied to various pharmaceutical targets.		
KBDOCK	● It is a database system that combines the Pfam domain characterization with coordinate data from the PDB to analyse and display 3D domain– domain interactions (DDIs). ● For a given query domain or pair of domains, KBDOCK retrieves and shows a non-redundant list of homologous DDIs or domain– peptide interactions in a common coordinate frame. ● It may also be utilized to search for and visualize interactions involving different, but structurally similar Pfam families. ● The current KBDOCK database was built from the June 2013 snapshot of the PDB and the latest version of Pfam (release 27.0). ● It gathers and classifies hetero and homo DDIs, just as all domain– peptide connections (DPIs).	15	[Ghoorah et al. 2013]
ReplicaDock	● Replica exchange Metropolis-Monte Carlo method for the low-resolution stage of protein-protein docking, which was	14	[Zhang et al. 2013]

	implemented within the RosettaDock program. ● ReplicaDock, uses temperature replica exchange to switch between bound and unbound thermodynamic states, and benchmarked its performance for sampling the low-resolution stage of protein-protein docking in RosettaDock.		
WinDock	● To make HTD more accessible to a broad community, WinDock, an integrated application was designed to help researchers perform structure-based drug discovery tasks under a uniform, user friendly graphical interface for Windows-based PCs. ● WinDock combines existing small molecules accessible three-dimensional (3D) libraries, homology modeling tools, and ligand-protein docking programs in a semi-automatic, intelligent way, which guides the user through the use of each integrated software component.	11	[Hu et al. 2007]
DockScore	● DockScore is a target scoring scheme that can be utilized to rank protein-protein docked poses. It considers several interface parameters, namely, surface area, evolutionary conservation, hydrophobicity, short contacts and spatial clustering at the interface for scoring. ● DockScore web server can be employed, subsequent to docking, to perform scoring of the docked solutions, starting from multiple poses as inputs. ● The web server for DockScore can be freely accessed at: http://caps.ncbs.res.in/dockscore/.	12	[Malhotra et al. 2015]

HDOCK	• HDOCK is a web server of the hybrid docking algorithms of template-based modeling and free docking, in which cases with deceiving templates can be protected by the free docking protocol. The server supports protein– protein and protein–DNA/RNA docking and acknowledges both sequence and structure inputs for proteins. The docking process is quick and expends about 10– 20 min for a docking run. Tested on the cases with weak homologous complexes of less than 30% sequence identity from five docking benchmarks. • The HDOCK pipeline tied with template-based modeling on the protein– protein and protein– DNA benchmarks and performed superior than template-based modeling with respect to the three protein– RNA benchmarks when the best 10 predictions were considered. • The performance of HDOCK turned out to be better when more predictions were considered. Combining the outcomes of HDOCK and template-based modeling by ranking first of the template based model additionally enhanced the predictive power of the server.	26	[Yan et al. 2017]
HiPCDock	• A High-Performance Computing (HPC)-based molecular docking scheme, termed HiPCDock was used for drug discovery and development. • It had been implemented to be used by both computational experts and experimental scientists. • Thus it is an automated, user-friendly and	11	[Zhang et al. 2009]

	efficient package for molecular docking based high throughput virtual screening in drug discovery.		
MoDock	• MoDock adopts an aggregate function to approximate the real solution of the original multi-objective and multi-constraint problem, which smooth the energy surface of the combined scoring functions. • At that point, method of centers and genetic algorithms are used to find the optimal solution. • Trial of MoDock against the GOLD test dataset reveals the multi-objective procedure improves the docking accuracy over the individual scoring functions.	11	[Gu et al. 2015]
LiGendock	• LiGenDock is based on pharmacophore models of binding sites, including a non-enumerative docking calculation. • It shows accompanying module LiGenPocket, aimed at the binding site analysis and at the structure-based pharmacophore definition. • The optimization procedure was reported that was carried out to improve the cognate docking and virtual screening performance of LiGenDock.	12	[Beato et al. 2013]
mPockDock	• mPockDock is a multi-conformational docking approach which reduces the rate of false-negatives in activity prediction. • mPockDock provide the AUC of 83.8%. • It has proved to be efficient for scaffold hopping.	10	[Chen et al. 2014]
CRDOCK	• CRDOCK is an ultrafast docking and virtual screening program that contains (1)	14	[Cabrera et

	a search engine that can use a variety of sampling methods and an initial energy evaluation function, (2) several energy minimization calculations for calibrating the binding poses, and (3) distinctive scoring functions. ● Testing CRDOCK on two broadly utilized benchmarks, the ASTEX diverse set and the Directory of Useful Decoys, yielded a success rate of ~75% in pose prediction and an average AUC of 0.66.		al. 2012]
DockQ	● DockQ is a continuous protein-protein docking model quality measure derived by combining F_{nat}, LRMS, and iRMS to a single score in the range that can be used to assess the quality of protein docking models. ● Utilizing DockQ on CAPRI models it is possible to almost give the original CAPRI classification into Incorrect, Acceptable, Medium and High quality.	10	[Basu et al. 2016]
ELMDOCK	● ELMDOCK is a tool which evaluates a rigid-body. It is a deterministic molecular docking method which relies solely on the three-dimensional structure of the individual components and the overall rotational diffusion tensor of the complex, obtained from nuclear spin-relaxation measurements. ● A docking technique, called ELMPATIDOCK, is based on the idea of combining the shape-related limitations from rotational diffusion with those from residual dipolar couplings, along with ambiguous contact/interface-related restrictions obtained from chemical shift	9	[Berlin et al. 2011]

	perturbations.		
FlexGAsDock	• In this method, the optimization of molecular docking was divided into two sub-problems based on the different effects on the protein–ligand interaction energy. • An adaptive genetic algorithm was created to solve the optimization issue and an updated docking program (FlexGAsDock) based on the hierarchical docking strategy was developed. • The docking results demonstrated that this strategy could be helpfully utilized for the efficient molecular drug designing.	8	[Kang et al. 2012]
MEGADOCK-GPU	• MEGADOCK is fast protein-protein docking programming yet more speed is needed for an interactome prediction, which is composed of millions of protein pairs. • Ultra-fast protein-protein docking software named MEGADOCK-GPU was developed by using general purpose GPU computing techniques. • A system was implemented that utilizes all CPU cores and GPUs in a computation node. • MEGADOCK-GPU on 12 CPU centers and 3 GPUs accomplished a figuring speed that was 37.0 occasions quicker than MEGADOCK on 1 CPU center. • The novel docking programming facilitates the utilization of docking techniques to help large-scale protein interaction network analyses.	8	[Shimoda et al. 2013]

