https://repository.uaeh.edu.mx/revistas/index.php./MJMR/issue/archive

Mexican Journal of Medical Research ICSa

Biannual Publication, Vol. 12, No. 24 (2024) 67-71

ISSN: 2007-5235

In Silico methods for the study of the interactions between drugs and their protein targets

Métodos *in Silico* para el estudio de las interacciones entre fármacos y sus blancos proteicos

Ricardo G. Martinez Daniel [a]

Abstract:

In Silico methods are a set of theoretical computer tools that analyze and correlate a series of physical, chemical and mathematical parameters to study the behavior of molecules with biological action called ligands, among which are drugs, biomolecules and their therapeutic targets The objective of this article is to show the most relevant ones, among which are; molecular docking, which relates the molecular dynamics between drug-protein structures, and the structure-activity relationship or SAR method and the quantitative method of structure-activity relationship or QSAR which correlates the physicochemical, electronic and steric parameters of drugs with their biological activity. With the mathematical results obtained through these methods, a series of predictions of important theorical-experimental relationships are generated for the development of drugs and the study of their behavior under different biological conditions.

Keywords:

Molecular docking, drug, protein, in Silico, relationship

Resumen:

Los métodos *in Silico* son un conjunto de herramientas informáticas teóricas que analizan y correlacionan una serie de parámetros físicos, químicos y matemáticos con el fin de estudiar el comportamiento de moléculas con acción biológica llamadas ligandos, entre las que se encuentran, fármacos, biomoléculas y sus blancos terapéuticos, estos estudios relacionan datos experimentales y datos teóricos de las interacciones ligando-proteína basados en descriptores moleculares. El objetivo del presente artículo es mostrar las más relevantes, entre las que se encuentran; el acoplamiento molecular, que relaciona la dinámica molecular entre estructuras fármaco proteína, el método de relación estructura actividad o SAR y el método cuantitativo de la relación estructura actividad o QSAR que correlacionan parámetros fisicoquímicos, electrónicos y estéricos de fármacos con su actividad biológica. Con los resultados matemáticos obtenidos mediante estos métodos se generan una serie de predicciones de inter relaciones teórico-experimentales importantes para el desarrollo de fármacos y el estudio de su comportamiento bajo distintas condiciones biológicas.

***Palabras Clave**:*

Acoplamiento molecular, fármaco, proteína, in Silico, relación

INTRODUCTION

In Silico methods provide a solution to the problem of excessive time and resources necessary for the study and development of new molecules and therapeutic targets based on experimental information, that sets the standard for structure-based molecule design (SBDD) for its acronym in English, in molecular dynamics, ligand-based molecule design (LBDD) and fragment-based molecule design (FBDD).[1,2] In this sense these methods include a variety of computational models based on the measurement of physicochemical, mathematical and biological parameters obtained from experimental models that study interactions at the molecular level between ligands (drugs) and proteins. The information resulting from this process is analyzed and interpreted by Machine Learning (ML) type algorithms that correlate parameters such as molecular weight, lipophilicity/hydrophilicity ratio (log P), number of hydrogens bonded donor-acceptor, polar surface and affinity to its target protein, resulting in an approximate accurate prediction of biological behavior in the process of reinforcing information from existing molecules or discovering new active molecules.[3] Among the most representative *in Silico* methods are those that

[a] Corresponding author, Consultorio Particular | Pachuca de Soto Hidalgo | México, https://orcid.org/0009-0008-6453-5422, Email: ricardoqfarma@gmail.com

Received: 01/03/2024, Accepted: 09/05/2024, Postprint: 13/05/2024, Published:

DOI: https://doi.org/10.29057/mjmr.v12i24.12393

study the interaction through a structural-based virtual screening (VS) high throughput screening (HTS), molecular docking and screens that analyzes the process of absorption, distribution, metabolism and elimination (ADME) of drugs in a biological system as in the case of SAR studies. This is way, to develop the methods one must start from the knowledge of the molecular structure of the ligands as well as the associated proteins previously studied and contained in libraries such as PubChem, ChEMBL and ZINC; nevertheless, the results are not limited to already existing information that it is possible to experiment with the design and structural nature of the ligands, their flexibility and their spatial dimension and rigidity, in this sense, it is essential to jointly prepare the target site of the ligand using X-ray crystallography available in databases such as PBD (www.rcsb.org). The digitization of the parameters corresponding to each of the molecular variables to be studied becomes the raw material for experimentation and the starting point for the different types of docking, validation and analysis of results.[4,5]

IN SILICO STUDIES

In Silico studies consist of a series of methods and methodologies based on chemistry, physics, pharmacology, mathematics and bioinformatics; that allow researchers to carry out biological assays computationally executed with mathematical algorithms established since the beginning of the 20th century in which correlations were established between the chemical structures of organic molecules and their behavior towards biomolecules of physiological and pharmacological importance. In these bioinformatics studies, the physical and chemical parameters of a molecule of interest are evaluated and how it behaves towards a target protein based on its three-dimensional geometry, type of bond, polarity, electronegativity and energy quantity.[6]

In Silico studies are used to generate a prediction approximate to reality in which researchers obtain an advantage by obtaining a simulated panorama of how a compound or compounds of interest would behave against a specific therapeutic target by analyzing the molecular docking in relation to their atomic, geometric and spatial structures. In this sense, these studies allow the researcher to reduce the resources necessary to develop test since the predictions provide information for the modelling of molecules with therapeutic potential; this is, if one wants study a group of drugs with similar characteristics approximations can be made regarding a minimal structure with an important activity as well as the substituents to choose to improve its activity.[7] *In Silico* studies provide a range of tools to relate a biological activity related to the chemical structure of molecules of interest in the case of SAR studies that relate these parameters to optimize the structure that presents the best activity and in this way only synthesize and study *in Vitro* the molecules with the most significant probability of being effective, however, these studies require a certain degree of prior experimental information that allows the information generated to be mutually supported and complemented between studies. Another type of study is the so-called Molecular Docking in which the interactions between ligands (drugs, small organic molecules, peptides, proteins, etc.) and proteins related to the structure of the ligand and the protein nature of the target to be studied are analyzed are carried out using specialized computer programs in which the different interactions are evaluated using algorithms that result in a complete analysis of the interactions between the ligand and the steric site of the protein with an action potential of either inhibition or induction.[8,9]

MOLECULAR DOCKING STUDIES

Studies in the molecular docking modality are based on the relationship between a ligand and its three-dimensional structure and its interaction with a general ligand protein, all at a molecular and strictly computational level in which there is a broad library of ligands and proteins targets, described and taken to a mathematical and computational language. The computer tools allow modeling ligands in 2D and 3D as well as three-dimensional modeling of a protein of interest in its quaternary structure in its amino acid sequence faithfully represented by X-ray crystallography (Figure 1).[9]

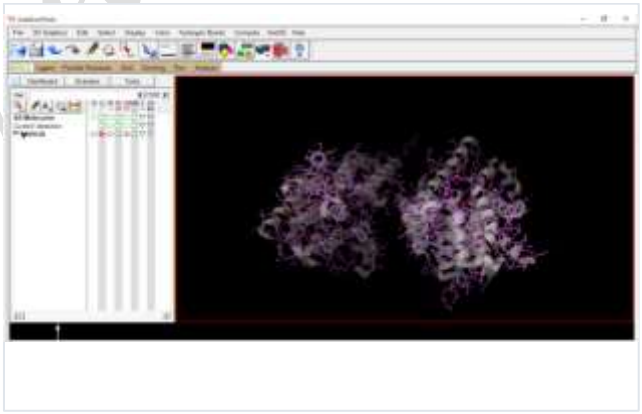

Figure 1. Free access computer program AutoDock showing an X-ray crystallography of the PDE-4B2 obtained from the PDB protein data bank.[7-9]

