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

6,900
Open access books available

186,000
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

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

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

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

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

Protein-Protein and Protein-Ligand Docking

Alejandra Hernández-Santoyo,
Aldo Yair Tenorio-Barajas, Victor Altuzar,
Héctor Vivanco-Cid and Claudia Mendoza-Barrera

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/56376

1. Introduction

Molecular interactions including protein-protein, enzyme-substrate, protein-nucleic acid, drug-protein, and drug-nucleic acid play important roles in many essential biological processes, such as signal transduction, transport, cell regulation, gene expression control, enzyme inhibition, antibody–antigen recognition, and even the assembly of multi-domain proteins. These interactions very often lead to the formation of stable protein–protein or protein-ligand complexes that are essential to perform their biological functions. The tertiary structure of proteins is necessary to understand the binding mode and affinity between interacting molecules. However, it is often difficult and expensive to obtain complex structures by experimental methods, such as X-ray crystallography or NMR. Thus, docking computation is considered an important approach for understanding the protein-protein or protein-ligand interactions [1-3]. As the number of three-dimensional protein structures determined by experimental techniques grows —structure databases such as Protein Data.

Bank (PDB) and Worldwide Protein Data Bank (wwPDB) have over 88000 protein structures, many of which play vital roles in critical metabolic pathways that may be regarded as potential therapeutic targets — and specific databases containing structures of binary complexes become available, together with information about their binding affinities, such as in PDBBIND [4], PLD [5], AffinDB [6] and BindDB [7], molecular docking procedures improve, getting more importance than ever [8].

Molecular docking is a widely used computer simulation procedure to predict the conformation of a receptor-ligand complex, where the receptor is usually a protein or a nucleic acid molecule and the ligand is either a small molecule or another protein (Figure 1).
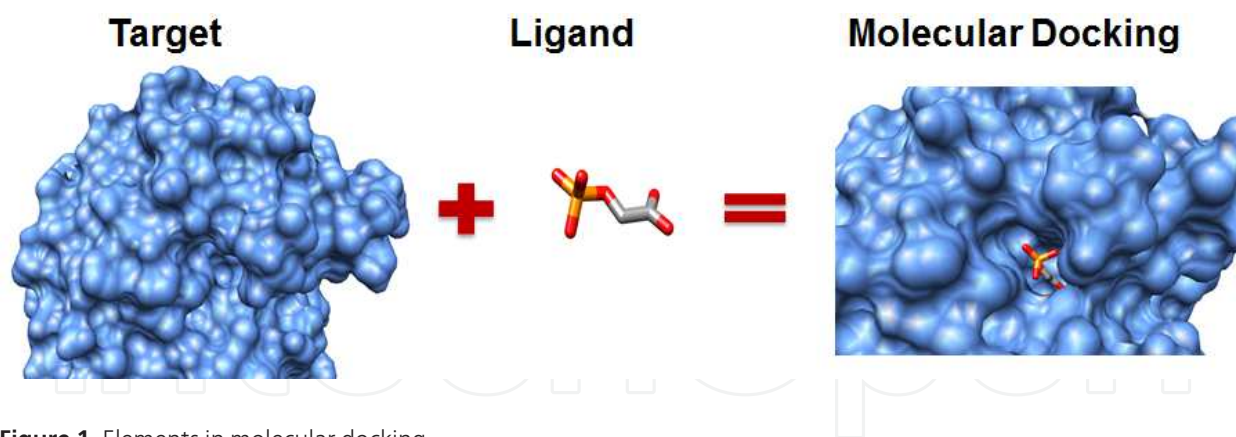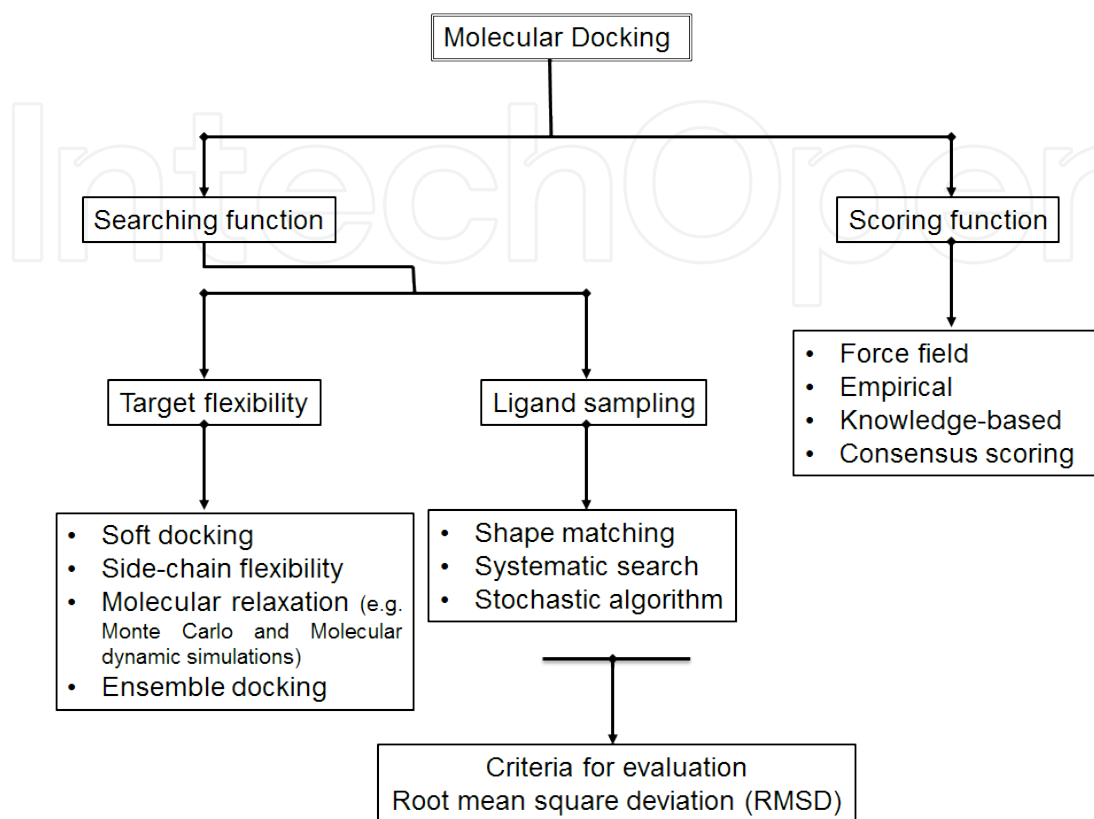
Figure 1. Elements in molecular docking.

The accurate prediction of the binding modes between the ligand and protein is of fundamental importance in modern structure-based drug design. The most important application of docking software is the virtual screening, in which the most interesting and promising molecules are selected from an existing database for further research. This places demands on the used computational method: it must be fast and reliable. Another application is the research of molecular complexes.