	• MEGADOCK-GPU is openly accessible at http://www.bi.cs.titech.ac.jp/megadock/gpu/.		
DockAFM	• The DockAFM tool sets up a connection between topographic images from AFM and the molecular dynamics of single proteins. • DockAFM computes the fit of input conformations of a given molecule with the topographic surface of AFM images. Thus, DockAFM can be utilized to benchmark protein 3D structures or models against an experimental data obtained by atomic force microscopy.	4	[Chaves et al. 2013]
HSYMDOCK	• HSYMDOCK is a web server of progressive symmetric docking algorithm that supports both Cn and Dn symmetry. • The HSYMDOCK server was broadly assessed on three benchmarks of symmetric protein complexes, including the 20 CASP11–CAPRI30 homo-oligomer targets, the symmetric docking benchmark of 213 Cn targets and 35 Dn targets, and a non-repetitive test set of 55 transmembrane proteins. • It was demonstrated that HSYMDOCK obtained a significantly better performance than other similar docking algorithms. • The server supports both sequence and structure inputs for the monomer/subunit.	1	[Yan et al. 2018]
MemDock	• MemDock is software for docking α-helical membrane proteins which takes into consideration the lipid bilayer environment for docking just as for	8	[Hurwitz et al. 2016]

	refining and positioning the docking candidates.		
UDock	• In UDock, the users can tackle simplified representations of protein structures and explore protein-protein interfaces' conformational space using a gamified interactive docking system with on the fly scoring. • It makes use of users' cognitive capabilities to provide relevant data for (1) the prediction of correct interfaces in binary protein complexes and (2) the identification of the experimental partner in interaction among a set of decoys.	8	[Levieux et al. 2014]
MpSDock	• MpSDock is software that runs on a scheme similar to consensus scoring that consists of a force-field-based scoring function and a knowledge-based scoring function. • This optimization technique can dynamically sample and regenerate decoy poses utilized in each iteration step of refining the scoring function, hence significantly improving both the effectiveness of the exploration of the binding conformational space and the sensitivity of the positioning of the native binding poses. • MpSDock can be used successfully in structure-based studies on novel designed simplified largazole analogues (SLAs) and benzodiazepine derivatives (BZDs) as human lysine deacetylase (hKDAC)-isoform-selective inhibitors. The tool is written in Bash code (available over the Internet) to be used in Linux operating systems.	7	[Bai et al. 2015]

DockAnalyse	• DockAnalyse is an unsupervised and programmed clustering application which is based on the DBscan clustering technique, which searches for continuities among the clusters generated by the docking output data representation. • The DBscan clustering method is extremely powerful and, also, solves some of the inconsistency problems of the classical clustering methods like, for instance, the treatment of outliers and the dependence of the previously characterized number of clusters. • To extract the significant solutions from the docking output datafile, an unsupervised and programmed clustering program called DockAnalyse, was created with the R software environment. • DockAnalyse was applied to choose the best docking solutions and, therefore, to model the dynamic protein-interaction mechanism among the given proteins. Tridimensional structure studies and representations were made using the following tools: - UCSF Chimera, PyMOL and RasMol. **Benchmark**: • In comparison to the crystallographic protein complex structure, which was obtained from the benchmark set, all of these satisfactory solutions showed a very low RMS (Root Mean Square) deviation. This means that only through DockAnalyse outputs could it be seen in these cases that the dockings were credible before realizing that the RMS deviation was so low.	14	[Amela et al. 2010] .
iMOLSDOCK	• MOLSDOCK is a docking tool that performs operation on rigid	4	[Paul et al. 2017]

	receptor/flexible ligand docking. iMOLSDOCK utilizes mutual orthogonal Latin squares (MOLS) to sample the conformation and the docking pose of the ligand and also the flexible residues of the receptor protein. ● The method then uses a variant of the mean field technique to analyze the sample to arrive at the optimum. It was benchmarked and approved that iMOLSDOCK with a dataset of 44 peptide-protein complexes with peptides.		
LightDock	● LightDock is a multi-scale protein–protein docking procedure fit for accommodating conformational flexibility and an assortment of scoring functions at various resolution levels. ● Implicit use of normal modes during the search and atomic/coarse-grained combined scoring functions yielded improved predictive outcomes with respect to state-of-the-art rigid-body docking, especially in flexible cases.	3	[Jiménez-García et al. 2017]
DarwinDock	● DarwinDock represents is a strategy for small-molecule docking that isolates pose generation and scoring into separate stages, which allow complete binding site sampling followed by efficient, hierarchical sampling. ● Their union criterion for complete sampling allows for various systems to be studied without earlier knowledge of how big a set of poses needs to be to span a given binding site, making the procedure more automatic. The bulky, nonpolar residues with alanine were replaced and	1	[Griffith et al. 2017]

	this process is called "alanization". ● This allows the ligand to interact more closely with polar side chains, which help to orient the ligand.		
LeDock	● LeDock is better than that using AutoDock Vina. Overall, reverse docking is a quick and efficient computational method to recognize the probable target of the compounds with anti-tumor activities, and it can be complementary to the biological testing strategies.	2	[Chen et al. 2017]
ProQDock	● ProQPred use the machine learning method Random Forest trained on previously calculated features from the programs ProQDock and InterPred. By combining some of ProQDock's features and the InterPred score from InterPred the ProQpred method generated a higher performance than both ProQDock and InterPred. ● This work also tried to predict the quality of the PPI model after refinement and the chance for a coarse PPI model to succeed at refinement. The result illustrated that the predicted quality of a coarse PPI model also was a relatively good prediction of the quality the coarse PPI model would get after refinement.	1	[Rörbrink et al. 2016]
Snapdock	● SnapDock is a highly efficient template-based protein– protein docking calculation which utilizes a Geometric Hashing-based structural arrangement plan to align the target proteins to the interfaces of non-redundant protein– protein interface libraries. ● Docking of a couple of proteins using the	2	[Estrin et al. 2017]