Directed Docking is studied when the researcher has a bibliographic base that supports the information on the exact steric site in which the protein interacts with a specific ligand, it is carried out to confirm the existing information on the mechanism and site of action of a ligand and a fraction of a protein target. However, molecular docking can also be performed without knowing thee specific site of interaction with a ligand through a blind docking study, in which a specific ligand is studied in a protein of interest in such a way that a mapping of the entire protein ant its probable binding sites with ligand is carried out, thus giving a prediction of how it interacts specifically with an amino acid residue to the analysis of the

functional groups that interact and the type of interaction link of each of them.[10]

The preparation of the ligand and the protein target is a crucial step in the molecular docking methodology because it defines the complexity of the study; starting with the ligand, its flexibility is stablished, referring to the mobility and position of atoms that amplify the modalities and positions of interaction with its protein target, ending with flexibility, solubility microenviroment of the medium and the protein target that defines the capacity of the hydrogens available to interact through hydrogen bonds and ruling out false positive interactions.[11] Molecular docking enables structure-based molecular design (SBDD) or ligand-based drug design (LBDD) ligands, the first consisting of structure-based virtual screening models (SBVS) that evaluates a structure and its interaction with an entire protein or a molecular dynamic in which the possible interactions between the structures of the ligand and the protein target are described. LBDD docking studies are evaluated in three modalities, the first, which refers to a virtual screening based on a ligand; the second, which consists of carrying out a SAR study; and the third, which consists of the generation of a pharmacophore, which is established as a pattern for the development of new molecules or ligands, able of interacting with a fraction of the protein (pharmacophore), serving as a template for the design and study of molecules with therapeutic potential.[12,13]

QSAR STUDIES

Quantitative Structure Activity Relationship Studies or QSAR are one of the *in Silico* methods that allows a relationship to be made between the structure of a ligand and the protein target through a series of physical and chemical factors that have the function of acting as biological descriptors; these factors are related to each other when compared with the existing experimental information and the theorical parameters established by the theorical QSAR method.[14,15]

QSAR studies are carried out by establishing molecular parameters to be evaluated and their relationship with the existing biological experimental information, the physicochemical data of the ligand running in a computer program that describes its behavior in different theorical scenarios and resulting in a prediction attached to reality supported by a SAR model based on the study objectives established by the researcher that ensure a correct extrapolation between empirical and theorical data, having a general scheme as show in figure 2.[16,18]

2D-QSAR

The study of 2D-QSAR represents a series of experimental and theorical methods from which information on the structure-activity relationship is obtained through lineal parameters that do not necessary represent a ligand-protein interactive activity in three dimensions. However, important parameters and results can be established to visualize the activity of molecules of therapeutic interest.[19,20]

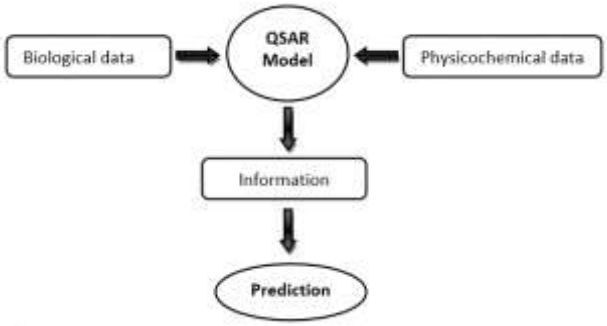

Figure 2. General scheme for the use of the QSAR method.[16-18]

PHYSICO-CHEMICAL PARAMETERS

Experimental information between molecules with biological potential allows real parameters to be digitized into theorical parameters to carry out a theorical-experimental contrast study between different biological parameters such as IC_{50} (binding affinity), minimum inhibitory concentration MIC, lethal dose 50 LD_{50}, volume of distribution Vd, bioavailability and physicochemical such as pKa (ionization constant), log P (lipophilicity/hydrophilicity ratio), log Kw (lipophilicity from HPLC measurement) and λ (hydrogen bond capability).[21]