Since the pioneering work of Kuntz *et al*. [9] during the early 1980s, significant progress has been made in docking research to improve the computational speed and accuracy. Over the last years several important steps beyond this point have been given. Handling efficiently the flexibility of the protein receptor is currently considered one of the major challenges in the field of docking. The binding-site location and binding orientation can be greatly influenced by protein flexibility. In fact, X-ray structure determination of protein–ligand complexes frequently reveals ligands with a buried surface area in the range of 70–100%, which can only be achieved as a consequence of protein flexibility [3]. There are many interesting docking suites and algorithms that have shown significant progress in predicting near-native binding poses by making use of biophysical and biochemical information combination with bioinformatics.

2. Theory

Modeling the interaction of two molecules is a complex problem. Many forces are involved in the intermolecular association, including hydrophobic, van der Waals, or stacking interactions between aromatic amino acids, hydrogen bonding, and electrostatic forces. Modeling the intermolecular interactions in a ligand-protein complex is difficult since there are many degrees of freedom as well as insufficient knowledge of the effect of solvent on the binding association. The process of docking a ligand to a binding site tries to mimic the natural course of interaction of the ligand and its receptor via the lowest energy pathway [3]. There are simple methods for docking rigid ligands with rigid receptors and flexible ligands with rigid receptors, but general methods of docking considering conformationally flexible ligands and

receptors are problematic. Docking protocols can be described as a combination of a search algorithm, and the scoring functions (Figure 2).

```
                        ┌─────────────────────┐
                        │  Molecular Docking  │
                        └─────────────────────┘
              ┌──────────────────┴───────────────────────────┐
     ┌─────────────────────┐                        ┌──────────────────┐
     │  Searching function │                        │ Scoring function │
     └─────────────────────┘                        └──────────────────┘
        ┌──────────┴──────────┐                              │
                                                   ┌──────────────────────┐
 ┌───────────────────┐  ┌──────────────────┐       │ •  Force field       │
 │  Target flexibility│  │  Ligand sampling │       │ •  Empirical         │
 └───────────────────┘  └──────────────────┘       │ •  Knowledge-based   │
                                                    │ •  Consensus scoring │
                                                    └──────────────────────┘
```

Target flexibility	Ligand sampling
• Soft docking • Side-chain flexibility • Molecular relaxation (e.g. Monte Carlo and Molecular dynamic simulations) • Ensemble docking	• Shape matching • Systematic search • Stochastic algorithm

```
                        ┌─────────────────────────────────────┐
                        │       Criteria for evaluation        │
                        │ Root mean square deviation (RMSD)    │
                        └─────────────────────────────────────┘
```

Figure 2. Methods used for protein-ligand docking.

The search algorithm should create an optimum number of configurations that include the experimentally determined binding modes. Although a rigorous searching algorithm would go through all possible binding modes between the two molecules, this search would be impractical due to the size of the search space and amount of time it might take to complete. As a consequence, only a small amount of the total conformational space can be sampled, so a balance must be reached between the computational expense and the amount of the search space examined. Some common searching algorithms include molecular dynamics, Monte Carlo methods, genetic algorithms, fragment-based, point complementary and distance geometry methods, Tabu, and systematic searches. On the other hand, scoring function consists of a number of mathematical methods used to predict the strength of the non-covalent interaction called the binding affinity. In all the computational methodologies, one important problem is the development of an energy scoring function that can rapidly and accurately describe the interaction between the protein and ligand. Several reviews on scoring are available in the literature [10-12].

There are three important applications of scoring functions in molecular docking. First is to determine the binding mode and site of a ligand on a protein. The second application is to

predict the absolute binding affinity between protein and ligand. This is particularly important in lead optimization. The third application, and perhaps the most important one in structure-based drug design, is to identify the potential drug hits/leads for a given protein target by searching a large ligand database. Over the course of the last years, different scoring functions have been developed that exhibit different accuracies and computational efficiencies. Some of these commonly-used functions are: force-field, empirical, knowledge-based and consensus scoring.

The protein-ligand docking procedure can be typically divided into two parts: rigid body docking and flexible docking [9].

1. *Rigid Docking.* This approximation treats both the ligand and the receptor as rigid and explores only six degrees of translational and rotational freedom, hence excluding any kind of flexibility. Most of the docking suites employ rigid body docking procedure as a first step.

2. *Flexible Docking.* A more common approach is to model the ligand flexibility while assuming having a rigid protein receptor, considering thereby only the conformational space of the ligand. Ideally, however, protein flexibility should also be taken into account, and some approaches in this regard have been developed. There are three general categories of algorithms to treat ligand flexibility: systematic methods, random or stochastic methods, and simulation methods [3]. Due to the large size of proteins and their multiple degrees of freedom, their flexibility may be the most challenging issue in molecular docking. The methods to address the flexibility of proteins can be grouped into: soft docking, side-chain flexibility, molecular relaxation and protein ensemble docking. They were described by Huang *et al* [1].

3. Experimental docking procedures

There are a number of excellent reviews of molecular docking methods and a large number of publications comparing the performance of a variety of molecular docking tools [1-3], [13]. Following, we will describe the four-step procedure adopted in this study to perform the molecular docking.

3.1. Target selection

Ideally, the target structure should be determined experimentally by either X-ray crystallography or nuclear magnetic resonance, which can be downloaded from PDB; however, docking has been performed successfully in comparison to homology models or threading. The model should have good quality. It can be tested using validation software such as Molprobity [14]. After selecting the model, it must be prepared by removing the water molecules from the cavity, stabilizing charges, filling the missing residues, and generating the side chains, all according to the available parameters. The receptor should be at this point biologically active and in the stable state.

3.2. Ligand selection and preparation

The type of ligands chosen for docking will depend on the goal. It can be obtained from various databases, e.g. ZINC or/and PubChem, or it can be sketched by means of Chemsketch tool [8]. Often it is necessary to apply filters to reduce the number of molecules to be docked. Examples include the net charge, molecular weight, polar surface area, solubility, commercial availability, similarity thresholds, pharmacophores, synthetic accessibility, and absorption, distribution, metabolism, excretion, and toxicology properties. Many times the researchers design their own molecules such as those generated by us in the example that will be described in this work in the section 5.

3.3. Docking

This is the last step, where the ligand is docked onto the receptor and the interactions are checked. The scoring function generates a score depending on the best selected ligand.

3.4. Evaluating docking results

The success of docking algorithms in predicting a ligand binding pose is normally measured in terms of the root-mean-square deviation (RMSD) between the experimentally-observed heavy-atom positions of the ligands and the one(s) predicted by the algorithm. The flexibility of the system is a major challenge in the search for the correct pose. The number of degrees of freedom included in the conformational search is a central aspect that determines the searching efficiency [3]. A good performance is usually considered when the RMSD is less than 2Å.

3.5. Docking software description

There are many algorithms available to assess and rationalize ligand-protein or protein-protein interactions, and their number is constantly increasing. Speed and accuracy are key features for obtaining successful results in docking approaches. Several algorithms share common methodologies with novel extensions focused on obtaining a fast method with accuracy as high as possible. The most common docking programs include AutoDock [15], DOCK [9], FlexX [16], GOLD [17], ICM [18], ADAM [19], DARWIN [20], DIVALI [21], and DockVision [22].

