



**SCHOOL OF TECHNOLOGY AND EXPERIMENTAL SCIENCES**

**DEPARTMENT OF PHYSICAL AND ANALYTICAL CHEMISTRY**

**Degree final project**

**INTRODUCTION TO COMPUTATIONAL  
CHEMISTRY.  
STUDY OF THE INTERACTIONS BETWEEN  
INHIBITORS AND ENZYMES**

**Author**

*Lledó Moreno Delgado*

**Tutor**

*Raquel Castillo Solsona*

**Chemical Degree**

*2020/2021*



## INDEX

1. Abbreviations .....	iii
2. Introduction .....	1
2.1. Computational chemistry.....	2
2.1.1. Molecular methods .....	2
2.1.1.1. Molecular Mechanics methods (MM) .....	2
2.1.1.2. Quantum Mechanics methods (QM) .....	7
2.1.1.3. Hybrid methods (QM/MM) .....	9
2.1.1.4. Molecular Dynamics (MD) .....	10
2.1.2. System studied .....	13
2.1.3. Methodology.....	15
3. Objectives .....	19
4. Results and analysis .....	23
4.1. System setup .....	25
4.1.1. Previous study .....	27
4.1.2. Minimization .....	28
4.1.3. Heat .....	28
4.2. Molecular dynamics (MD) .....	28
4.2.1. Results of MD simulations .....	30
4.2.1.1. RMSD .....	30
4.2.1.2. Interactions .....	33
5. Conclusions .....	47
6. References .....	51
7. Annexes .....	55



# ***Abbreviations***



<b>MM</b>	<i>Molecular Mechanics methods</i>
<b>QM</b>	<i>Quantum Mechanics methods</i>
<b>QM/MM</b>	<i>Quantum Mechanics/Molecular Mechanics (Hybrid methods)</i>
<b>MD</b>	<i>Molecular Dynamics</i>
<b>VEGF</b>	<i>Vascular Endothelial Growth Factor</i>
<b>VEGFR</b>	<i>Vascular Endothelial Growth Factor Receptor</i>
<b>VEGFR-2</b>	<i>Vascular Endothelial Growth Factor Receptor-2</i>
<b>RTKs</b>	<i>Tyrosine Kinase Receptors</i>
$\hat{H}$	<i>Hamiltonian operator</i>
<b>HF</b>	<i>Hartree-Fock</i>
<b>DFT</b>	<i>Density Functional Theory</i>
<b>VMD</b>	<i>Visual Molecular Dynamics</i>





## ***Introduction***

The work carried out in this study consists of a preliminary introductory step in the area of computational chemistry in order to apply the knowledge acquired, and to be able to model different inhibitors of the system responsible for vascular endothelial growth known as "Vascular Endothelial Growth Factor Receptor-2" (VEGFR-2) involved in the angiogenesis process that takes place in cancer cells.

## 2.1. Computational chemistry

In the last decades, computational chemistry has become an essential tool in the field of research. This is because it has a very important role in facilitating a large number of calculations that are almost impossible to solve by hand or with the help of a calculator.

The transformation into a method of great importance lies in the availability of high-power and quality computers, as well as the development of new methods and optimal and efficient computer programs.<sup>[1]</sup>

Every time, the number of users who consume this type of tool increases, since its use is so simple that it can be used without previous knowledge or with very basic notions.

Generally, computational chemistry is applied when a mathematical method is adequately developed to allow its automation and be carried out on a computer. However, it should be noted that they are not exact calculations but approximate models because they are qualitative or approximately quantitative methods, which provide useful data for researchers.<sup>[2]</sup>

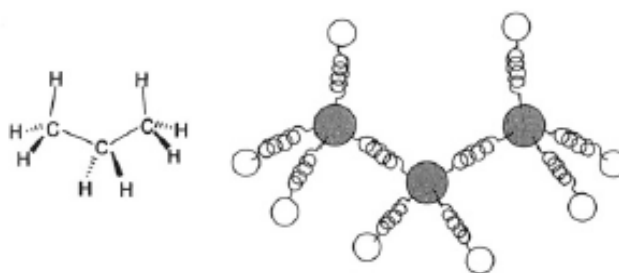
### 2.1.1. Molecular methods

In this study, different molecular methods have been used to carry out the research and thus be able to solve mathematical calculations and predict the behavior of the molecules, which are shown below.