LIPOPHILICITY/HYDROPHILICITY

One of the initial parameters that are evaluated in the QSAR study is the lipophilicity/hydrophilicity ratio of a ligand of interest to evaluate its pharmacokinetic behavior. Experimental or mathematical methods such as the Hansch and Fujita equation that represent a relationship between lipophilicity and hydrophilicity of each of the substituents or functional groups in a primary or seed molecule and its existing or probable modifications to be used subsequently. The fragmentary parameter π describes the lipophilicity/hydrophilicity value of the substituents of a molecule or drug; it is calculated by a difference of logarithms between an experimental value of a base structure studied in an organic solvent/water solution (π) and that same base with a substituent added by chemical synthesis (πx) resulting in a value of X that will be the lipophilicity/hydrophilicity value of a substituent that is interpreted as: Positive π values (+) represents a lipophilic substituent and negative π values (-) represent a hydrophilic substituent. In this sense, this information can be extrapolated to define this same parameter to a more complex molecule by applying a general equation:

$$\mathrm{Log}P_x = \mathrm{Log}P_H + \sum \pi x$$

The value to be searched $\mathrm{Log}P_x$ will be equal to the lipophilicity/hydrophilicity value of the $\mathrm{Log}P_H$ base plus the sum of each of the π values of the different substituents.[22,23]

ELECTRONIC PARAMETERS

This parameter studied by Hammett in 1964 analyzes the relationship between the acceptance and donation of electrons of a molecule by studying a differentiation between the ionization of unsubstituted benzoic acid based on the pKa values and the difference in ionization with some substituents. It is important to analyze this ionization due to the interactions of the different functional groups in a molecule and the amino acid residues of its therapeutic target. The sigma parameter σ is measured, which is the logarithmic difference of the electronic effect of a base molecule Ka_H and that of a functional group or substituent Ka_x in which positive values (+) describe electron-acepting substituents and negative values (-) describe electronic-donating substituents, thus interpreting the relationship between the substituents of a ligand and its electronic interaction. On the other hand, Swain and Lupton calculated a sigma value that is the sum of two electronic effects in a molecule with heterocycle aromatic characteristics, the inductive effect F and the resonance effect R that describes the movement of electrons within a molecule and its possible interaction with functional groups of amino acid residues of its protein target.[24]

STERIC PARAMETERS

This parameter shows the hydrolysis measured experimentally with ethyl acetate with a hydrogen substituent by Taft in aqueous and acidic media where Kx is measured, which represents the speed at which this reaction occurs starting from H as a substituent until its reaction with n X substituents. The speed then depends on the size of X, that is, the size of the atom or atoms of each unknown X. However, the size of a substituent is relative since it may be a flat substituent in 2D which does not faithfully represent an interaction that is close to reality, so it is necessary to include other parameters such a Verloop or sterimol in which the size and electronic shape of the substituent are represented on three axes (Figure 3), also supported by a molar refractivity parameter in which the molecular wight, refractive index and density are related.[25,26]

3D-QSAR

3D-QSAR considers three-dimensional parameters of the molecule to make a prediction based on the polar surface and a protein binding site. This method adds greater complexity to the evaluation of interactions between ligands and proteins through two families of methods, alignment-dependent and alignment-independent.[27]

ALIGNMENT-DEPENDENT

This method developed by Cramer et al., is a three-dimensional QSAR method called CoMFA (Comparative Molecular Field Analysis) in which the 3D structure and its quantitative relationship in structural and steric fields of the ligands are considered in superposition in each molecular field. A statistical analysis is used to analyze the biological response/activity correlation and the molecular energy interaction field. Activity predictions is evaluated by partial least squares.[28]

