

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

6,900

Open access books available

185,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

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

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.  
For more information visit [www.intechopen.com](http://www.intechopen.com)



# Past, Present, and Future of Molecular Docking

*Thuluz Meza Menchaca, Claudia Juárez-Portilla  
and Rossana C. Zepeda*

## Abstract

The interface of any given ligand and protein—normally considered a macromolecule—of a known or predicted/modeled structure can be computed by determining each potential ligand position, resulting in an array of possibilities which are finally expressed in numerical energy values based on their thermodynamic affinity. Over the past few decades, this premier approach technique has proved to be crucial as an automated method in drug design and discovery, as well as in other fields. Data are retrieved from contour surface calculations for each ligand probe and can be analyzed to delineate regions of attraction on the basis of energy levels. Negative energy levels from contours are used to infer protein-ligand affinity clefts and are therefore relevant to drug design. Accordingly, molecular docking, framed as the “new microscope,” is part of a group of *in silico* computational techniques that enable the behavior of molecular chemistry to be analyzed and predicted in an inexpensive manner. From the starting point of framing the key terms in the binomial macromolecule-ligand docking approach, this chapter presents an introductory description of the progress made in this field of research over the past several years, in addition to present and future perspectives. This chapter presents a broad plethora of possibilities arising from the old docking alternatives to the current software technology and critically dissects and discusses the emerging trends. Despite the emergence of more degrees of freedom, a number of flexible conglomerates have not been well developed, and there are still computational limitations to solve, including several features in the focused technique. The present goals, such as molecular flexibility, binding entropy, and the presence of ions and solute conditions, are revisited with the purpose of anticipating the challenges, goals, and achievements in this field over the next few years or decades.

**Keywords:** molecules, modeling, structure, proteins

## 1. Introduction

In biology, dissimilar molecules dock and interact to enable the perpetuation of the primordial logistics of living organisms. Molecular docking methodologies can be used to identify the interaction between a small ligand and a target molecule and to determine whether they could behave in combination as the binding site of two or more constituent molecules with a given structure. The comparison of docking molecules for proteins, other drug-like molecules, or even fragments from the original molecule enables a pool of prominent candidates to be calculated with listed values.

Interestingly, a wide spectrum of molecular binding interactions can be explored with this technique, including lipid-protein, lipid-lipid, enzyme-substrate, drug-enzyme, drug-nucleic acid, protein-nucleic acid, nucleic acid-nucleic acid, protein-drug, and protein-protein potential affinities, with key functions in every molecular biological or biochemical stage, as well as structural coupling [1–2].

The analysis of the binding scores between the constituent molecules in molecular recognition is essential to explain the constitutive processes and subsequently suggest a possible therapy in the context of a particular disease. The molecular docking in silico approach seeks the optimization of this process, not only in terms of techniques but also in relation to time and economic resources. For instance, there is no microscope with a sufficient power of resolution to capture an image at the dynamic (real-time) molecular level, and accordingly, theoretical and computational approaches can be used to predict the best binding and most probable trajectories. Faster techniques and reduced resources are related to efficiency, in contrast to in vitro approaches, in which the examination of every synthesized and purified protein can have higher time and material costs. On average, traditional in vitro research can take about a decade to complete and can cost around 800 million USD; in silico method importantly diminishes these costs [3]. As such, due to the difficulties in determining the structures of complexes, in silico approaches, including molecular docking, are suitable for predicting binding modes by investigating thousands of ligand positions using the lowest energy score analyzed.

Since 1975, the development of high-throughput protein purification X-ray crystallography and nuclear magnetic resonance spectroscopy has continued to advance, predominantly contributing to a better understanding of the structural details of macromolecules and complexes with ligands [4]. Molecular docking, as with many other in silico tools, has become more common and easier to apply to the field of drug discovery; however, it is not entirely dependent on molecular structure databases. It is not impossible to work with molecules that are absent from the databases, as they can be modeled by using one or multiple similar structures to build a novel chimeric output that can mimic the original molecule. In the docking process, the parameters can be further adjusted to test the function of the drug molecule versus a particular target molecule.