	22600 interface PIFACE library is performed in less than 2 minutes on the average. An adaptable version of the algorithm permitting hinge movement in one of the proteins is exhibited as well.		
TagDock	• TagDock was used to compute all geometrically possible docking poses between the domains and evaluated those compatible with experimental distance constraints. The docking represents that were reliable with the limitations were then additionally refined. • In TagDock approaches, an ensemble of solutions with RMSD 2.8 and 1.6 Å, respectively, were obtained. In addition, the average of the ensemble solutions obtained using the two approaches have an RMS deviation of 2.4 Å. The final averaged solution obtained by TagDock-based modeling.	7	[Smith et al. 2013]
evERdock	• It is used for the evaluation of protein-protein complex model structures generated by protein docking prediction (decoys).	1	[Takemura et al. 2018]
PATIDOCK	• PATIDOCK is used for efficiently docking a two domain complex based solely on the novel idea of using the difference between the experimental alignment tensor and the predicted alignment tensor computed by Prediction of Alignment Tensor using Integration (PATI). • The alignment tensor fundamentally contains enough information to accurately dock a two-domain complex, and the two domains can be docked very quickly by pre-computing the right set of data.	22	[Berlin et al. 2010]

ZDOCKpro	• ZDOCKpro is a unique protein-protein docking program that depends on the ZDOCK and RDOCK programs created at Boston University by Professor Zhiping Weng. • It is a valuable tool for protein modelers and structural biologists who need to perform protein-protein docking, just as for bioinformaticians who study protein pathways and computational chemists who inspect protein or peptide ligand docking.	21	[Gay et al. 2007]

1643 **Table 8: Comprehensive List of Scoring Functions.**

1644

S.No.	Scoring Function- Name	PubMed ID
1	DrugScore	30513206
2	CAII-specific fingerprint-based (IFP)	29937490
3	CASF-based scoring function	29517771
4	DITScoreRR	29506237
5	HawkRank	29282565
6	PLANTS	29165067
7	ITScore2	29127582
8	Graph-approach Scoring Function	28921375
9	Template-based Scoring function	28905425
10	GalaxyDock BP2 Scoring Function	28623486
11	RpveScore	28120375
12	AutoDock-GIST	27886114
13	GOLD-based scoring function	27879015
14	DeltaVina RF20	27859414
15	HADDOCK Score	27802573
16	PocketScore	27549813
17	GRIM	27480696
18	GBSA score	27618247
19	QSAR score	27762146
20	SAnDReS	27686428
21	HADDOCK Score	27630991
22	D(3)DOCKxb	27501852
23	XBSF	27195023
24	Vinardo	27171006
25	Glide-Schrodinger Scoring Function	27035259
26	AutoDock4-based scoring function	26629955
27	PMF-based Score	26418299
28	ITScore	26389744
29	AutoDock and AutoDock-Vina based scoring function	26302746
30	Mscomplex	26252196
31	GeauxDock	26250822
32	DARC-scoring function	26181386
33	XBScore	25957658
34	GOLD-based scoring function	28706666
35	SAXS-based scoring function	25897115
36	NMR-based scoring function	25877959
37	Force-field based scoring function	25753725

38	Knowledge-based scoring functions	25746437
39	SCC-DFTB	25296988
40	HWK	25229183
41	Surflex-Dock scoring function	25207678
42	STScore	24623011
43	Wilma-SIE	24474162
44	FFT-based scoring function	24227686
45	PLANS Scoring Function	24163807
46	XBPMF	24072554
47	London dG	23975271
48	SFCscore	23705795
49	SAXS-based scoring function	23633577
50	AuPosSOM	23055752
51	MM-ISMSA	26605745
52	Evolutionary Trace (ET)-based scores	22809378
53	SPIDER	22581643
54	ChemPLP	22371207
55	Scoring function based on weighted residue network	22272103
56	The HYDE	22203423
57	Cscore	22144250
58	eHiTS	22076470
59	MedusaScore	22017385
60	NNScore	22017367
61	ZRANK	21739500
62	S1 and S2 scoring	21644546
63	dG prediction	21612285
64	FACTS	21541955
65	RPScore	21432933
66	Interaction-motif based scoring function	20525216
67	PM6-DH2	21286784
68	RF-Score	20236947
69	MM-GB/SA	20180264
70	AutoDock4-based scoring function	20150996
71	ZDOCK 3.0, ZRANK, ITScore-PP, EMPIRE, and RosettaDock	19938153
72	Glide XP	19421721
73	AMBER	19320460
74	KBSF	19255647
75	EON Scoring	19235177
76	MOSFOM	19210777

77	POLSCORE	19128216
78	Nscore	19066998
79	ROTA	18704939
80	MedusaScore	18672869
81	bootstrap-based consensus scoring (BBCS)	18426197
82	GoldScore	18410085
83	HPNet	18329160
84	ITScore-PP	18247354
85	ChemScore	18041758
86	LigScore2	17985863
87	F-Score	17685604
88	HINT	17346861
89	Thr184	17257425
90	Glide 4.0 XP	17034125
91	Surflex-Dock scoring function	17004701
92	MolDock-PP	16722650
93	DQ3.2beta	16510499
94	RPScore	14635126
95	X-Cscore	12197663
96	GOLD-based Cscore	11858637
97	PMF-Scoring	10896316
98	DOCK-based PMF Score	10411471
99	ITScorePEP	30368849
100	Glide SP	30347931
101	Molegro	30245350
102	CNN	29992528
103	3dRPC-Score	29186336
104	RF-NA-Score	29137330
105	Convex-PL	28921375
106	NMRScore	28406291
107	Wscore	25395044
108	Smog	10072678
109	LUDI (Böhm's score)	10072678

