Chapter 1

Introductory Chapter: Molecular Docking - Overview, Background, Application and What the Future Holds

Dimitrios Vlachakis

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.78266

1. Introduction

Molecular docking is on the frontline of computational biology and drug discovery. The explosion of structural and chemical information in recent years has rendered the use of efficient algorithms and large supercomputer facilities of uttermost importance in the drug discovery process. Medicinal chemists can now screen *in silico* hundreds of thousands of compounds on a repertoire of receptor molecules and putative pharmacological targets. It goes without saying that molecular docking comes in many shapes and sizes, thus allowing the researcher to balance out speed and exhaustiveness of calculation. Molecular docking can be performed online of freeware servers using just a web browser or it can be fully parameterized on a virtual machine on a cloud supercomputer for high resolution calculation. The main factor that changes here is the grid resolution and the rigidity and flexibility of both the ligand and the receptor.

2. Molecular docking in a nutshell

Let us start by setting the basis on molecular properties that are required to comprehend the molecular docking chapters that follow in this book. The geometry and the overall structure of a molecule are described by its bond distances, dihedral angles and bond angle [1]. This unique set of angles and distances create a set of coordinates that define the positioning of each atom in that molecular structure in three-dimensional (3D) space. The energy condition of this molecule can also be assessed and evaluated. The energy of a molecule includes all forms of energies, such as kinetic motion (described by vibration, rotation and translation) and forms of the potential energy of the molecule [2]. The potential energy of a molecule

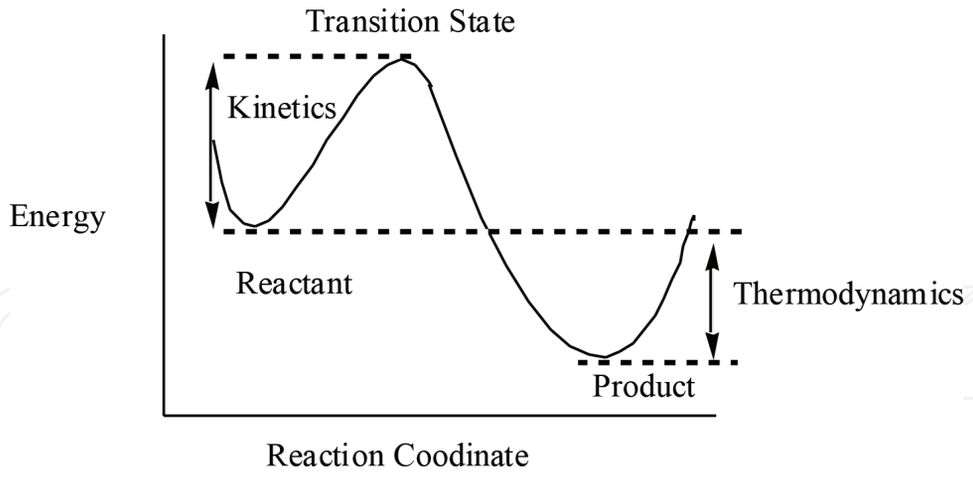


Figure 1. Energy changes during the course of a chemical reaction.

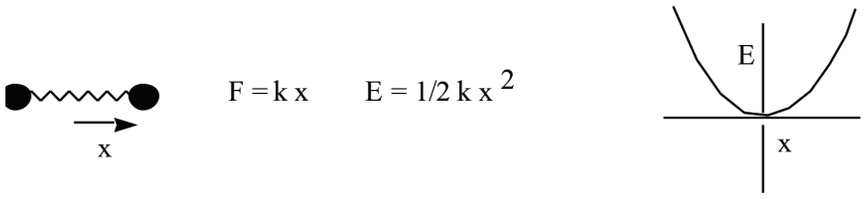
can be defined by the analysis of the electrostatic interaction between charges, the magnetic interactions between spinning charges and finally the potential energy of the bonds of the molecule. The total energy is indicative of the reactivity and stability of that a molecule or a system. **Figure 1** is depicts a reaction coordinate diagram that indicates the energy changes during the course of a chemical reaction [3].