After the molecular docking has been performed, the software executes a systematic search on the algorithm, in which the ligand conformation is recurrently approached until the minimum energy conformation is identified. The final result will have a negative value of  $\Delta G$  (U total in kcal/mol), in which a number of electrostatic and van der Waals energy variables will have been synthesized. These energies are related through the interaction between two molecules. This association allows a final scoring function to classify the candidate positions through the driving forces of the specific interactions to be obtained. The structural shape and electrostatic forces of both the ligand and the target molecule at specific binding-site surfaces are key aspects in biological complementarity systems. In the drug discovery field, several key aspects must be considered when predicting whether the molecule will bind with the receptor target, such as the structural shape and electrostatic interactions of the protein-ligand, ligand-ligand, or protein-protein. In this sense, several physicochemical parameters, including the van der Waals forces, Coulombic interactions, and the formation of hydrogen bonds, play relevant roles. The combination of all these values and potential binding is predicted by a docking score. Essentially, for drug design, it is possible to use a rigid system in which a rotational and translational space in six dimensions is explored to fit the ligand into a specific binding structure site [5].

The constantly growing number of biological targets for the design of rational structure-based ligands in public databases has gained interest in the research

community. In the drug discovery field, the essential processes in computational docking are the design of the ligand and the search for targets of the existing candidate ligands. The latter are used to predict a reliable binding affinity, in which the best possible physicochemical prediction of how the target and ligand will interact is made. A strategy to enhance the selection of drug candidate ligands is based on the scores obtained from *in silico* approaches. These scores not only significantly reduce the amount of inefficient compounds synthesized but also decrease the amount of unnecessary biological tests by taking into account valuable information about crucial binding elements in a given ligand-receptor conglomerate. Molecular docking approaches are used to calculate the scores of ligand-binding types and linking affinities. The estimation of reliable ligand-binding associations and modes is a difficult challenge. During the last few decades, the scientific community has gradually shown an increasing interest in molecular docking methods, illustrated by the increase in references and the number of publications in the field [6]. Nevertheless, there is currently no standard consensus regarding the criteria that should be used to classify a docking mode as correct or incorrect. Most docking methods are based on the use of general scoring functions to predict molecular suitability for a wide range of applications. In order to accomplish what is needed, a reliable scoring function, reasonable protein flexibility, and a treatment for ligand conformational changes are required.

In the context of molecular biology, the interactions between molecules are key to understanding the mechanisms that underlie a particular biomedical event. The latest achievements have been the improvement of computational methods essential to the process of drug discovery, modeling in the preclinical stage, and the actual analysis of putative binding interactions. It is possible to conduct exploratory work by examining the best score function values or by using a large set of multivariate experimental data. In both cases, it is possible to analyze how changes in ligands or macromolecules can have an effect on their interactions by validating the associated biological processes, with the aim of gaining a better understanding of the interplay between the biomolecular functions of the bioactive candidates through the characterization of the kinetics and binding score values imperative to their molecular recognition. In order to better understand the historical and conceptual implications of the development of this interesting and well-established technique, past and present achievements must be considered, as well as the current limitations with the potential to change the course of the technological methods developed in the future. In comparison to “wet lab” experimental procedures such as, e.g., microarray technology or even sequencing, virtual screening is inexpensive and efficient. However, several considerations need to be taken into account [7]. Overall, computational methods have been a recurrent option due to the focus approximation of the analysis.

## **2. The development of molecular docking techniques**

As one of the most commonly used approaches since the 1980s, the experimental data obtained through molecular docking techniques have grown at an increasing rate since the approach was first established. Programs configured through different algorithms for molecular docking analysis have been developed on an almost yearly basis, significantly improving pharmaceutical research [6]. The first algorithms were designed for protein-protein interactions. Along with the scoring function, which is used to determine the best binding poses, algorithms designed to calculate the best geometrically complementary shapes as rigid bodies are necessary to identify the most favorable orientations and conformational bindings with the potential to confer a putative drug candidate.