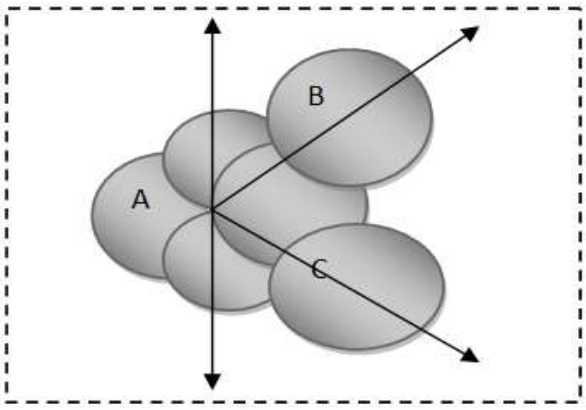

Figure 3. Graphic representation of the Verloop parameter in which the spatial conformations of the substituents and molecular bonds must be establshed.[25,26]

ALIGNMENT-INDEPENDENT

The CoMSIA (Comparative Molecular Similarity Indices Analysis) method developed by Klebe et al., is very similar to CoMFA but with the addition of electrostatic, steric, hydrophobic, hydrogen acceptor and hydrogen donor molecular fields generated using the Gaussian distance function (Figure 4)[29,30]

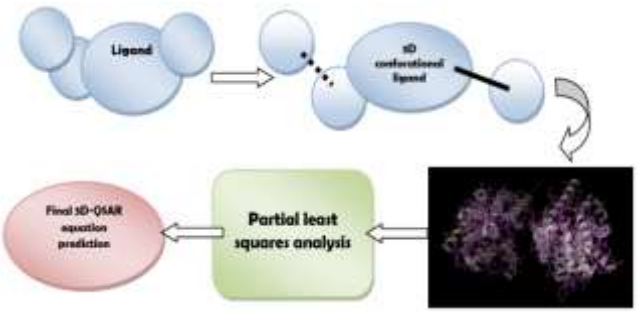

Figure 4. General scheme of the QSAR-3S method, presentation of the ligand and its 3D description conformations for docking to its target protein.[29,30]

CONCLUSION

In Silico methods represent a strategic advantage for researchers who wish to develop new molecules since they provide the necessary tools to study the behaviors during the interactions of drugs with their protein targets; although it is necessary to have a prior experimental basis, extrapolation, with theorical parameters they allow the researcher to obtain accurate predictions close to reality, which are very useful when designing and synthesizing drugs. However, the predictions should be taken as theoretical and complementary support to the existing experimental information.

REFERENCES

[1] Mihąşan M. What in silico molecular docking can do for the banch-working biologist. J. Biosci. 2012;37(6):1089-95.

[2] Kontoyianni M. Docking and virtual screening in drug discovery. Methods Mol. Biol. 2017;1647:255-66.

[3] Choi J, Lee J. V-Dock fast generation of novel drug-like molecules using machine-learning-based docking score and molecular optimization. Int. J. Mol. Sci. 2021;22(21):11635.

[4] Batool M, Ahmad B, Choi S. A structure-based drug discovery paradigm. Int. J. Mol. Sci. 2019;20(11):2783.

[5] Wang T, Wu MB, Chen ZJ, Chen H, Lin JP, Yang LR. Fragment-based drug discovery and molecular docking in drug design. Curr. Pharm. Biotechnol. 2015;16(1):11-25.

[6] Yuriev E, Holien J, Ramsland PA. Improvements, trends and new ideas in molecular docking 2012-2013 in review. J. Mol. Recognit. 2015;28(10):581-04.

[7] Yan Y, Zhang D, Zhou P, Li B, Huang SY. HDOCK: a web server for protein-protein and protein-DNA/RNA docking based on hybrid strategy. Nucleic Acid Res. 2017;45(1):365-73.

[8] Schweiker SS, Levonis SM. Navigating the intricacies of molecular docking. Future Med. Chem. 2020;12(6):469-71.

[9] Dallakyan S, Olson AJ. Small-molecule library screening by docking with PyRx. Methods Mol. Biol. 2015;1263:243-50.

[10] Menin S, Pavan M, Salmaso V, Sturlese M, Moro S. Thermal titration molecular dynamics (TTMD): not your usual post-docking refinement. Int. J. Mol. Sci. 2023;24(4):3596.