4. Application of molecular docking to a particular case — Biopolymers docked to dengue virus E protein

In the last decades, the incidence of Dengue disease has dramatically increased around the world. About 2.5 billion persons (two fifth of the world population) are exposed to the risk of contracting the disease. Every year, dengue virus (DENV) infects more than 50 million people, with approximately 22 000 fatal cases [23]. The disease is endemic in more than 100 countries of Africa, America, Oriental Mediterranean, Southeast Asia, and the Western Pacific Ocean with the last two regions being the most affected by the disease. Before 1970 only nine countries

suffered from the Hemorrhagic Dengue (HD) epidemics, number that in 1995 was multiplied for more than four. There are four antigenically distinct, but closely related, serotypes of dengue virus (DENV), which is a Flavivirus member of the family Flaviviridae [24]. Each serotype has genotypes, which are virulent at several levels; nevertheless, the factors of virulence are not totally established [25]. A better understanding of the mechanisms and the molecules involved in the key steps of the DENV transmission cycle may lead to the identification of new anti-dengue targets [26]. In fact, the presence of two or more serotypes in the same geographical region implies a growing risk to population of contracting Hemorrhagic Dengue or Dengue Shock Syndrome (SSD) due to a phenomenon known as the Antibody – Dependent Enhancement (ADE). As a result, the diagnosis and treatment of dengue disease has become a world-wide global problem to deal with. The mature DENV virion contains three structural proteins: capsid protein (C), membrane protein (M), and envelope protein (E). In particular, the DENV E glycoprotein (51-60 kDa~ 495 aa), found on the viral surface, is important in the initial attachment of the viral particle to the host cell, as it contains two N-linked glycosylation sites at Asn-67 and Asn-153. While the glycosylation site at position 153 is conserved in most flaviviruses, the site at position 67 is thought to be unique for dengue virus. N-linked oligosaccharide side chains on flavivirus E proteins have been associated with viral morphogenesis, infectivity, and tropism [27, 28]. In addition, E protein is closely associated with the lipid envelope containing a cellular receptor-binding site (s) and a fusion peptide [29]. It can be found in a form of a homodimer on the surface of the mature virion, and inside the cell, it creates a prM-E heterodimer together with the prM protein. E protein is the principal component of the virion surface, containing the antigenic determinants (epitopes) responsible for the neutralization of the virus and the hemagglutination of erythrocytes, inducing thereby an immunological response in the infected host [29]. On native virions, the elongated three-domain E molecule is positioned tangentially to the virus envelope in a head-to-tail homodimeric conformation. Upon penetration of the virion into the target cell endosome, E dimers are converted into stable target-cell membrane-inserted homotrimers that reorient themselves vertically to promote virus-cell fusion at low pH [30]. Furthermore, there is a great deal of evidence that E protein contains the majority of molecular markers for pathogenicity. Comparing the nucleotide sequence of the E protein gene in different flaviviruses has demonstrated a perfect conservation of 12 cysteine residues, which form six disulfide bridges. The structural model for the E protein was refined by Mandl and co-workers [31], who correlated the structural properties of different epitopes with disulfide bonds [32].

4.1. Biopolymers as potential adjuvants carriers

The aim of this work is to study the docking of monomers of polyvinylpyrrolidone (PVP), chitosan (CS), and chitosan-tripropylphosphate-chitosan (CS/TPP/CS) with E protein of dengue virus in order to use them as potential adjuvant carriers. Given their structure, these polymers have specific molecular anchor sites that are expected to be exploited to induce antigenic specificity to the conserved regions of dengue virus. Several authors report that the E protein produces immunity and confers protection against infection in mice with low levels of neutralizing antibodies [33-35]. Because of the dual role of its receptors as well as the cell

entry through membrane fusion, the E protein, apart from being the most exposed protein, is the main target against which the neutralizing antibodies are produced to inhibit its functions.

At present, the biggest challenge in developing an efficient dengue vaccine is to achieve a life-long protective immune response to all 4 serotypes (DEN1-4) simultaneously. Although several vaccines are currently being developed, so far only a chimeric dengue vaccine for live attenuated yellow fever (YF) has reached stage 3 in clinical trials. The candidate vaccines can be divided into the following types: (a) live attenuated, (b) DEN-DEN and DEN-YF live chimeric virus, (c) inactivated whole virus, (d) live recombinant, (e) DNA, and (f) subunit vaccines [36].

Chitosan is a polycationic polymer comprising of D-glucosamine and N-acetyl-D-glucosamine linked by β(1,4)-glycosides' bonds. It is produced by deacetylation of chitin, which is extracted from the shells of crabs and shrimp. It is a linear, hydrophilic, positively charged, water soluble biopolymer, can form thin films, hydrogels, porous scaffolds, fibers, and micro and nanoparticles in mild acidity conditions. As a polycationic polymer, it has a high affinity to associate macromolecules such as insulin, pDNA, siRNA, heparin, among others, with antigenic molecules, protecting them in turn from hydrolytic and enzymatic degradation [37].

Polyvinylpyrrolidone (N-vinyl-2-pyrrolidone, PVP) has chemical, physical and physiological properties which have been exploited in various industries, including but not limited to medical, pharmaceutical, cosmetic, food, and textile, due to its biological compatibility, low toxicity, tackiness, resistance to thermal degradation in solutions as well as inert behavior in salt and acidic solutions [38, 39]. It is a water soluble homopolymer with a wide range of molecular weights (2.5 to 1.200 kDa), molecules between 12 and 1350 monomers, and end-to-end distances ranging from 2.3 to 93 nanometers. It is physically and chemically stable; it tolerates heating and air atmospheres for up to 16 hours at 100 °C, as well as the change of appearance for 2 months at 24 ° C and 15% HCl. When heated with strong bases such as lithium carbonate, trisodium phosphate or sodium metasilicate, it generates a precipitate due to the ring opening and subsequent crosslinking of chains. Yen-Jen *et al.* studied its effect as a drug deliverer and intracellular acceptor [40].

4.1.1. Molecular docking

In this work, the molecular docking calculations were performed using the AutoDock program. In particular, it uses a Lamarckian genetic algorithm (LGA) and a force field function based approximately on the AMBER force field, which consists of five terms: 1) the 12-6 dispersion term of Lennard-Jones, 2) a 12-10 directional hydrogen bonding term, 3) an electrostatic Coulomb potential, 4) an entropic term, and 5) one term of desolvation pairs. The scaling factor of these terms is empirically calibrated using a set of 30 structurally-known protein-ligand complexes, which affinities have been experimentally determined. The AutoDock program has become widely used due to its good precision and high versatility; moreover, the latest version of AutoDock (version 4.0) added flexible functions to the side chains in the receptor.