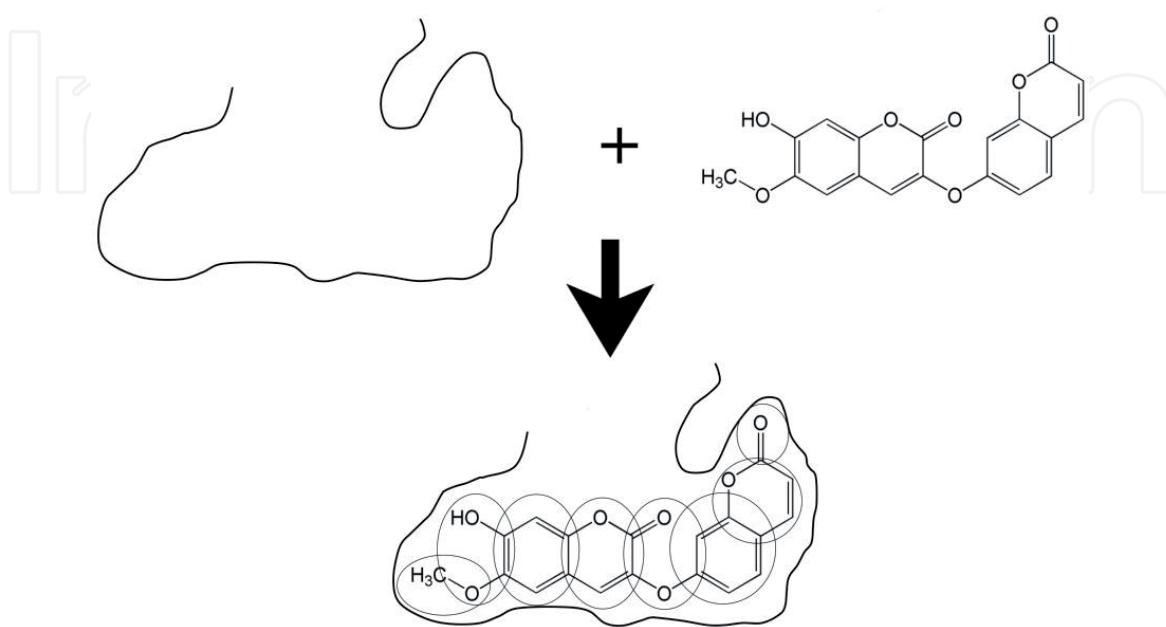


The gradual achievement of more powerful and complex algorithms with the addition of further parameters has paralleled computational technological advances over the last few decades. In order to achieve optimum flexibility, *in silico* methods use different tools with different approaches. Docking software depends on the algorithms employed, which comprise three different kinds: systematic, stochastic, or deterministic.

In the beginning, calculation algorithms that consider docking complexes to be rigid structures were used. In rigid docking, the objective is to match the ligand to the protein receptor, with the main aim being the generation of as many poses as possible in order to achieve the optimum of all poses. Through this process, all possibilities are considered heuristically to identify a group of complementary matches that present the most favorable van der Waals forces between the ligand and the macromolecule receptor. Intermolecular interaction calculations avoid any flexibility but nevertheless have a level of freedom dependent on a 3x3 matrix plus the vector rotation. This means that three rotational and three translational degrees of freedom cover all possible moves in three-dimensional space within the active site. However, no binding is permitted, as the macromolecular structures are simplistically represented as solid structures located under a center of mass and longitude [8].

The earliest work was performed using structural shape contacts, in which the fitting of outlines enables the best possible complementary configuration between two proteins to be identified [9]. A little later, a shape matching strategy algorithm was used by Kuntz and collaborators in UCSF<sup>8</sup> to continue searching for possible configurations using the geometric distance between the ligand atoms and the macromolecule or receptor spheres (**Figure 1**).

In this method, the ideal intersection or match between the ligand and receptor is viewed as a “negative image” that represents the active site. The image is produced by covering the receptor surface region and overlapping spheres with a solvent, in which a part of the overlapping spheres comprises the actual binding site. This constitutes the fundamentals of the DOCK search algorithm [10]. A few years later, Kuntz also developed a more advanced approach by conferring flexibility to the ligand; however, this variant is still categorized as “flexible docking.”



**Figure 1.**

Top left, binding site; top right, ligand. Down below conjugate with geometrical fitness functional group related proposed by the earliest docking algorithm model.

Subsequently, the investigation of HIV-1 protease using this approach was notable for leading to the technique's exponential use in drug discovery [11].

Following the pioneering work from Kuntz, a different approach was taken a decade later in order to develop an improved new geometric recognition method, which was developed through an algorithm called Fourier transformation [12]. For the first time, the molecules could be described by a digital model, allowing their interior and exterior parts to be distinguished. This novel method allows faster calculation by determining the surface of contact, overlap, and approximation using the six degrees of freedom. In this method, molecules are considered rigid bodies, and the changes in structure have the degrees of freedom. This technique makes it possible to process atomic coordinates, and Zdock represents an example of this approach. Nevertheless, rigid-body algorithms are very erratic and ineffective in terms of any structural and conformational change arising due to the interface between the ligand and the receptor. In this context, new alternatives to enable torsions and angle movement became a matter of interest. In the same period, a new semiflexible docking innovation was achieved using the HADDOCK protocol [13], which involves rigid-body docking complemented by semiflexible optimization in order to describe possible torsion angles in the main backbone and side chains. Unlike the previous Fourier transformation method [12], which uses a grid, this method adopts a Cartesian approach with particular coordinates, in which one of the two molecules is flexible and the solvent can be selected. One of the two molecules therefore needs to be small in order to be computationally possible in terms of the number of conformational variations. Other methods also attempt to describe flexible bodies undergoing rotational conformational, rotational, and translational changes, mimicking the nature of biological molecules. In this category, both the ligand and the receptor that are modeled by simulating protocols are flexible. However, the flexibility needs to be lowered to make computational configuration possible. In the end, flexible docking approaches offer a more precise technique capable of imitating in vivo behavior of the possible structural conformations.