1645

1646

1647

1648

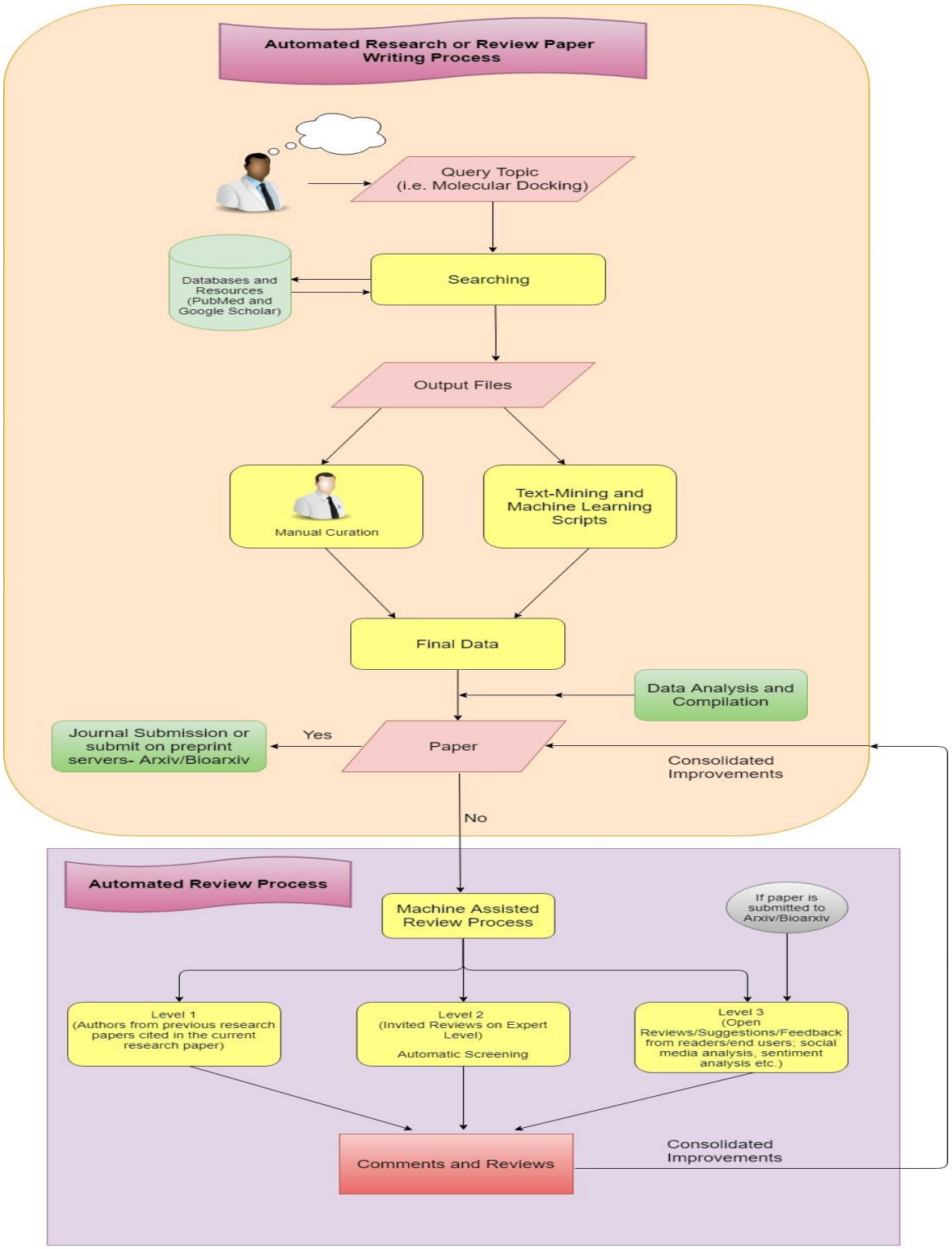
Automated Research or Review Paper Writing Process
Query Topic (i.e. Molecular Docking)
Databases and Resources (PubMed and Google Scholar)
Searching
Output Files
Manual Curation
Text-Mining and Machine Learning Scripts
Final Data
Data Analysis and Compilation
Journal Submission or submit on preprint servers- Arxiv/Bioarxiv
Yes
Paper
Consolidated Improvements
No
Automated Review Process
Machine Assisted Review Process
If paper is submitted to Arxiv/Bioarxiv
Level 1 (Authors from previous research papers cited in the current research paper)
Level 2 (Invited Reviews on Expert Level) Automatic Screening
Level 3 (Open Reviews/Suggestions/Feedback from readers/end users; social media analysis, sentiment analysis etc.)
Comments and Reviews
Consolidated Improvements

Figure 1

Figure 1: We outline two modules used to achieve the stated goals of this review paper. The first module is designed using Perl and Python based text mining scripts which were developed as in-house system. This module is supported by SVM based system which find relationships between two entities in a given text data using the training datasets. This module delivers important information for processing of human expert. This hybrid approach helps in completing the literature mining task in much less time (of the order of 10 fold reduction).

The next module is designed to capture feedbacks from the users at three levels using web based forms. The first level of feedback is obtained from the domain experts cited in the review paper. The second level of feedback is obtained from experts who have been invited by the journal's editor during the peer review process phase. The third level of feedbacks is obtained from the potential end users or general readers. The data from each level is combined to produce final output to determine ranking of the given research manuscript. Individual weights are also assigned to each level of feedback so as to adjust the relative importance. Intuitively, highest weights are assigned to feedback obtained from level 2, followed by level 1 and level 3. The final output or rank of the given manuscript is computed as weighted sum of all levels. This ranking is dynamic in nature and could vary over a period depending upon the continuous feedback obtained from the users (level 3).