[11] Tingle BI, Irwin JJ. Large-scale docking in the cloud J. Chem. Inf. Model. 2023;63(9):2735-41.

[12] Potemkin V, Grishina M. Principles for 3D/4D QSAR classification of drugs. Drug Discov. Today 2008;13(21-22): 952-59.

[13] Bahia MS, Kaspi O, Touitou M, Binayev I, Dhail S, Spiegel J, et al. A comparison between 2D and 3D descriptors in QSAR modeling based on bio-active conformations. Mol. Inform. 2023;42(4):2200186.

[14] Nikonenko A, Zankov D, Baskin I, Madzhidov T, Polishchuk P. Multiple conformer descriptors for QSAR modelling. Mol. Inform. 2021;40(11):2060030.

[15] Clark RD. Prospective ligand and target-based 3D QSAR: state of the Art 2008. Curr. Top. Med. Chem. 2009;9(9):791-810.

[16] Wilkes JG, Stoyanova-Slovova IB, Buzatu DA. Alignment-independent technique for 3D QSAR analysis. J. Comput. Aided Mol. Des. 2016;30(4): 331-45.

[17] Cramer RD, Wendt B. Pushing the boundaries of 3D QSAR. J. Comput. Aided Mol. Des. 2007;21(1-3):23-32.

[18] Wang YL, Wang F, Shi XX, Jia CY, Wu FX, Hao GF, et al. Cloud 3D-QSAR: a web tool for the development of quantitative structure-activity relationship models in drug discovery. Brief. Bioinform. 2021;22(4):276.

[19] Myint KZ, Xie XQ. Recent advances in fragment-based QSAR and multi-dimensional QSAR methods. Int. J. Mol. Sci. 2010;11(10):3846-66.

[20] Cramer RD. Rethinking 3D QSAR. J. Comput. Aided Mol. Des. 2011;25(3): 197-201.

[21] Bak A. Two decades of 4D-QSAR: A dying art or staging a comeback? Int. J. Mol. Sci. 2021;22(10):5212.

[22] Doweyko AM. 3D-QSAR illusions. J. Comput. Aided Mol. Des. 2004;18(7-9):587-96.

[23] Ginex T, Vazquez J, Gilbert E, Herrero E, Luque FJ. Lipophilicity in drug design: an overview of lipophilicity descriptors in 3D-QSAR studies. Future Med. Chem. 2019:11(10): 1177-93.

[24] Polanski J. Receptor dependent multidimensional QSAR for modeling drug-receptor interactions. Curr. Med. Chem. 2009;16(25):3243-57.

[25] Verma J, Khedkar VM, Coutinho EC. 3D-QSAR in drug design-a review. Curr. Top. Med. Chem. 2010;10(1):95-115.

[26] Lu S, Liu HC, Chen YD, Yuan HL, Sun SL, Gao YP, et al. Combined pharmacophore modeling, docking, and 3D-QSAR studies of PLK1 inhibitors. Int. J. Mol. Sci. 2011;12(12):8713-39.

[27] Damale MG, Harke SN, Kalam Khan FA, Shinde DB, Sangshetti JN. Recent andvances in multidimensional QSAR (4D-6D): a critical review. Mini Rev. Med. Chem. 2014;14(1):35-55.

[28] Bordás B, Komíves T, Lopata A. Ligand-based computer-aided pesticide design. A review of applications of the CoMFA and CoMSIA methodologies Pest. Manag. Sci. 2003;59(4):393-400.

[29] Caballero J. 3D-QSAR (CoMFA and CoMSIA) and pharmacophore (GALAHAD) studies on the differential inhibition of aldose reductase by flavonoid compounds. J. Mol. Graph. Model. 2010;29(3):363-71.

[30] Xu M, Zhang A, Han S, Wang L. Studies of 3D-quantitative structure-activity relationships on a set of nitroaromatic compounds: CoMFA, advanced CoMFA and CoMSIA. Chemosphere 2002;48(7):707-15.