In flexible docking, there are two different logarithmic approaches, deterministic incremental construction and stochastic. Systematic incremental construction algorithms are most commonly used, which gradually develop binding predictions on the basis of all possible ligand-binding poses covering all specified areas, e.g., DOCK [14], Glide [15], LUDI [16] FlexX [17], Hammerhead [18], and Surflex [19], in which on-the-fly incremental ligand construction is implemented. In this method, the number of analyses grows in line with increases in the degrees of freedom as part of anchor-and-grow methods. In a different example, in eHiTS, the ligand is fragmented, and each piece is tested for rigid docking, commonly based on library screening for the best conformations to religate the fragments and test their flexibility.

A different approach randomizes probabilistic or stochastic algorithms to selectively reject or accept configurations through the criteria spectrum, in which computational efforts are optimized, e.g., AutoDock [20], DARWIN [21], Monte Carlo [22], and GOLD [23]. By the middle of the 1990s, this technique was the point of origin of a diverse set of methods that are most commonly present in the genetic algorithm, named after Darwin's theory of evolution, in which the ligand is interpreted as a chromosome and its fragments are considered genes [24]. Every gene exhibits conformational behavior due to its torsional/translational nature. During computational analyses, the information is transmitted and altered through stochastic crossover and mutational events evolving through specific parameters. The changes improve the conformational binding pose from the ligand and the receptor, e.g., Lamarckian (AutoDock). In the case of the Monte Carlo stochastic variant that produces randomized translational conformations, the most thermodynamically

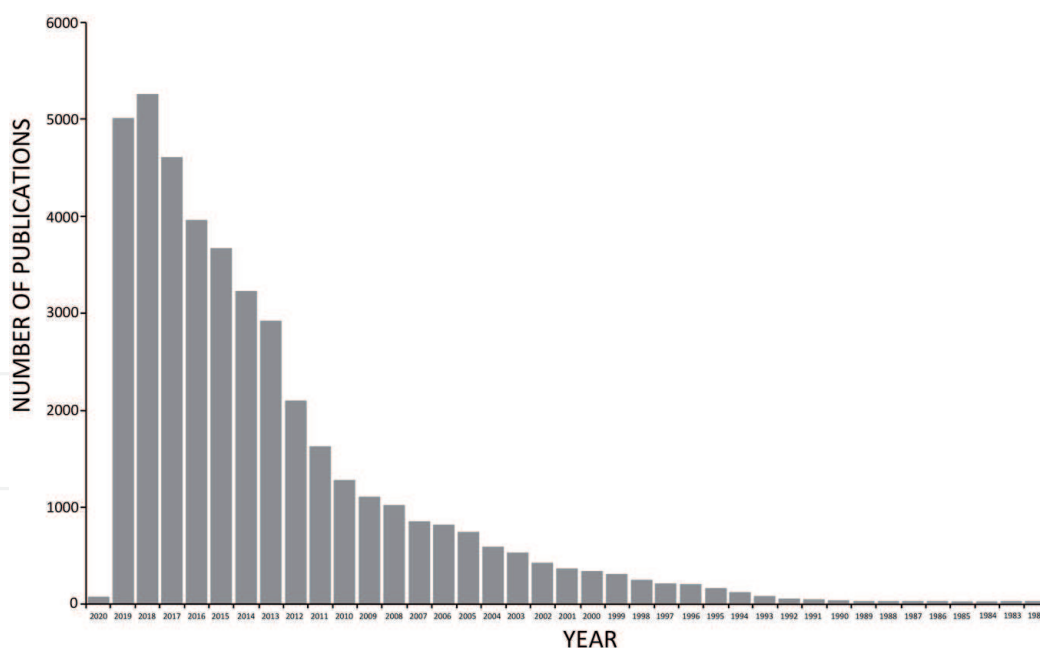
stable potential bindings are explored by focusing on the local minimum energy using a decision criteria parameter that is based on a temperature reaction, called Metropolis. The flexibility also alternates with rigid rotation, displaying several parameters at once. A more recent development is the deterministic method, which has been used for Newton equation simulations and also employs Monte Carlo methods that can measure trajectories, using Amber, Charm, and GROMACS; however, this scope forms the focus of the present work, and wide reviews have been provided by other researchers [25–27].