REFERENCES SPECIFIC TO TABLE 7:

1. Alexander, J. C. (1998). *U.S. Patent No. 5,813,072*. Washington, DC: U.S. Patent and Trademark Office.
2. Amela, I., et al. (2010). DockAnalyse: an application for the analysis of protein-protein interactions. *BMC structural biology*, *10*(1), 37.
3. Antes, I. (2010). DynaDock: A new molecular dynamics-based algorithm for protein–peptide docking including receptor flexibility. *Proteins: Structure, Function, and Bioinformatics*, *78*(5), 1084-1104.
4. Atanasova, M., et al. (2013). EpiDOCK: a molecular docking-based tool for MHC class II binding prediction. *Protein Engineering, Design & Selection*, *26*(10), 631-634.
5. Bai, F., et al. (2015). An accurate metalloprotein-specific scoring function and molecular docking program devised by a dynamic sampling and iteration optimization strategy. *Journal of chemical information and modeling*, *55*(4), 833-847.
6. Ballante, F., & Marshall, G. R. (2016). An automated strategy for binding-pose selection and docking assessment in structure-based drug design. *Journal of chemical information and modeling*, *56*(1), 54-72.
7. Banitt, I., & Wolfson, H. J. (2011). ParaDock: a flexible non-specific DNA—rigid protein docking algorithm. *Nucleic acids research*, *39*(20), e135-e135.
8. Baspinar, A., et al. (2014). PRISM: a web server and repository for prediction of protein–protein interactions and modeling their 3D complexes. *Nucleic acids research*, *42*(W1), W285-W289.
9. Basu, S., & Wallner, B. (2016). DockQ: a quality measure for protein-protein docking models. *PloS one*, *11*(8), e0161879.
10. Beato, C., et al. (2013). Use of experimental design to optimize docking performance: The case of ligendock, the docking module of ligen, a new de novo design program.
11. Ben-Shimon, A., & Niv, M. Y. (2015). AnchorDock: blind and flexible anchor-driven peptide docking. *Structure*, *23*(5), 929-940.
12. Berlin, K., et al. (2010). Structural assembly of molecular complexes based on residual dipolar couplings. *Journal of the American Chemical Society*, *132*(26), 8961-8972.

13. Berlin, K., et al. (2011). Fast approximations of the rotational diffusion tensor and their application to structural assembly of molecular complexes. *Proteins: Structure, Function, and Bioinformatics*, *79*(7), 2268-2281.
14. Bissantz, C. (2002). *Development and application of new methods for the virtual screening of chemical databases* (Doctoral dissertation, ETH Zurich).
15. Brylinski, M. (2013). Nonlinear scoring functions for similarity-based ligand docking and binding affinity prediction.
16. Burkhard, P., et al. (1999). The discovery of steroids and other novel FKBP inhibitors using a molecular docking program1. *Journal of molecular biology*, *287*(5), 853-858.
17. Cabrera, A. C., et al. (2012). CRDOCK: An Ultrafast Multipurpose Protein–Ligand Docking Tool.
18. Chang, D. T. H., et al. (2005). MEDock: a web server for efficient prediction of ligand binding sites based on a novel optimization algorithm. *Nucleic acids research*, *33*(suppl_2), W233-W238.
19. Chaves, R. C., & Pellequer, J. L. (2013). DockAFM: benchmarking protein structures by docking under AFM topographs. *Bioinformatics*, *29*(24), 3230-3231.
20. Chelliah, V., et al. (2006). Efficient restraints for protein–protein docking by comparison of observed amino acid substitution patterns with those predicted from local environment. *Journal of molecular biology*, *357*(5), 1669-1682.
21. Chen, H. M., et al. (2007). SODOCK: swarm optimization for highly flexible protein–ligand docking. *Journal of computational chemistry*, *28*(2), 612-623.
22. Chen, R. ,et al. (2003). ZDOCK predictions for the CAPRI challenge. *Proteins: Structure, Function, and Bioinformatics*, *52*(1), 68-73.
23. Chen, Y. C., et al. (2014). Docking to multiple pockets or ligand fields for screening, activity prediction and scaffold hopping. *Future medicinal chemistry*, *6*(16), 1741-1755.
24. Chen, Y. Z., & Zhi, D. G. (2001). Ligand–protein inverse docking and its potential use in the computer search of protein targets of a small molecule. *Proteins: Structure, Function, and Bioinformatics*, *43*(2), 217-226.
25. Cheng, T. K. (2008). *Towards a better understanding of non-synonymous single nucleotide polymorphisms: exploring new methods for analysing their pathological implications* (Doctoral dissertation, University of Cambridge).

26. Coelho, E. D., et al. (2016). Computational discovery of putative leads for drug repositioning through drug-target interaction prediction. *PLoS computational biology*, *12*(11), e1005219.
27. Comeau, S. R., et al.(2004). ClusPro: an automated docking and discrimination method for the prediction of protein complexes. *Bioinformatics*, *20*(1), 45-50.
28. Cuzzolin, A., et al. (2015). DockBench: an integrated informatic platform bridging the gap between the robust validation of docking protocols and virtual screening simulations. *Molecules*, *20*(6), 9977-9993.
29. Di Muzio, E., et al. (2017). DockingApp: a user friendly interface for facilitated docking simulations with AutoDock Vina. *Journal of computer-aided molecular design*, *31*(2), 213-218.
30. Di Tommaso, P., et al. (2015). The impact of Docker containers on the performance of genomic pipelines. *PeerJ*, *3*, e1273.
31. Ding, F., & Dokholyan, N. V. (2012). Incorporating backbone flexibility in MedusaDock improves ligand-binding pose prediction in the CSAR2011 docking benchmark. *Journal of chemical information and modeling*, *53*(8), 1871-1879.
32. Douguet, D., et al. (2006). Dockground resource for studying protein–protein interfaces. *Bioinformatics*, *22*(21), 2612-2618.
33. Estrin, M., & Wolfson, H. J. (2017). SnapDock—template-based docking by Geometric Hashing. *Bioinformatics*, *33*(14), i30-i36.
34. Fahmy, A., & Wagner, G. (2002). TreeDock: a tool for protein docking based on minimizing van der Waals energies. *Journal of the American Chemical Society*, *124*(7), 1241-1250.
35. Fradera, X., et al. (2004). Unsupervised guided docking of covalently bound ligands. *Journal of computer-aided molecular design*, *18*(10), 635-650.
36. Fukui, R., et al. (2009, July). iDock: a multifunctional intermediate instrument to improve efficiency of domestic delivery and storage system. In *Proc. of IEEE/ASME Int. Conf. on Advanced Intelligent Mechatronics* (pp. 1939-1945).
37. Gao, Y., et al. (2007). DOCKGROUND system of databases for protein recognition studies: Unbound structures for docking. *Proteins: Structure, Function, and Bioinformatics*, *69*(4), 845-851.