Here the products are in the lowest or global minimum, the transition state is at energy maximum and the reactants are at an energy minimum. The dotted lines in the above diagram are indicative of the reactivity of the system (its kinetics) and the thermodynamic stability of the system. Through molecular modelling it is possible to quantify the above characteristics of the system and, for example, predict its reactivity. There are two fields in molecular modelling that attempt to do this: molecular mechanics and quantum mechanics [4].

The docking algorithm is basically split into two main parts: the searching algorithm and the scoring algorithm [5]. The searching algorithm will explore all conformations of the ligand within the space available [6]. Practically, it is impossible to perform all these calculations for every compound so most of the rotational and translational states of each compound will be explored within a given threshold of identical conformations. Each compound is not a rigid body but is a dynamic structure that exists in an ensemble of different conformations. The user can define how fine the docking algorithm will be by altering the various parameters of the task. Very fine calculations are much more accurate, but also much more time consuming. The most popular docking algorithm approaches can involve a coarse grained molecular dynamics simulation or a linear combination of many structures or a genetic algorithm that generates new conformations as it moves along.

The second feature of the docking algorithm is its scoring function [7]. The scoring function must be able to accurately evaluate each different conformation using certain forcefields and rules from physics and return a value that will describe the energy of the system at the given conformation. Low energies indicate better, more stable interactions.

http://dx.doi.org/10.5772/intechopen.78266



Molecular mechanics are based on the ball and spring representation of molecular systems. Here, the atoms are considered to be little balls, with varying properties according to the element, and the bonds are considered to be the springs that make the two interconnecting balls interact with each other. The ball and spring model is described by Hook's law, which evaluates and quantifies the energy of the stretching of the spring [8].

The force constant is the constant k. The energy that is contained in the spring and the restoring force of the spring are proportional to the force constant. The force constant will determine the strength of the bond that the spring represents [9]. The vibrational frequency of the spring is described as:

$$n = \frac{1}{2\pi}\sqrt{\frac{k}{\mu}} \tag{1}$$

The vibrational frequency (n) has been estimated to be proportional to the square root of the force constant (k) and inversely proportional to the reduced mass of the atoms that participate in a bond [10].

All of the above can be combined and through potential energy functions of various structural features, such as bond lengths, bond angles and non-bonded interactions, can describe a forcefield (**Figure 2**) [11]. There are many different ways to set a forcefield depending on the needs of the system under investigation. Usually the factors affecting the energy of a molecular system (bonds, angles, dihedrals, non-bonded, etc.), are evaluated separately and they will contribute to the value of the total energy of the system [12]. The most popular forcefields are the MM2, which is suitable for small molecules, hydrocarbons and some simple heteroatom

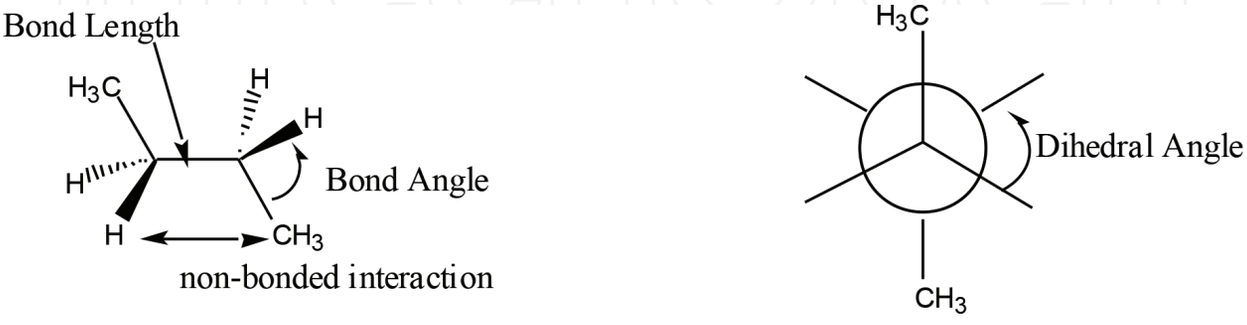


Figure 2. Total energy is affected by bond distances, bond angles, dihedral angles and finally non-bonded interactions.

functional groups, AMBER or CHARMM, which are parameterised to be used for peptides, nucleic acids and generic macromodels [13].