4.1.2. Model preparation

In this study we used the E protein of dengue virus. It consists of a dimer with 394 amino acids (aa) per monomer and, as mentioned before, it is the main component of the dengue virus envelope. E protein enters the cell by fusion with the membrane due to a previous conformational change produced by a low pH, generating thereby a change of form from dimer to trimer, in which the fusion peptide between the II and III domains is exposed. When the pH is lower than 6.3, dimers dissociate from dimer phase, making the I and II domains rotate outwards and exposing the fusion loop, which interacts with the endosomal membrane of the cell. Domain III then rotates backwards to pull the I and II domains, which were already bound to the cell membrane by the fusion peptide, thus attaching the cell membrane with the membrane of the virus in order to release the RNA [27, 29, 41-45]. It is important to mention that Bressanelli showed that the virus domains remain at neutral pH but their relative orientation is altered [27]. For best results during the molecular docking process, we optimize the original model of the dengue virus protein E (PDB code 1OKE) with a number of refinements and validations cycles with Phenix and Molprobity programs respectively. Figure 3 shows the corrected model of the dimer and trimer.
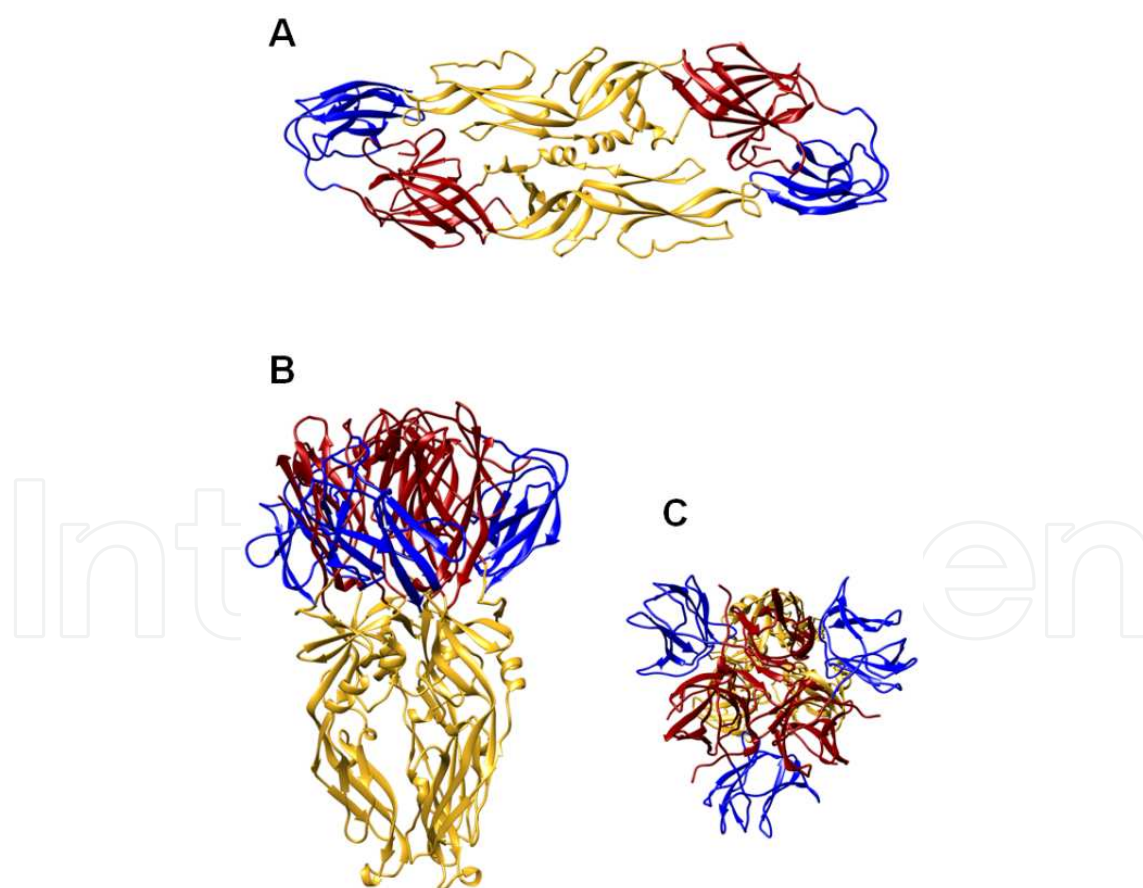
Figure 3. Dengue virus protein E. (A) Dimeric protein after geometric and refining corrections with Phenix program. (B) Trimeric protein that represents a postfusion state (C) Top view of the trimeric form. Domain I (aa$_{1-52,132-193,280-296}$) is in red, domain II (aa$_{52-132, 193-280}$) in yellow and domain III (aa$_{296-394}$) in blue color.

4.1.3. Ligands preparation

To prepare the ligands, we utilized the linear PVP and CS monomers, and CS/TPP/CS chains (Figure 4). The coordinates of those ligands were obtained using the SMILES program [46].
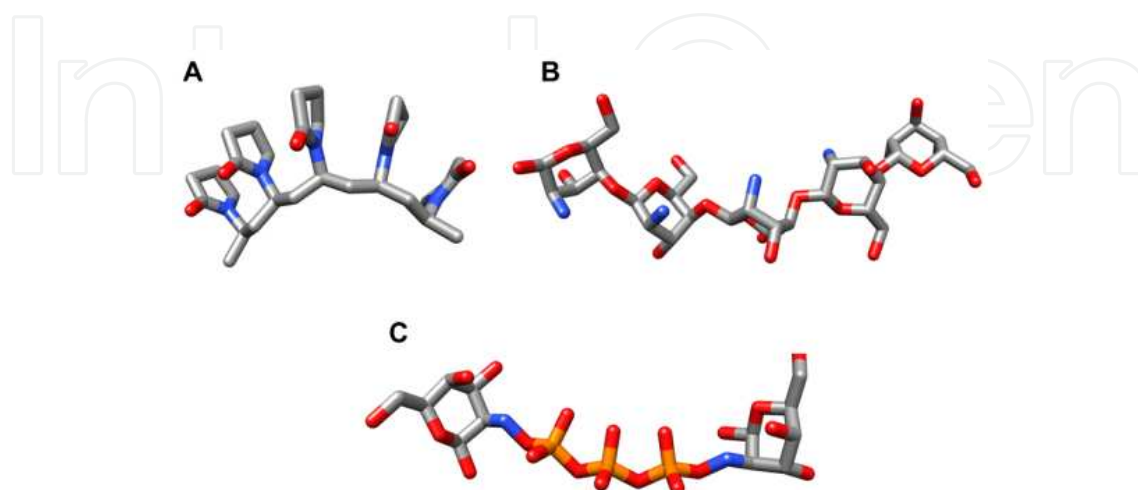
Figure 4. Ligands used in molecular docking. (A) PVP monomer, (B) CS monomer, and (C) CS/TPP/CS chain.