38. Gardiner, E. J., et al. (2003). GAPDOCK: A genetic algorithm approach to protein docking in CAPRI round 1. *Proteins: Structure, Function, and Bioinformatics*, *52*(1), 10-14.
39. Garzon, J. I., et al. (2009). FRODOCK: a new approach for fast rotational protein–protein docking. *Bioinformatics*, *25*(19), 2544-2551.
40. Gay, E., et al. (2007). Structural determinates for apolipoproteinE-derived peptide interaction with the α7 nicotinic acetylcholine receptor. *Molecular pharmacology*.
41. Ghoorah, A. W., et al. (2013). KBDOCK 2013: a spatial classification of 3D protein domain family interactions. *Nucleic acids research*, *42*(D1), D389-D395.
42. Goodsell, D. S., et al. (1996). Automated docking of flexible ligands: applications of AutoDock. *Journal of Molecular Recognition*, *9*(1), 1-5.
43. Goto, J., et al. (2008). ASEDock-docking based on alpha spheres and excluded volumes. *Journal of chemical information and modeling*, *48*(3), 583-590.
44. Griffith, A. R. (2017). *DarwinDock and GAG-Dock: Methods and Applications for Small Molecule Docking* (Doctoral dissertation, California Institute of Technology).
45. Grosdidier, A., et al. (2011). SwissDock, a protein-small molecule docking web service based on EADock DSS. *Nucleic acids research*, *39*(suppl_2), W270-W277.
46. Grosdidier, A., Zoete, V., & Michielin, O. (2011). SwissDock, a protein-small molecule docking web service based on EADock DSS. *Nucleic acids research*, *39*(suppl_2), W270-W277.
47. Gu, J., et al. (2015). MoDock: a multi-objective strategy improves the accuracy for molecular docking. *Algorithms for Molecular Biology*, *10*(1), 8.
48. Hagelueken, G., et al. (2013). mtsslSuite: In silico spin labelling, trilateration and distance-constrained rigid body docking in PyMOL. *Molecular physics*, *111*(18-19), 2757-2766.
49. Hongjian, L et al (2014). "istar: A web platform for large-scale protein-ligand docking."*PLoS One* 9.1: e85678.
50. Hsu, K. C., et al. (2011). iGEMDOCK: a graphical environment of enhancing GEMDOCK using pharmacological interactions and post-screening analysis. *BMC bioinformatics*, *12*(1), S33.

51. Hu, B., & Lill, M. A. (2014). PharmDock: a pharmacophore-based docking program. *Journal of cheminformatics*, *6*(1), 14.
52. Hu, Z., & Southerland, W. (2007). Windock: Structure-based drug discovery on windows-based PCs. *Journal of computational chemistry*, *28*(14), 2347-2351.
53. Huang, B., & Schroeder, M. (2008). Using protein binding site prediction to improve protein docking. *Gene*, *422*(1), 14-21.
54. Huang, S. Y., & Zou, X. (2014). A knowledge-based scoring function for protein-RNA interactions derived from a statistical mechanics-based iterative method. *Nucleic acids research*, *42*(7), e55-e55.
55. Hurwitz, N., et al. (2016). Memdock: an α-helical membrane protein docking algorithm. *Bioinformatics*, *32*(16), 2444-2450.
56. Inbar, Y., et al. (2005). Prediction of multimolecular assemblies by multiple docking. *Journal of molecular biology*, *349*(2), 435-447.
57. Irwin, J. J., et al. (2009). Automated docking screens: a feasibility study. *Journal of medicinal chemistry*, *52*(18), 5712-5720.
58. Jiang, F., Lin, W., & Rao, Z. (2002). SOFTDOCK: understanding of molecular recognition through a systematic docking study. *Protein engineering*, *15*(4), 257-263.
59. Jiménez-García, B., et al. (2013). pyDockWEB: a web server for rigid-body protein–protein docking using electrostatics and desolvation scoring. *Bioinformatics*, *29*(13), 1698-1699.
60. Jiménez-García, B., et al. (2017). LightDock: a new multi-scale approach to protein–protein docking. *Bioinformatics*, *34*(1), 49-55.
61. Kang, L., et al. (2012). A hierarchical method for molecular docking using cloud computing. *Bioorganic & medicinal chemistry letters*, *22*(21), 6568-6572.
62. Khan, J. M., & Ranganathan, S. (2010). pDOCK: a new technique for rapid and accurate docking of peptide ligands to Major Histocompatibility Complexes. In *Immunome Research* (Vol. 6, No. 1, p. S2). BioMed Central.
63. Kim, D. S., et al. (2011). BetaDock: shape-priority docking method based on beta-complex. *Journal of Biomolecular Structure and Dynamics*, *29*(1), 219-242.
64. Labbé, C. M., et al. (2015). MTiOpenScreen: a web server for structure-based virtual screening. *Nucleic acids research*, *43*(W1), W448-W454.

65. Lee, H., et al. (2015). GalaxyPepDock: a protein–peptide docking tool based on interaction similarity and energy optimization. *Nucleic acids research*, *43*(W1), W431-W435.
66. Levieux, G., et al. (2014). Udock, the interactive docking entertainment system. *Faraday discussions*, *169*, 425-441.
67. Li, H., et al. (2004). GAsDock: a new approach for rapid flexible docking based on an improved multi-population genetic algorithm. *Bioorganic & medicinal chemistry letters*, *14*(18), 4671-4676.
68. Li, H., Gao, Z., et al. (2006). TarFisDock: a web server for identifying drug targets with docking approach. *Nucleic acids research*, *34*(suppl_2), W219-W224.
69. Li, L., et al. (2003). RDOCK: refinement of rigid-body protein docking predictions. *Proteins: Structure, Function, and Bioinformatics*, *53*(3), 693-707.
70. Li, L., et al. (2011). ASPDock: protein-protein docking algorithm using atomic solvation parameters model. *BMC bioinformatics*, *12*(1), 36.
71. Liu, M., & Wang, S. (1999). MCDOCK: a Monte Carlo simulation approach to the molecular docking problem. *Journal of computer-aided molecular design*, *13*(5), 435-451.
72. Liu, Y., et al. (2013). FIPSDock: a new molecular docking technique driven by fully informed swarm optimization algorithm. *Journal of computational chemistry*, *34*(1), 67-75.
73. London, N., et al. (2011). Rosetta FlexPepDock web server—high resolution modeling of peptide–protein interactions. *Nucleic acids research*, *39*(suppl_2), W249-W253.
74. London, N., et al. (2014). Covalent docking of large libraries for the discovery of chemical probes. *Nature chemical biology*, *10*(12), 1066.
75. Luo, H., et al. (2016). Molecular docking for identification of potential targets for drug repurposing. *Current topics in medicinal chemistry*, *16*(30), 3636-3645.
76. Lyskov, S., & Gray, J. J. (2008). The RosettaDock server for local protein–protein docking. *Nucleic acids research*, *36*(suppl_2), W233-W238.
77. Malhotra, S., et al. (2015). DOCKSCORE: a webserver for ranking protein-protein docked poses. *BMC bioinformatics*, *16*(1), 127.