Overall through molecular mechanics the total energy of a molecule is described as a sum of all the contributions that may arise from loss of equilibrium in bond distances, also known as stretching contribution, bond angles, known as bending contribution, dihedral angles, the torsion contribution and finally non-bonded interaction contributions [14].

$$E^{total} = \sum_{i}^{bonds} E_i^{stretch} + \sum_{i}^{bond\ angles} E_i^{bend} + \sum_{i}^{dihedral\ angles} E_i^{torsion} + \sum_{i}\sum_{j}^{non\text{-}bonded\ atoms} E_{ij}^{non\text{-}bonded} \quad (2)$$

The energy that is stored in chemical bonds of a molecule can describe the stretch, bend, and torsion energy whereas it is the steric attraction or repulsion that represents the non-bonded energy [15]. The latter is broken down to two different categories: the van der Waals (VDW) and electrostatic interactions [16].

A very steep energy barrier is generated at the van der Waals radius of each atom. Moreover a very shallow energy well is produced at larger separations (**Figure 3**). The inherent steric size of atoms and elements is dictated by their VDW radii. The same metric is used to describe weak attractive forces between atoms in close proximity [17]. A trivial example of the weak van der Waals attractive forces is the condensation of a gas into liquids. Furthermore it is the van der Waals radii of each element that is used for its visualisation purposes in space filling models of the molecule they participate. Steric repulsion takes place only in the case where two atoms come closer than the sum distance of their VDW radii [18].

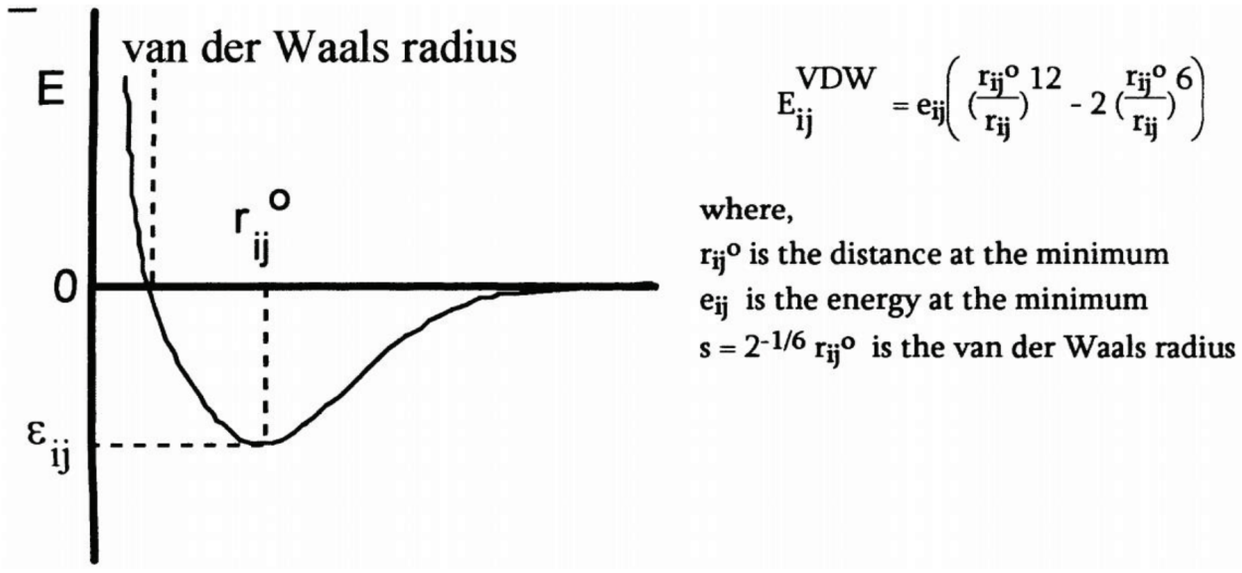


Figure 3. The van der Waals interactions plot and formula.