### **3. Molecular docking at present: a diverse and common approach**

The drug discovery informatics market had an estimated value of 713.4 million USD in 2016 [28]. The presence of *in silico* tools that can allow the computation of data flowing from diverse methodology pathways in parsimony with medical chemistry can be synergistic in terms of upgrading the market and are well-known in the scientific literature. In this manner, molecular docking has been consolidated as a useful technique among sequence analysis platforms, molecular modeling, and clinical training management. The use of molecular docking in each of these fields is enhancing drug discovery in the pharmaceutical and biotechnology sector. As it comprises several stages and workflows, the discovery of new drugs relies on *in silico* tools and molecular docking in particular to simplify the overall process.

A crucial factor is the steadily rising number of structures stored in the Protein Data Bank (PDB). The PDB is the most robust, currently storing over 151,000 structures and counting. The 3D structure information bank includes a large set of proteins, lipids, carbohydrates, and nucleic acids, in both single structures and complexes [29]. On the other hand, nearly a hundred different forms of molecular docking software are available, which offer analogous implementations with various implementation options. There has been rapid progress in developing faster architecture based on graphics processing unit clusters, more adequate algorithms for optimized computational analysis, and the tracking of ligand-receptor binding expressed in scoring functions.

Although there is a need to maintain computational equipment, the associated expenses are certainly lower than the costs of “wet lab” experiments, and molecular docking is therefore an affordable technique. One of the most challenging tasks in bioinformatics sciences is undoubtedly the development of new and effective drugs, which is currently an almost mandatory step before wet lab experiments. In structure-based drug modeling, obtaining the most accurate and efficient model of ligand-receptor binding is a crucial step and is a suitable starting point for further evaluation to test new compounds or drug candidates, but also and no less importantly, to discard the improbable candidates. Molecular-ligand docking is a significant tool in pharmacology at present and an important area of drug discovery that has comprised a central node of important achievements over the current century. As an interdisciplinary process of multiple joint efforts mainly from the pharmaceutical sector, biotechnological companies, and academic researchers, as well as many other fields, the process is highly complex and requires the most accurate and precise tools and methodologies. This has been enhanced by an increasing number of protein coordinates and the high number of available software programs that are constantly evolving with more sophisticated levels and a wider field of applications, in combination with more numerous candidates. In order to discover new drugs, as well as improve the existing ones, it is necessary to understand the targets as well as the nature of the possible drug candidates. *In silico* bioinformatics approaches have attracted increased interest due to the results of post-genomic era sequencing.



**Figure 2.**  
 Chart bar displaying paper publications per year (1982–2020) (NCBI, accessed on January 12, 2019).

Due to the limited set of protein-coding genes, the complexity is much higher due to posttranscriptional modifications, prosthetic groups, multimeric complexes, and other various phenomena, clearly demonstrating the need to better understand their nature to fulfill biomedical objectives. Interestingly this year's (2019) publications account for the first time a pause in the upper trend of docking publication number (**Figure 2**). This may be symptomatic on how the future holds already crucial challenges.

#### 4. Future challenges, endeavors, and perspectives

The drug discovery informatics market is estimated to grow from 1.5 billion in 2016 to 2.84 billion by 2022 and may continue expanding. Accordingly, there is currently a rising demand for the discovery and implementation of novel informatics solutions. The major factors driving the expansion of the global market include the transition from pure research to clinical treatment. More skilled professionals, interdisciplinary backgrounds, and the high pricing of informatics software may have a crucial impact on the growing market. At present, a number of well-established applications have been made available for free or as paid software or services. However, many challenges remain to be addressed to enable the full potential of this powerful technique to be realized.

Nevertheless, in the case of pharmacology, the synergistic aspect is an important chemical phenomenon in which two different biomolecules with different origins can have an exponential effect in combination that is greater than their separate effects. If it is determined that a particular structure is more favorable [30] in terms of the docking score and it may be correlated with synergism, this can be secondary, due to the fact that a molecular docking procedure has not been developed to examine it in a particular scoring function. A linear/quadratic formula could be developed to measure synergy by discriminating between synergistic, additive, or antagonistic effects, which can be expressed both qualitatively and quantitatively. In this sense, further work is needed to investigate how the chemosensitivity between a macromolecule and ligand could be detected once more than one ligand



is included. Although unmanageable amounts of data make this process difficult, it is possible to analyze the small targets that are the most restricted to the binding site being examined, especially in drug-protein analysis. System biology models that depend on a drug synergy test need to be developed in a more comprehensive manner, perhaps by including qualitative features in combination with the quantitative. In this sense, a novel input could be developed in computational docking analysis to enable, e.g., the measurement of molecular signaling that has been established to be part of several components, ligands, or targets. These systematic synergy modeling methods could support drug synergy research with the aim of improving the accuracy of experimental results.