78. Martin, L., et al. (2006). kinDOCK: a tool for comparative docking of protein kinase ligands. *Nucleic acids research*, *34*(suppl_2), W325-W329.
79. Mashiach, E., et al. (2008). FireDock: a web server for fast interaction refinement in molecular docking. *Nucleic acids research*, *36*(suppl_2), W229-W232.
80. Mashiach, E., et al. (2010). FiberDock: flexible induced-fit backbone refinement in molecular docking. *Proteins: Structure, Function, and Bioinformatics*, *78*(6), 1503-1519.
81. Meier, R., et al. (2010). ParaDockS: a framework for molecular docking with population-based metaheuristics. *Journal of chemical information and modeling*, *50*(5), 879-889.
82. Morley, S. D., & Afshar, M. (2004). Validation of an empirical RNA-ligand scoring function for fast flexible docking using RiboDock®. *Journal of computer-aided molecular design*, *18*(3), 189-208.
83. Morris, G. M., et al. (2009). AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *Journal of computational chemistry*, *30*(16), 2785-2791.
84. Oda, A., et al. (2009). Validation of ArgusLab efficiencies for binding free energy calculations. *Chem-Bio Informatics Journal*, *9*, 52-61.
85. Ohue, M., et al. (2014). MEGADOCK: an all-to-all protein-protein interaction prediction system using tertiary structure data. *Protein and peptide letters*, *21*(8), 766-778.
86. Ouyang, X., et al. (2013). CovalentDock: automated covalent docking with parameterized covalent linkage energy estimation and molecular geometry constraints. *Journal of computational chemistry*, *34*(4), 326-336.
87. Paul, D. S., & Gautham, N. (2017). iMOLSDOCK: Induced-fit docking using mutually orthogonal Latin squares (MOLS). *Journal of Molecular Graphics and Modelling*, *74*, 89-99.
88. Paul, N., & Rognan, D. (2002). ConsDock: A new program for the consensus analysis of protein–ligand interactions. *Proteins: Structure, Function, and Bioinformatics*, *47*(4), 521-533.
89. Pierce, B. G., & Weng, Z. (2013). A flexible docking approach for prediction of T cell receptor–peptide–MHC complexes. *Protein Science*, *22*(1), 35-46.
90. Pierce, B. G., et al. (2014). ZDOCK server: interactive docking prediction of protein–protein complexes and symmetric multimers. *Bioinformatics*, *30*(12), 1771-1773.

1882 91. Popov, P., et al. (2014). DockTrina: Docking triangular protein trimers. *Proteins: Structure, Function, and Bioinformatics*, *82*(1), 34-44.
1884 92. Rao, S. N., et al. (2007). Validation studies of the site-directed docking program LibDock. *Journal of chemical information and modeling*, *47*(6), 2159-2171.

1886 93. Ravindranath, P. A., et al. (2015). AutoDockFR: advances in protein-ligand docking with explicitly specified binding site flexibility. *PLoS computational biology*, *11*(12), e1004586.
1889 94. Rörbrink, M. (2016). Improving detection of promising unrefined protein docking complexes.
1891 95. Ruiz-Carmona, S., et al. (2014). rDock: a fast, versatile and open source program for docking ligands to proteins and nucleic acids. *PLoS computational biology*, *10*(4), e1003571.
1894 96. Sandeep, G., et al. (2011). AUDocker LE: A GUI for virtual screening with AUTODOCK Vina. *BMC research notes*, *4*(1), 445.
1896 97. Schneidman-Duhovny, D., et al. (2005). PatchDock and SymmDock: servers for rigid and symmetric docking. *Nucleic acids research*, *33*(suppl_2), W363-W367.
1898 98. Schneidman-Duhovny, D., et al.(2005). Geometry-based flexible and symmetric protein docking. *Proteins: Structure, Function, and Bioinformatics*, *60*(2), 224-231.

1900 99. Scholz, C., et al. (2015). DOCKTITE A Highly Versatile Step-by-Step Workflow for Covalent Docking and Virtual Screening in the Molecular Operating Environment. *Journal of chemical information and modeling*, *55*(2), 398-406.
1903 100. Sgobba, M., et al. (2012). Application of a post-docking procedure based on MM-PBSA and MM-GBSA on single and multiple protein conformations. *European journal of medicinal chemistry*, *58*, 431-440.
1906 101. Shimoda, T., et al. (2013, September). MEGADOCK-GPU: acceleration of protein-protein docking calculation on GPUs. In *Proceedings of the International Conference on Bioinformatics, Computational Biology and Biomedical Informatics* (p. 883). ACM.

102. Shin, W. H., et al. (2011). LigDockCSA: protein–ligand docking using conformational space annealing. *Journal of computational chemistry*, *32*(15), 3226-3232.

103. Sircar, A., & Gray, J. J. (2010). SnugDock: paratope structural optimization during antibody-antigen docking compensates for errors in antibody homology models. *PloS computational biology*, *6*(1), e1000644.

104. Smith, J. A., et al. (2013). TagDock: an efficient rigid body docking algorithm for oligomeric protein complex model construction and experiment planning. *Biochemistry*, *52*(33), 5577-5584.