4.1.4. Molecular docking

Molecular docking was performed by means of the AutoDock program that combines rapid grid-based energy evaluation and efficient search of torsional freedom. This program uses a semi-empirical free energy force field to evaluate the conformations during the docking simulations. The force field is quantified using a large number of protein-inhibitor complexes, for which the inhibition constants (K_i), are known. The force field evaluates the union in two steps, first when the ligand and the protein are separated. Then, the intramolecular energies are estimated for the transition from the unbound state to the protein-ligand bound state. In the second step, intermolecular energies are evaluated by combining the ligand with the protein conformations bound to themselves. The force field includes six pairs-wise of evaluations (Vi) and an estimated loss of conformational entropy after binding (ΔS_{conf}):

$$\Delta G = \left(V_{bound}^{L\text{-}L} - V_{unbound}^{L\text{-}L} \right) + \left(V_{bound}^{P\text{-}P} - V_{unbound}^{P\text{-}P} \right) + \left(V_{bound}^{P\text{-}L} - V_{unbound}^{P\text{-}L} + \Delta S_{conf} \right) \tag{1}$$

where L refers to the ligand and P to the "protein" in a ligand-protein docking. Each of the pair-wise energetic terms includes evaluations for dispersion/repulsion, hydrogen bonding, electrostatics, and desolvation [47].

The calculations can be summarized in the following four steps: (1) preparation of files using AutoDockTools coordinates, (2) pre-calculation of atomic affinities by using AutoGrid, (3) docking of ligands by using AutoDock, and (4) analysis of the results applying AutoDockTools.

4.2. Results

4.2.1. Amino acids of interest in the dengue virus infection mechanism

In the loop conformation, several amino acids are involved in trimerization of unit E of DENV. These amino acids are of particular interest since they are allocated in between I and II domains, the fusion loop of the host cell located between domains II and III, and aa268-270 (kl loop). Also are important, the loop of fusion to the host cell located between domains II and III, which subsequently is exposed in the trimer with the aa98-111 fusion peptide, and the C-terminal of domain III, which holds the protein to the virus membrane. Other important amino acids were mentioned by Mazumder [48], who made a structural analysis of the dengue virus E protein of the 4 serotypes in order to find the conserved and exposed sites as well as the epitopes in the T-cell. In our study, we additionally considered the sites of interest described by Yorgo Modis [42, 43] (Table 1).

In addition to the ten conserved regions presented in Table 1, we predicted around 740 E proteins of the 4 serotypes, some of which are included in the same Table 1. Their sequences were quantified using Shannon's entropy [48] with a variation from 0.3 to 1.1 bits.

Amino acid	Dimer configuration	Trimer configuration	Function
L191,T268-I270	This region is known as kl loop, without the presence of the ligand (β-Octylglucoside, BOG), it forms a salt bridge and a hydrogen bond with beta strand I and j of the counterpart dimer.	The ligand is present; however, the kl loop does not adopt the open conformation present in the dimer-ligand pair.	A hinge allowing movement of domains I and II, as well as conformational changes when varying the pH of the endosome.
V382-G385	C terminal that attaches domain III to premembrane (prM) virus.	It combines to create a trimeric contact with the other two domains.	It holds and folds the domains I and II, acting as a zipper. It is the most variable region among the 4 serotypes.
D98-G111	The loop is protected between the domain II and III; it contains a fusion peptide that is formed by hydrophobic residues.	It is exposed only in the trimeric conformation during the conformational changes, and it maintains its structure.	It is the region of highest interest since it is the binding receptor for the host cell. In the trimeric form it fuses and binds the cell and virus's membranes. It is the most conserved region among the 4 serotypes.
N37, P207	Exposed.	Exposed.	Conserved epitope in the 4 serotypes.

Amino acid	Dimer configuration	Trimer configuration	Function
Q211 D215	Exposed.	Not exposed.	Conserved epitope in the 4 serotypes.
H244, K246	Not exposed.	Exposed.	Conserved epitope in the 4 serotypes.
C30,121,105,116,285 and 333	30 hidden, 105 semi-exposed, 116 hidden, 121 semi-exposed, 285 semi-exposed, 333 semi-exposed.	All semi-hidden.	Disulfide bonds in the 4 serotypes. It provides structure to the protein.
R9 E368	Salt bridge, conserved region.	Salt bridge, conserved region.	Structural stability, interactions between the domain I and III.
H144,317	Hydrogen bridges, conserved region.	Hydrogen bridges, conserved region.	Structural stability of the main chain with the opposite domain.
N67 N153	Interacts with N acetyl-D-glucosamine (NAG)	Only N67 interacts with NAG.	Glycosylation sites.

Table 1. Sites of interest identified in the structure of the dengue virus E protein according to Yorgo Modis [42, 43, 48].

The analyzed proteins can be identified as: N8-G14, V24-D42, R73-E79, V97-S102, D192-M196, V208-W220, V252-H261, G281-C285, E314-T319, E370-G374, and K394-G399; whereas the hidden amino acids, which change to exposed amino acids in the trimer, can be listed as follows: M1, H244, K246, G254, G330 and K344; and the exposed residues that remain hidden include the following: S16, Q52, Q167, S169, P243, D290, Q293, S331 and E343. It is worth mentioning that we have identified at least 14 conserved negative sites in at least 3 of the 4 serotypes (C3, C60, R73, T189, F213, A267, F306, T319, S376, F392, K394, S424, G445, and V485). The importance of this discovery relies on the fact that it has demonstrated that the epitopes with negative sites work better as vaccines than those with positive sites as they are less likely to change due to their functional restrictions.
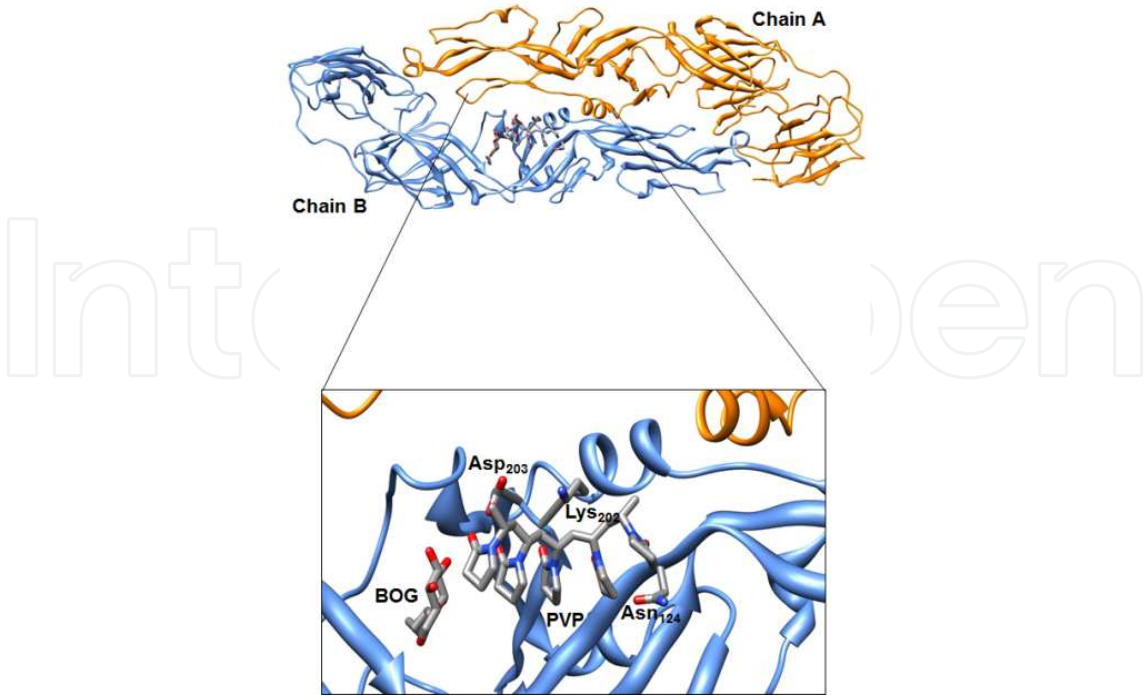
4.2.2. Dengue virus E protein — PVP docking