An improvement of the molecular structure databases is necessary for further development. Filters are needed to ensure the structural models they contain are of a better quality, as this will influence the reliability of the results. The PDB was established in 1971 as a pioneer crystal structure database, and today it is the most common source for molecular in silico modeling, harboring more than 150,000 experimentally proven 3D models. However, there is no guarantee that the chosen structures are error-free, including even those with excellent geometrical parameters, and this must be taken into account. High-quality statistics are not an indication that the structure is perfect. Therefore, an improvement of their quality, protocols, and validation would allow the construction of better models that could be valuable in the inevitable task of structure refinement. However, a better model will not be more informative in terms of more detailed biological information, which means that the interpretation of a scientist will be necessary. However, the confirmation of outcomes and the precision of the docking tool in a certain interaction can be tested. Although docking strategies have become more complex, false positives are a recurrent issue with this technique, and as such, refining the structures stored in the PDB will undoubtedly lead to an improvement and better results from pharmacodynamics studies [31].

Those who devote their time to molecular docking are well aware of the large number of docking techniques. In the years to come, docking experiments will need to be more consistent in terms of the outputs generated by different docking methods. Using meta-experimental databases, including a large-scale and diverse variety of targets and ligands, comparisons of scoring functions have shown that accuracy and reportability are far from being reached. A standardized common workflow that follows the same procedures and is associated with the same advantages and issues is therefore necessary. A streamlined validation process to define standard test protocols needs to be agreed for every aspect of the docking method; otherwise there will be a lack of reproducibility in the output process used by each research group and for each given software [32].

The interaction model of the ligand and the active site must achieve the most optimum site of recognition. Docking ensembles using rigid proteins can be slightly inaccurate. Through the ensemble, the protein can fluctuate according to the relative energy, with more time spent in the lowered energy structure. On the other hand, the conformations of ligands fluctuate partially, making the whole ensemble more stable. This can be misleading for dockings that are not flexible, due to the fact that a given conformation may not be the most stable choice in the structure. Up-to-date docking scores have been oriented for machine learning scoring and mainly consist of four building blocks: descriptors, a model, a training set, and a test set. Currently, SFCscore, NNscore, or RFscore represents prominent examples of nonlinear and nontrivial correlations of data in order to avoid obstacles to interpretation [33]. Techniques that provide free access to the scoring function are still a minority and more options are needed, particularly those with open access. The number of poses needs to be exhaustive; however, this has not

been well-established. In this sense, we can state that the sensitivity of the original conformation of the ligands remains unanswered. Furthermore, in the case of multidomain proteins, proteins are frequently composed of more than a single effector domain, and this should be taken into consideration.

With regard to a different aspect, how water is placed around the binding site is not a straightforward problem to solve, although recent studies have proposed the use of this parameter as functionally valid in specific contexts [34] within and around the conglomerate binding site. X-ray crystallography is the most extensively used tool for predicting 3D conformational structure; however, the actual output is only partially informative, due to the fact that the density limits are out of resolution and, on occasion, the electron density can be of insufficient quality. Future efforts need to endorse novel alternatives to increase the capacity and parameters that can be used in every aspect of a given analysis, not only in terms of water but also the physiological solutes found in nature and even protonation, in addition to the pH potency spectra.

An understanding of the biological functions and roles of a protein in a particular cell or tissue is highly relevant in determining the role of a protein's structure, including all of its functional domains. Genome-wide studies have demonstrated that multidomains are present in over 70% of eukaryotic proteins. Nevertheless, protein-folding studies usually consider only single domains and are therefore not focused on the mechanisms in multidomains that can even influence the folding structure [34]. Very crucial obstacles are involved in multidomain docking analyses. In some examples, the understanding of intermolecular movement can be restricted by rigid docking methodologies that lack the ability to consider the effect of multiple domains in a single macromolecule. A given protein is not always present in a static and simplistic single conformational shape but can be present in a collection of scaffolds, stages, and intersections of conformational shapes. As a consequence, the free energy landscape can be profoundly affected, distinctively changing the scoring function's output. This continues to present a major issue [35].