#### 2.1.1.1. Molecular Mechanics methods (MM)

Molecular mechanics (MM) is based on a mathematical model of a molecule that applies the laws of classical physics and considers atoms as a set of spherical balls with a mass and partial charge joined by springs that correspond to the bonds that are established among them.

The electrons in this type of method are not taken into account, so it implicitly uses the Born-Oppenheimer approximation.<sup>[3]</sup>



**FIG 1.** *Molecular mechanics (the force field method) considers a molecule to be a collection of balls (the atoms) held together by springs (the bonds)*<sup>[3]</sup>

The application range of this type of model is quite wide, since it can be applied to a multitude of systems due to its high efficiency. In this way, it allows to apply the calculations on liquid and solid samples and solutions.<sup>[1]</sup>

In addition, it serves to generate the potential energy function under which the molecules move to carry out the molecular dynamics calculations.<sup>[3]</sup>

One of the basic principles on which this model is based is to express the energy of a molecule as a function of its resistance to different parameters such as resistance of the bonds to stretching or bending forces. Therefore, it is based on the search for minimum energy geometry obtained from a mathematical expression.<sup>[1]</sup>

From here, a new concept called " *force field* " arises, which develops from the energy and the parameters of the mathematical expression.<sup>[3]</sup>

This concept has a great diversity of meanings, however, two of them should be highlighted:

On the one hand, the first definition applies to any system described by its own force field "an equation used to calculate the energy of a specific model system, comprised of a collection of atoms or atom-sized particles, and to calculate the derivative of energy, i.e. the forces acting on the atoms".<sup>[1]</sup>

On the other hand, in the second definition each force field is used for a different type of molecule, regardless of the number of them. In this way, it can be defined as "a set of generic rules for constructing a molecular mechanics energy expression for a member of a general class of model

*systems, including a definition of which energy terms should be used, and which parameter values should be used in the energy terms".<sup>[1]</sup>*

An example of this would be the force field for liquid water, which can be applied to many different sizes of liquid water models from one or two molecules to millions of them.

Since force fields are designed to predict properties of unknown molecules, the lack of parameters is often an obvious problem. This can be reflected in the expression of the energy of the force fields, since it requires many parameters such as:

- Force constants  $k_{stretch}$  and  $k_{bend}$
- The equilibrium bond lengths  $r_o$
- Angles  $\theta_o$
- The torsional constants  $A_{dihedral}$  and associated  $n$  and  $\alpha_o$
- The terms for improper torsion
- The non-bonded parameters  $\sigma$ ,  $\epsilon$  and  $q$  <sup>[1]</sup>

However, in practice there are different procedures to minimize the number of parameters required from quantum calculations. The main simplification, used to reduce the number of necessary variables, introduces the concept of *type atoms*, which are those that share the same expected bonding and interaction properties.<sup>[1]</sup>

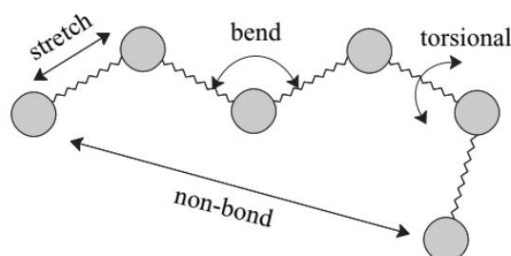
Due to the number of parameters on which it depends, the force field energy is described as the sum of terms responsible for the energy required to distort a molecule in a certain way.

$$E_{FF} = E_{str} + E_{bend} + E_{tors} + E_{vdw} + E_{el} + E_{cross} \quad [4] \quad (1)$$

$E_{str}$  is the energy function for stretching a bond between two atoms,  $E_{bend}$  represents the energy required for bending an angle,  $E_{tors}$  is the torsional energy for rotation around a bond,  $E_{vdw}$  and  $E_{el}$  describe the interactions

between non-bound atoms, and finally  $E_{\text{cross}}$  describes coupling between the first three terms.

This can be seen graphically reflected in the following figure:



**FIG 2.** Illustration of the fundamental force field energy terms <sup>[4]</sup>

From this energy function of the nuclear coordinates, by means of optimization processes, geometries and relative energies can be calculated. Stable molecules correspond to surface minima of potential energy and can be located by minimizing this function as a function of nuclear coordinates.

The number of terms on which the global potential energy equation of a system depends varies depending on the type of molecule.