105. Sohn, W., et al. (2013). Endockscope: using mobile technology to create global point of service endoscopy. *Journal of endourology*, *27*(9), 1154-1160.

106. Solernou, A., & Fernandez-Recio, J. (2011). pyDockCG: new coarse-grained potential for protein–protein docking. *The journal of Physical Chemistry B*, *115*(19), 6032-6039.

107. Springer, C., et al. (2005). PostDOCK: a structural, empirical approach to scoring protein ligand complexes. *Journal of medicinal chemistry*, *48*(22), 6821-6831.

108. Sprous, D. G., et al. (2004). OptiDock: virtual HTS of combinatorial libraries by efficient sampling of binding modes in product space. *Journal of combinatorial chemistry*, *6*(4), 530-539.

109. Takemura, K., et al. (2018). Binding free energy analysis of protein-protein docking model structures by evERdock. *The Journal of chemical physics*, *148*(10), 105101.

110. Taylor, R. D., et al. (2002). A review of protein-small molecule docking methods. *Journal of computer-aided molecular design*, *16*(3), 151-166.

111. Thomsen, R. (2003). Flexible ligand docking using evolutionary algorithms: investigating the effects of variation operators and local search hybrids. *Biosystems*, *72*(1-2), 57-73.

112. Thomsen, R., & Christensen, M. H. (2006). MolDock: a new technique for high-accuracy molecular docking. *Journal of medicinal chemistry*, *49*(11), 3315-3321.

113. Tietze, S., & Apostolakis, J. (2007). GlamDock: development and validation of a new docking tool on several thousand protein− ligand complexes. *Journal of chemical information and modeling*, *47*(4), 1657-1672.

114. Torchala, M., et al. (2013). SwarmDock: a server for flexible protein–protein docking. *Bioinformatics*, *29*(6), 807-809.

115. Trosset, J. Y., & Scheraga, H. A. (1999). PRODOCK: software package for protein modeling and docking. *Journal of computational chemistry*, *20*(4), 412-427.

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

117. Tuszynska, I., et al. (2015). NPDock: a web server for protein–nucleic acid docking. *Nucleic acids research*, *43*(W1), W425-W430.

118. Van der Meer, P. F., et al. (2000). Compodock, a new device for sterile docking. *Transfusion*, *40*(6), 682-686.

119. Verkhivker, G. M., Bouzida, D., Gehlhaar, D. K., Rejto, P. A., Arthurs, S., Colson, A. B., & Rose, P. W. (2000). Deciphering common failures in molecular docking of ligand-protein complexes. *Journal of computer-aided molecular design*, *14*(8), 731-751.

120. Vistoli, G., et al. (2010). Homology modeling and metabolism prediction of human carboxylesterase-2 using docking analyses by GriDock: a parallelized tool based on AutoDock 4.0. *Journal of computer-aided molecular design*, *24*(9), 771-787.

121. Vorobjev, Y. N. (2010). Blind docking method combining search of low-resolution binding sites with ligand pose refinement by molecular dynamics-based global optimization. *Journal of computational chemistry*, *31*(5), 1080-1092.

122. Wang, Z., et al. (2016). Comprehensive evaluation of ten docking programs on a diverse set of protein–ligand complexes: the prediction accuracy of sampling power and scoring power. *Physical Chemistry Chemical Physics*, *18*(18), 12964-12975.

123. Wu, G., & Vieth, M. (2004). SDOCKER: a method utilizing existing X-ray structures to improve docking accuracy. *Journal of medicinal chemistry*, *47*(12), 3142-3148.

124. Xu, Y., et al. (2016, April). SDN docker: Enabling application auto-docking/undocking in edge switch. In *Computer Communications Workshops (INFOCOM WKSHPS), 2016 IEEE Conference on* (pp. 864-869). IEEE.

125. Xue, L. C., et al. (2014). DockRank: Ranking docked conformations using partner-specific sequence homology-based protein interface prediction. *Proteins: Structure, Function, and Bioinformatics*, *82*(2), 250-267.

126. Yadav, D. K., et al. (2010). Development of QSAR model for immunomodulatory activity of natural coumarinolignoids. *Drug design, development and therapy*, *4*, 173.

127. Yan, C., Xu, X., & Zou, X. (2016). Fully blind docking at the atomic level for protein-peptide complex structure prediction. *Structure*, *24*(10), 1842-1853.

128. Yan, Y., et al. (2017). HDOCK: a web server for protein–protein and protein–DNA/RNA docking based on a hybrid strategy. *Nucleic acids research*, *45*(W1), W365-W373.

129. Yan, Y., et al. (2018). HSYMDOCK: a docking web server for predicting the structure of protein homo-oligomers with Cn or Dn symmetry. *Nucleic acids research*.

130. Yang, J. M., & Chen, C. C. (2004). GEMDOCK: a generic evolutionary method for molecular docking. *Proteins: Structure, Function, and Bioinformatics*, *55*(2), 288-304.

131. Yu, J., et al. (2016). InterEvDock: a docking server to predict the structure of protein–protein interactions using evolutionary information. *Nucleic acids research*, *44*(W1), W542-W549.

132. Zhang, C., & Lai, L. (2011). SDOCK: A global protein-protein docking program using stepwise force-field potentials. *Journal of computational chemistry*, *32*(12), 2598-2612.

133. Zhang, L., et al. (2017). Virtual screening approach to identifying influenza virus neuraminidase inhibitors using molecular docking combined with machine-learning-based scoring function. *Oncotarget*, *8*(47), 83142.

134. Zhang, S., & Du-Cuny, L. (2009). Development and evaluation of a new statistical model for structure-based high-throughput virtual screening. *International journal of bioinformatics research and applications*, *5*(3), 269-279.

135. Zhang, Z., & Lange, O. F. (2013). Replica exchange improves sampling in low-resolution docking stage of RosettaDock. *PLoS One*, *8*(8), e72096.

136. Zhao, Y., & Sanner, M. F. (2007). FLIPDock: docking flexible ligands into flexible receptors. *Proteins: Structure, Function, and Bioinformatics*, *68*(3), 726-737.