The docking of PVP molecules with the E protein of dengue virus has demonstrated that the interface of the I and II domain was the most energetically favorable site for the binding (Figure 5). The interaction between protein and ligand takes place by establishing 8 hydrogen bonds with the Asn124, Lys202, and Asp203 amino acids (Table 2). This region is extremely important for the pivotal role it plays in the conformational changes triggered by low pH, which in turn is closely related to the infectivity of the virus. In particular, the PVP molecule, which interacts with aa124,202,203 in the E protein-BOG ligand complex, could act as a blocker of the kl aa268-270 pitchfork activity, which is responsible for the conformational changes in the E protein at low pH. In other words, it could inhibit their function to work as a hinge for conformational changes due to its proximity to amino acids through steric hindrance, preventing thereby

the hinge action between the I and II domain, which in turn could stiffen the area. Alternatively, if BOG ligand is absent, the molecule could be internalized into the hydrophobic pocket and replace it, but the subsequent molecular prediction simulations would be required to determine how it could act in the presence of low pH, in order to find out whether the conformational changes would appear or be inhibited. The PVP is well-known to be highly stable at acid pH and high temperatures, so its structural integrity is assured to remain intact; the loop or kl pitchfork amino acids mutate, resulting in an increase of the pH threshold, at which conformational changes occur. It is achieved by replacing long hydrophobic side chains by the short ones. As the result, the site can be consistently represented as a potential trigger in the virus replication cycle and a good candidate to inhibit its function (Figure 5).

PVP molecule	Amino acid	Distance (Å)
O1	OD1-Asp 203	2.75
O2	OD1-Asp 203	2.44
O3	N-Asp 203	2.64
O3	N-Lys 202	2.58
O4	N-Lys 202	3.37
O4	O-Asn 124	2.92
O5	O-Asn 124	3.32
O5	N-Asn 124	2.87

Table 2. Polar interactions of the PVP molecule with dengue virus E protein.

Figure 5. Dengue virus E protein with PVP ligand. The lower part shows a close up of the docking area.

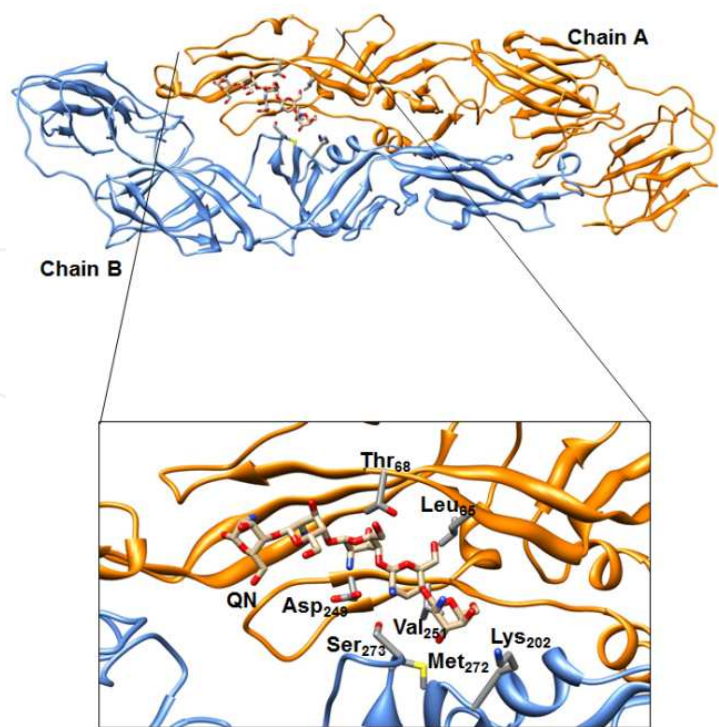*4.2.3. Dengue virus E protein — CS docking*

The docking of CS molecules in the E protein of dengue virus resulted in the interaction with the interface of domain I and II of the protein (see Figure 6). The CS ligand binds to seven amino acids of E protein by ten hydrogen bonds (see Table 3). The elongated CS molecule settles into a channel formed in the II domain surface of the protein. Additionally, it interacts with amino acids near the kl hinge or loop of I and II domain interface. There is a remarkable familiarity between the BOG and NAG complexes. Amino acids-CS molecule interactions, which are shown in Table 10 (aa65,68,202,249,251,272,273), suggest that the mechanism of action of this molecule is similar to PVP ligand. Additionally, it is very close to the conserved region V252-H261 that forms a channel in the 4 serotypes. This finding is of the highest importance since it could very well serve as a ligand for the 4 serotypes, and it could be even more useful in the development of a chimera vaccine with the four domains III of E protein, which would be similar to the chimeric vaccine developed in India at the International Centre for Genetic Engineering and Bionanotechnology.

CS molecule atom	Amino acid atom	Distance (Å)
O1 ring 1	O- Lys 202- B	3.47
O1 ring 1	O-Met 272-B	2.97
N ring 2	OG-Ser 273-B	3.24
O1 ring 2	N-Val 251-A	3.14
O1 ring 2	O-Val 251-A	1.97
O3 ring 2	O-Leu 65-A	3.28
O5 ring 2 and O2 ring 3	OD2-Asp 249-A	3.5
O1 ring 3	OG1-Thr 68-A	2.77
N ring 3	OD2-Asp 249-A	2.51
N ring 3	OD1-Asp 249-A	3.33

Table 3. CS molecule interactions with dengue virus E protein.

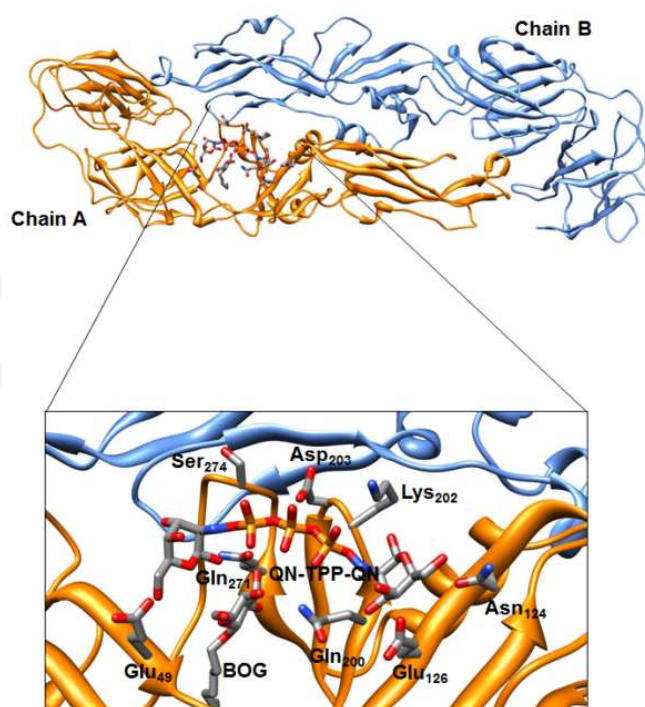*4.2.4. Dengue virus E protein — CS/TPP/CS docking*

In this case, we used the CS and TPP monomers taking into account that the CS units form bindings by means of 1-4 beta bonds. Similarly to the E protein–PVP docking, the molecular docking between the CS/TPP/CS ligand and the E protein was carried out between the domains I and II, although we observed more interactions in the case of PVP monomer. Table 4 and Figure 7 illustrate seven interactions between the amino acids and the BOG. The CS-TPP-CS complex interacts with aa49, 124, 126, 200, 202, 203, 271 amino acids, and the docking results suggest that these three molecules are attracted the most to the area formed by the hydrophobic pocket, indicating that the latter molecule has a direct interaction with the BOG ligand oxygen.