http://dx.doi.org/10.5772/intechopen.78266

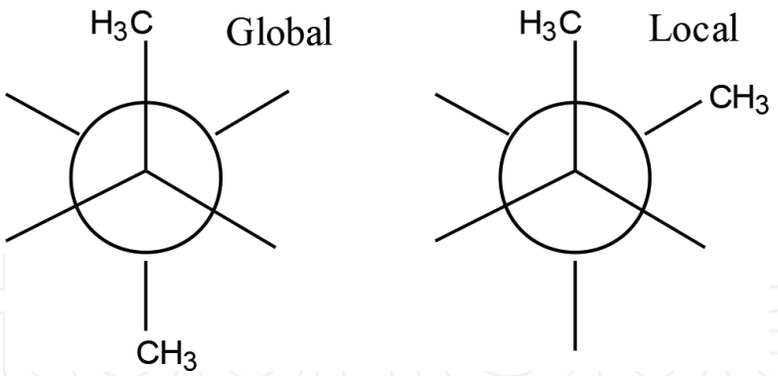


Figure 4. Two different conformations of butane.

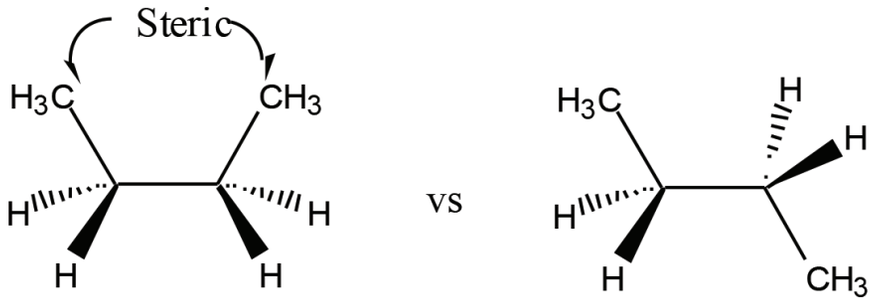


Figure 5. Steric hindrance of a small organic compound.

As soon as the set of the internal coordinates of a molecular system has been determined, computer algorithms can be used to help find those coordinates which will account for the lowest energy of the system [19]. All bond angles, lengths, dihedral angles and the relative energy between various different conformations of a given system will be evaluated in order to determine the minimum energy conformation [20]. It is crucial to understand that reducing the strain energy of a given molecular system does not mean that the system will reach energy minimum (also known as global minimum). An example is the following figure (**Figure 4**) with two different conformations of butane.

An energy minimisation algorithm will allow the rotation of groups, when their bonding allows. The rotation of the groups will give the molecule the opportunity to explore different conformations that will account for different energy values, thus allowing the compound to move towards its global minimum conformation [21].

Molecular modelling is very useful for investigating, comparing, analysing and visualising chemical structures and for giving qualitative and quantitative information about biological systems [22].

Figure 5 shows a characteristic example of steric hindrance. Two dimensional models like this only contain qualitative information. Quantitative information can arise through molecular mechanics and in conjunction with a computer, where the physical properties of the molecules can be evaluated and analysed based on a set of predefined criteria concerning various chemical

properties (such as bonding, charges, steric hindrance) [23]. Molecular Modelling can be used to study the geometry, the energy and the chemical properties *in silico* so efficiently that nowadays it is possible to predict the outcome of chemical reactions, design reactions, determine the unknown three-dimensional structures of proteins, screen and design new and effective drugs [23].