To improve modeling, the role played by multiple molecules in the context of a certain reaction is an indispensable step that must be considered. At the current stage of technology, this does not fall under the current scope of molecular docking, due to the fact that the processes are far too complex and it is difficult to manage all of the interactions that occur during a molecular binding and reaction. In order to mimic how chemistry works in nature, the inclusion of more than two factors (ligand/macromolecule) where methodologically possible would be a priority to enable the possible interactions in a molecular group to be predicted. Although a few software packages use this approach, in the future, it needs to become more common in other methods to address the binding modes of ligands in assessments with higher stoichiometry using multiple ligand complexes against the molecular target. Additionally, as stated earlier in this work, it would be of great interest to evaluate the synergy of ligand combination conjugates.

## 5. Conclusion

Over the last four decades, molecular docking has improved quite remarkably, contributing to the enhancement and improvement of pharmacology in addition to many different areas of applied and molecular biology. After the first complete draft of the Human Genome Project was announced in 2003, the scientific community concluded that there are far fewer protein-coding genes than expected and it has therefore been swift to study how molecules interact by investigating more possible target bindings of a given molecule. The increasing demand for molecular

docking has paralleled the revolutionary advancement of its technological background. Nevertheless, several biochemical and physical properties of proteins, particularly at the surface of contact, need to be included in docking algorithms in conjunction with those already present. On the other hand, the question of how to diminish unnecessary calculations and outputs from undesirable rotations and therefore translations is a big challenge to be considered in the near future, especially in virtual screening. The right implementation needs to be standardized, and closer multi- and interdisciplinary teams must overcome this challenge in order to fine-tune this already widely explored technique.

### **Author details**

Thuluz Meza Menchaca<sup>1\*</sup>, Claudia Juárez-Portilla<sup>2</sup> and Rossana C. Zepeda<sup>2</sup>

1 Laboratorio de Genómica Humana, Facultad de Medicina, Universidad Veracruzana, Veracruz, Mexico

2 Centro de Investigaciones Biomédicas, Universidad Veracruzana, Veracruz, Mexico

\*Address all correspondence to: [thuluz@gmail.com](mailto:thuluz@gmail.com)

### **IntechOpen**

© 2020 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

## References

- [1] Hedger G, Sansom MSP. Lipid interaction sites on channels, transporters and receptors: Recent insights from molecular dynamics simulations. *Biochimica et Biophysica Acta*. 2016;**1858**(10):2390-2400
- [2] Alejandra Hernández-Santoyo AYT-B, Altuzar V, Vivanco-Cid H, Mendoza-Barrera C. Protein-protein and protein-ligand docking, protein engineering. In: Ogawa T, editor. *Technology and Application*. IntechOpen; 2013
- [3] Lavecchia A, Di Giovanni C. Virtual screening strategies in drug discovery: A critical review. *Current Medicinal Chemistry*. 2013;**20**(23):2839-2860
- [4] Love RA, Koetzle TF, Williams GJB, Andrews LC, Bau R. Neutron diffraction study of the structure of Zeise's salt,  $KPtCl_3(C_2H_4) \cdot H_2O$ . *Inorganic Chemistry*. 1975;**14**(11):2653-2657
- [5] Alberg DG, Schreiber SL. Structure-based design of a cyclophilin-calcineurin bridging ligand. *Science*. 1993;**262**(5131):248-250
- [6] Mark Andrew Phillips MAS, Woodling DL, Xie Z-R. Has molecular docking ever brought us a medicine? In: Vlachakis DP, editor. *Molecular Docking*. IntechOpen; 2018
- [7] Wong CF. Flexible receptor docking for drug discovery. *Expert Opinion on Drug Discovery*. 2015;**10**(11):1189-1200
- [8] Kuntz ID, Blaney JM, Oatley SJ, Langridge R, Ferrin TE. A geometric approach to macromolecule-ligand interactions. *Journal of Molecular Biology*. 1982;**161**(2):269-288
- [9] Levinthal C, Wodak SJ, Kahn P, Dadivanian AK. Hemoglobin interaction in sickle cell fibers. I: Theoretical approaches to the molecular contacts. *Proceedings of the National Academy of Sciences of the United States of America*. 1975;**72**(4):1330-1334
- [10] Ewing TJA, Kuntz ID. Critical evaluation of search algorithms for automated molecular docking and database screening. *Journal of Computational Chemistry*. 1997;**18**(9):1175-1189
- [11] DesJarlais RL, Seibel GL, Kuntz ID, Furth PS, Alvarez JC, Ortiz de Montellano PR, et al. Structure-based design of nonpeptide inhibitors specific for the human immunodeficiency virus 1 protease. *Proceedings of the National Academy of Sciences of the United States of America*. 1990;**87**(17):6644-6648
- [12] Katchalski-Katzir E, Shariv I, Eisenstein M, Friesem AA, Aflalo C, Vakser IA. Molecular surface recognition: Determination of geometric fit between proteins and their ligands by correlation techniques. *Proceedings of the National Academy of Sciences of the United States of America*. 1992;**89**(6):2195-2199
- [13] Dominguez C, Boelens R, Bonvin AM. HADDOCK: A protein-protein docking approach based on biochemical or biophysical information. *Journal of the American Chemical Society*. 2003;**125**(7):1731-1737
- [14] Venkatachalam CM, Jiang X, Oldfield T, Waldman M. LigandFit: A novel method for the shape-directed rapid docking of ligands to protein active sites. *Journal of Molecular Graphics & Modelling*. 2003;**21**(4):289-307
- [15] Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, et al. Glide: A new approach for rapid, accurate docking and scoring. 1. Method