Figure 6. Docking of dengue virus E protein with CS ligand. The interaction takes place at an interface between the two monomers.

CS/TPP molecule atom	Amino acid atom	Distance (Å)
N ring 1	O-Ser 274	3.43
O3 ring 1	OE2-Glu 49	2.90
O5 ring 1	NE2-Gln 271	3.08
O5 ring 1	OE1-Gln 271	2.66
O5 ring 1	O3-BOG	2.74
O1 TPP1	O-Ser274	3.03
O3 TPP1	O4-BOG	3.14
O1 TPP2	OD1-Asp 203	2.83
O3 TPP2	O4-BOG	3.30
O1 TPP3	O-Lys 202	2.19
O1 TPP3	OD1-Asp 203	3.18
O3 TPP3	NE2-Gln 200	3.43
O4 TPP	O-Lys 202	2.51
O1 ring 2	N-Lys 202	2.89
O3 ring 2	ND2-Asn 124	2.89
O3 ring 2	OD1-Asn 124	2.38
O4 ring 2	OE2-Glu 126	2.49

Table 4. Docking of CS/TPP/CS molecule with dengue virus E protein.

Figure 7. CS/TPP/CS molecule docking with the dengue virus E protein.

5. Conclusions

We have reviewed the key concepts and current experimental procedures, including the recent advances in protein flexibility, ligand sampling, and scoring function. In addition, challenges and possible future directions were addressed in this chapter. As an example of protein ligand study we analyzed the interaction between the dengue virus E protein and Polyvinylpyrroli-done and Chitosan biopolymers and we confirmed that PVP, CS, and CS/TPP/CS biopolymers can fulfill the function of adjuvant carriers in the potential development of a chimeric dengue vaccine against the 4 serotypes of dengue virus. Furthermore, the ring-shaped molecules have shown affinity to or preference for a place of vital importance in the virus's cycle of infection and replication, which placed us on the path to develop an inhibitor of the aforementioned conformational changes (see Figures 5-7). Their binding to the E protein is possible due to the great affinity they present to simulated molecules. However, further analysis of molecular simulation is required to determine the behavior of the protein without the presence of BOG ligand or in different environmental conditions in the presence of low pH.

Acknowledgements

Author details

Alejandra Hernández-Santoyo[1], Aldo Yair Tenorio-Barajas[2], Victor Altuzar[2],
Héctor Vivanco-Cid[3] and Claudia Mendoza-Barrera[2*]

*Address all correspondence to: omendoza@uv.mx

1 Instituto de Química, Universidad Nacional Autónoma de México, Mexico, D.F., Mexico

2 Laboratorio de Nanobiotecnología, Centro de Investigación en Micro y Nanotecnología, Universidad Veracruzana, Boca del Rio, Veracruz, Mexico

3 Instituto de Investigaciones Médico-Biológicas, Universidad Veracruzana, Boca del Río, Veracruz, Mexico

References

[1]   Huang, S. Y, & Zou, X. Advances and challenges in protein-ligand docking. International Journal of Molecular Sciences (2010). , 11(8), 3016-3034.

[2]   Halperin, I, Ma, B, Wolfson, H, & Nussinov, R. Principles of docking: An overview of search algorithms and a guide to scoring functions. Proteins (2002). , 47(4), 409-443.

[3]   Sousa, S. F, Fernandes, P. A, & Ramos, M. J. Protein-ligand docking: current status and future challenges. Proteins (2006). , 65(1), 15-26.

[4]   Wang, R, Fang, X, Lu, Y, Yang, C. Y, & Wang, S. The PDBbind Database: Methodologies and updates. Journal of Medicinal Chemistry (2005). , 48(12), 4111-4119.

[5]   Puvanendrampillai, D, & Mitchell, J. B. L. D Protein Ligand Database (PLD): additional understanding of the nature and specificity of protein-ligand complexes. Bioinformatics (2003). , 19(14), 1856-1857.

[6]   Block, P, Sotriffer, CA, Dramburg, I, Klebe, G, & Affin, . : a freely accessible database of affinities for protein-ligand complexes from the PDB. Nucleic Acids Research 2006; 34 D522-526.

[7]   Liu, T, Lin, Y, Wen, X, Jorissen, RN, Gilson, MK, & Binding, . : a web-accessible database of experimentally determined protein-ligand binding affinities. Nucleic Acids Research 2007; 35 D198-201.

[8]   Dias, R. de Azevedo WF Jr. ((2008). Molecular docking algorithms. Curr Drug Targets. Dec; , 9(12), 1040-7.

[9] Kuntz, I. D, Blaney, J. M, Oatley, S. J, Langridge, R, & Ferrin, T. E. A geometric approach to macromolecule-ligand interactions. Journal of Molecular Biology (1982). , 161(2), 269-88.

[10] Jain, A. N. Scoring functions for protein-ligand docking. Current Protein Peptide Science (2006). , 7(5), 407-20.

[11] Gilson, M. K, & Zhou, H. X. Calculation of protein-ligand binding affinities. Annual Review of Biophysics and Biomolecular Structure (2007). , 36-21.

[12] Huang, S. Y, Grinter, S. Z, & Zou, X. Scoring functions and their evaluation methods for protein-ligand docking: recent advances and future directions. Physical Chemistry Chemical Physics (2010). , 12(40), 12899-908.

[13] Taylor, R. D, Jewsbury, P. J, & Essex, J. W. A review of protein-small molecule docking methods. Journal of Computer-Aided Molecular Design. (2002). , 16(3), 151-66.

[14] Chen, V. B, & Arendall, W. B. rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC. MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallographica Section D: Biological Crystallography. (2010). Pt 1):12-21.

[15] Morris, G. M, Huey, R, Lindstrom, W, Sanner, M. F, Belew, R. K, Goodsell, D. S, & Olson, A. J. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. Journal of Computational Chemistry (2009). , 16-2785.

[16] Rarey, M, Kramer, B, Lengauer, T, & Klebe, G. A fast flexible docking method using an incremental construction algorithm. Journal of Molecular Biology (1996). , 261(3), 470-89.