All in all, the future is bright for molecular docking. New technologies are being developed and employed in the race against drug discovery and lethal diseases. Data mining, machine or deep learning, hyper-computers and cloud computers are just few of the emerging technologies in modern molecular docking.

Author details

Dimitrios Vlachakis

Address all correspondence to: dimvl@aua.gr

Laboratory of Genetics, Department of Biotechnology, Agricultural University of Athens, Athens, Greece

References

[1] Branden C, Tooze J. Introduction to Protein Structure. 2nd ed. New York: Garland Publishing Inc; 1999

[2] Bratley P, Fox BL, Schrage LE. A Guide to Simulation. New York: Springer-Verlag; 1987

[3] BrooksIII CL, Karplus M, Pettitt BM. A Theoretical Perspective of Dynamics, Structure, and Thermodynamics. New York: Wiley Interscience; 1988

[4] Burkert U, Allinger NL. Molecular Mechanics. Washington D.C.: American Chemical Society; 1980

[5] Ryckaert JP, Ciccotti G, Berendsen HJC. Numerical integration of the Cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes. Journal of Computational Physics. 1977;**23**:327-341

[6] Hess B, Bekker H, Berendsen HJC, Fraaije JGEM. LINCS: A linear constraint solver for molecular simulations. Journal of Computational Chemistry. 1997;**18**:1463-1472

[7] Ferguson DM, Raber DJ. A new approach to probing conformational space with molecular mechanics: Random incremental pulse search. Journal of the American Chemical Society. 1989;**111**:4371-4378

[8] Cantor CR, Schimmel PR. Biophysical Chemistry. Vol. 1-3. San Francisco: W.H. Freeman and Company; 1980

[9] Cohen NR, editor. Guidebook on Molecular Modeling in Drug Design. San Diego: Academic Press; 1996

[10] Deuflhard P, Hermans J, Leimkuhler B, Mark AE, Reich S, Skeel RD, editors. Computational Molecular Dynamics: Challenges, Methods, Ideas – Proceedings of the 2nd International Symposium on Algorithms for Macromolecular Modelling, Berlin, May 21-24, 1997. Vol. 4. Lecture Notes in Computational Science and Engineering. Berlin, Heidelberg: Springer-Verlag; 1999

[11] Creighton TE, editor. Protein Folding. New York: W.H. Freeman & Company; 1992

[12] Eisenberg D, Crothers D. Physical Chemistry with Applications to the Life Science. Menlo Park, California: Benjamin Cummings; 1979

[13] Fersht A. Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding. New York: W. H. Freeman and Company; 1999

[14] Frenkel D, Smit B. Understanding Molecular Simulations. From Algorithms to Applications. San Diego, California: Academic Press; 1996

[15] Gierasch LM, King J, editors. Protein Folding, Deciphering the Second Half of the Genetic Code. Washington D.C.: AAAS; 1990

[16] Gould H, Tobochnik J. An Introduction to Computer Simulation Methods: Applications to Physical Systems. Part 1 and 2. Reading, MA: Addison-Wesley; 1988

[17] Grosberg AY, Khokhlov AR. Giant Molecules. Here, There, and Everywhere…. San Diego, California: Academic Press; 1997

[18] Haile JM. Molecular Dynamics Simulations: Elementary Methods. New York: Wiley; 1992

[19] Kalos M, Whitlock PA. Monte Carlo Methods. New York: John Wiley & Sons; 1986

[20] Leach AR. Molecular Modelling. Principles and Applications. Essex, England: Addison Wesley Longman; 1996

[21] Lipkowitz KB, Boyd DB, editors. Reviews in Computational Chemistry. New York: VCH Publishers; 1990

[22] Allen MP, Tildesley DJ. Computer Simulation of Liquids. New York: Oxford University Press; 1987

[23] Bates AD, Maxwell A. DNA Topology. In Focus Series. New York: Oxford University Press; 1993

IntechOpen