and assessment of docking accuracy. *Journal of Medicinal Chemistry*. 2004;**47**(7):1739-1749

[16] Bohm HJ. The computer program LUDI: A new method for the de novo design of enzyme inhibitors. *Journal of Computer-Aided Molecular Design*. 1992;**6**(1):61-78

[17] Kramer B, Rarey M, Lengauer T. Evaluation of the FLEXX incremental construction algorithm for protein–ligand docking. *Proteins: Structure, function, Bioinformatics*. 1999;**37**(2):228-241

[18] Welch W, Ruppert J, Jain AN. Hammerhead: Fast, fully automated docking of flexible ligands to protein binding sites. *Chemistry & Biology*. 1996;**3**(6):449-462

[19] Spitzer R, Jain AN. Surflex-dock: Docking benchmarks and real-world application. *Journal of Computer-Aided Molecular Design*. 2012;**26**(6):687-699

[20] Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, et al. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *Journal of Computational Chemistry*. 2009;**30**(16):2785-2791

[21] Taylor JS, Burnett RM. DARWIN: A program for docking flexible molecules. *Proteins*. 2000;**41**(2):173-191

[22] Liu M, Wang S. MCDOCK: A Monte Carlo simulation approach to the molecular docking problem. *Journal of Computer-Aided Molecular Design*. 1999;**13**(5):435-451

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

[24] Oshiro CM, Kuntz ID, Dixon JS. Flexible ligand docking using a genetic algorithm. *Journal of Computer-Aided Molecular Design*. 1995;**9**(2):113-130

[25] Salomon-Ferrer R, Case DA, Walker RC. An overview of the Amber biomolecular simulation package. *Wiley Interdisciplinary Reviews: Computational Molecular Science*. 2013;**3**(2):198-210

[26] Brooks BR, Brooks CL 3rd, Mackerell AD Jr, Nilsson L, Petrella RJ, Roux B, et al. CHARMM: The biomolecular simulation program. *Journal of Computational Chemistry*. 2009;**30**(10):1545-1614

[27] Berendsen HJC, van der Spoel D, van Drunen R. GROMACS - a message-passing parallel molecular-dynamics implementation. *Computer Physics Communications*. 1995;**91**(1-3):14

[28] Drug Discovery Informatics Market Size, S. T. A. R. Drug Discovery Informatics Market Size, Share & Trends Analysis Report by Workflow (Discovery, Development), by Mode (In-House, Outsourced), by Services, by Region, Vendor Landscape, and Segment Forecasts 2018-2025. 2016. 978-1- 68038-749-0

[29] Berman HM, Bhat TN, Bourne PE, Feng Z, Gilliland G, Weissig H, et al. The protein data bank and the challenge of structural genomics. *Nature Structural Biology*. 2000;**7**:957-959

[30] Wang X, Song K, Li L, Chen L. Structure-based drug design strategies and challenges. *Current Topics in Medicinal Chemistry*. 2018;**18**(12):998-1006

[31] Markosian C, Di Costanzo L, Sekharan M, Shao C, Burley SK, Zardecki C. Analysis of impact metrics for the protein data Bank. *Scientific Data*. 2018;**5**:180212

[32] Rentzsch R, Renard BY. Docking small peptides remains a great challenge: An assessment using AutoDock Vina. *Briefings in Bioinformatics*. 2015;**16**(6):1045-1056

[33] Wojcikowski M, Ballester PJ, Siedlecki P. Performance of machine-learning scoring functions in structure-based virtual screening. *Scientific Reports*. 2017;**7**:46710

[34] Han JH, Batey S, Nickson AA, Teichmann SA, Clarke J. The folding and evolution of multidomain proteins. *Nature Reviews. Molecular Cell Biology*. 2007;**8**(4):319-330

[35] Mobley DL, Dill KA. Binding of small-molecule ligands to proteins: “what you see” is not always “what you get”. *Structure*. 2009;**17**(4):489-498