[17] Jones, G, Willett, P, Glen, R. C, Leach, A. R, & Taylor, R. Development and validation of a genetic algorithm for flexible docking. Journal of Molecular Biology (1997). , 267(3), 727-748.

[18] Abagyan, R, Totrov, M, & Kuznetsov, D. ICM-A new method for protein modeling and design: Applications to docking and structure prediction from the distorted native conformation. Journal of Computational Chemistry (1994). , 15(5), 488-506.

[19] Mizutani, M. Y, Tomioka, N, & Itai, A. Rational automatic search method for stable docking models of protein and ligand. J. Mol. Biol. (1994). , 243, 310-326.

[20] Taylor, J. S, & Burnett, R. M. DARWIN: A program for docking flexible molecules. Proteins (2000). , 41, 173-191.

[21] Clark, K. P. Flexible ligand docking without parameter adjust-ment across four ligand-receptor complexes. J. Comput. Chem. (1995). , 16, 1210-1226.

[22] Hart, T. N, & Read, R. J. A multiple-start Monte Carlo docking method. Proteins (1992). , 13, 206-222.

[23] World Health OrganizationWHO: Media Centre, Fact sheets: Dengue and severe dengue. http://www.who.int/accessed 15 August (2012).

[24] Huang, J. H, Wey, J. J, Sun, Y. C, Chin, C, Chien, L. J, & Wu, Y. C. Antibody responses to an immunodominant nonstructural 1 synthetic peptide in patients with dengue fever and dengue hemorrhagic fever. Journal of Medical Virology (1999). , 57(1), 1-8.

[25] Muñoz, M. L, Cisneros, A, Cruz, J, Das, P, Tovar, R, & Ortega, A. Putative dengue virus receptors from mosquito cells. FEMS Microbiology Letter (1998). , 168(2), 251-258.

[26] Cao-lormeau, V. M. Dengue viruses binding proteins from Aedes aegypti and Aedes polynesiensis salivary glands. Virology Journal (2009).

[27] Bressanelli, S, et al. Structure of a flavivirus envelope glycoprotein in its low-pH-induced membrane fusion conformation. The EMBO Journal (2004). , 23(4), 728-738.

[28] Mondotte, J. A, Lozach, P. Y, Amara, A, & Gamarnik, A. V. Essential Role of Dengue Virus Envelope Protein N Glycosylation at Asparagine-67 during Viral Propagation. Journal Virology (2007). , 81(13), 7136-7148.

[29] Mukhopadhyay, S, Kuhn, R. J, & Rossman, M. G. A structural perspective of the flavivirus life cycle. Nature Reviews Microbiology (2005). , 13-22.

[30] Stiasny, K, Allison, S. L, Schalich, J, & Heinz, F. X. Membrane Interactions of the Tick-Borne Encephalitis Virus Fusion Protein E at Low pH. Journal of Virology (2002). , 76(8), 3784-3790.

[31] Mandl, C. W, Guirakhoo, F, Holzmann, H, Heinz, F. X, & Kunz, C. Antigenic structure of the flavivirus envelope protein E at the molecular level, using tick-borne encephalitis virus as a model. Journal of Virology (1989). , 63(2), 564-571.

[32] Acosta-bas, C, & Gómez-cordero, I. Biología y métodos diagnósticos del dengue. Revista Biomedica (2005). , 16-113.

[33] Kelly, E. P, Greene, J. J, King, A. D, & Innis, B. L. Purified dengue 2 virus envelope glycoprotein aggregates produced by baculovirus are immunogenic in mice. Vaccine (2000). , 18(23), 2549-2559.

[34] Putnak, R, et al. Immunogenic and protective response in mice immunized with a purified, inactivated, Dengue-2 virus vaccine prototype made in fetal rhesus lung cells. The American Journal of Tropical Medicine Hygiene (1996). , 55(5), 504-510.

[35] Staropoli, I, Grenckiel, M. P, Mégret, F, & Deubel, V. Affinity-purified dengue-2 virus envelope glycoprotein induces neutralizing antibodies and protective immunity in mice. Vaccine (1997).

[36] Heinz, F. X, & Stiasny, K. Flaviviruses and flavivirus vaccines. Vaccine, (2012). , 30(29), 4301-4306.

[37] Sundar, S, Kundu, J, & Kundu, S. C. Biopolimeric nanoparticles. Science and Technology of Advanced Materials (2010).

[38] Robinson, B. V, Sullivan, F. M, Borzelleca, J. F, & Schwart, S. L. PVP : a Critical Review of the Kinetics and Toxicology of Polyvinylprrolidone (Povidone). Chelsea, MI: Lewis Publishers; (1990).

[39] Buhler, V. Polyvinylpyrrolidone Excipients for Pharmaceuticals: Povidone, Crospovidone, and Copovidone. Berlin, New York: Springer-Verlag; (2005).

[40] Wang, Y. J, Chien, Y. C, Wu, C. H, & Liu, D. M. Magnolol-loaded core-shell hydrogel nanoparticles: drug release, intracellular uptake, and controlled cytotoxicity for the inhibition of migration of vascular smooth muscle cells. Molecular Pharmaceutics (2011). , 8(6), 2339-2349.

[41] Harrison, S. C. The pH sensor for flavivirus membrane fusion. The Journal of Cell Biology (2008). , 183(2), 177-179.

[42] Modis, Y, Ogata, S, Clements, D, & Harrison, S. C. Structure of the dengue virus envelope protein after membrane fusion. Nature (2004). , 427-313.

[43] Modis, Y, Ogata, S, Clements, D, & Harrison, S. C. A ligand-binding pocket in the dengue virus envelope glycoprotein. Proceeding of the National Academic Science of the United State of America (2003). , 100(12), 6986-6991.

[44] Yu, I. M, Holdaway, H. A, Chipman, P. R, Kuhn, R. J, Rossmann, M. G, & Chen, J. Association of the pr Peptides with Dengue Virus at Acidic pH Blocks Membrane Fusion. Journal of Virology (2009). , 83(23), 12101-12107.

[45] Sánchez- San Martin CLiu CY, Kielian M. Dealing with low pH: entry and exit of alphaviruses and flaviviruses. Trends in Microbiology (2009). , 17(11), 514-521.

[46] Weininger, D. SMILES, a Chemical Language and Information System. 1. Introduction to Methodology and Encoding Rules. Journal of Chemical Information and Computational Science (1998). , 28(1), 31-36.

[47] Garrett, M, Morris, D. S. G, & Michael, E. Pique, William "Lindy" Lindstrom, Ruth Huey, Stefano and W.E.H. Forli, Scott Halliday, Rik Belew and Arthur J. Olson, User Guide AutoDock Version 4.2, Automated Docking of Flexible Ligands to Flexible Receptors. (2010). , 49.

[48] Mazumder, R, Hu, Z. Z, Vinayaka, C. R, & Sagripanti, J. L. Frost SDW, Kosakovsky P, Wu CH. Computational analysis and identification of amino acid sites in dengue E proteins relevant to development of diagnostics and vaccines. Virus Genes (2007). , 35(2), 175-186.