The expression of global potential energy for a simple system takes into account the following characteristics: <sup>[1]</sup>

$$v(r_A, r_B, \theta) = k_{\text{stretch}} (r_A - r_o)^2 + k_{\text{stretch}} (r_B - r_o)^2 + k_{\text{bend}} (\theta - \theta_o)^2 \quad [1] \quad (2)$$

This equation takes into account different parameters given such as the distances ( $r_A$  and  $r_B$ ) between the atoms as well as the angle ( $\theta$ ) formed between them. On the other hand, the values of  $k_{\text{stretch}}$  and  $k_{\text{bend}}$  are known constants of the system to be able to describe the energy variations produced by the movement. Finally, the unknowns  $r_o$  and  $\theta_o$  are reference values corresponding to the equilibrium structure of the molecule. <sup>[1]</sup>

In the case of considering a molecule a little more complicated, it is necessary to take into account a greater number of terms. One of the reasons is because terms of a new type are introduced depending on the formed dihedral angle, which is expressed by:

$$V(\alpha) = A_{dihedral} \{1 + \cos(n(\alpha - \alpha_0))\} \quad [1] \quad (3)$$

Where  $A_{dihedral}$  is the constant that defines the stiffness of the system at rotations produced around the central bond that forms the dihedral angle,  $\alpha$  is the reference angle and  $n$  is an integer that defines the periodicity of the torsion term.

Considering the interaction between molecules, it is necessary to take into account additional energy terms, that is, the so-called repulsion and attraction interactions.

The *repulsion* term is due to the Pauli principle and arises from the superposition between the electron clouds of two different molecules.

$$V_{repulsion}(r_{AB}) = \frac{A_{AB}}{r_{AB}^n} \quad [1] \quad (4)$$

The *attraction* term is due to the effects of the dispersion of London forces between unbound atoms.

$$V_{attraction}(r_{AB}) = \frac{-B_{AB}}{r_{AB}^m} \quad [1] \quad (5)$$

Although they can be treated separately, the two terms are often grouped together in the following equation known as the Lennard-Jones potential.

$$V_{Lennard-Jones}(r_{AB}) = 4 \varepsilon \left\{ \left( \frac{\sigma_{AB}}{r_{AB}} \right)^{12} - \left( \frac{\sigma_{AB}}{r_{AB}} \right)^6 \right\} \quad [1] \quad (6)$$

Where the constant  $\varepsilon$  describes the interaction force between two atoms considered, and  $\sigma$  describes the cut-off distance at which the interaction changes from negative, for large  $r_{AB}$  to positive, at small  $r_{AB}$ .

Normally, electrostatic energy terms between electrically charged atoms are also taken into account in force fields due to Coulomb's law.

Thus, partial charges are assigned to each of the atoms involved, and by adding different terms, the following expression of potential energy arises:

$$V_{Coulomb}(r_{AB}) = \frac{1}{4\pi\epsilon} \times \frac{q_A q_B e^2}{r_{AB}} \quad [1] \quad (7)$$

### 2.1.1.2. Quantum Mechanics methods (QM)

Once the concept of molecular mechanics has been introduced for its application to the calculation of potential energy, two new concepts appear, explained below along with their applications in chemistry. [3]

While molecular mechanics is based on classical physics, quantum mechanics, *ab initio*, semi-empirical and density functional methods are based on modern physics as they belong to quantum chemistry. [3]

Quantum chemistry and mechanics are the main responsible for the movement of electrons under the impact of the electromagnetic force exerted by nuclear charges. To understand the behavior of electrons, it is necessary to understand the equation on which it depends, that is, the Schrödinger equation. [3]

$$\hat{H}_{elec} \Psi(r) = E \Psi(r) \quad [1] \quad (8)$$

where

$$\hat{H}_{elec} = \left(-\frac{\hbar^2}{8\pi^2m} \nabla^2 + V\right) \quad [3] \quad (9)$$

The symbol  $\hat{H}$  is the operator known as the Hamiltonian operator, which represents that an operation is going to be performed, the result of which will be the multiplication of E by  $\Psi$ .

The operation to be carried out on  $\Psi$  (that is (x, y, z)) is “differentiate it twice with respect to x, to y and to z, add the partial derivatives and multiply the sum by  $-\hbar^2/8\pi^2m$ ; then add this result to V times  $\Psi$ ”. The expression  $\hat{H}\Psi$  means  $\hat{H}$  of  $\Psi$ , not  $\hat{H}$  times  $\Psi$ .

There are different methods to solve the Schrodinger equation, these are the *ab initio*, semi-empirical, and density functional theory methods.

On the one hand, *Ab initio* methods are those approximate quantum mechanical calculations derived from theoretical principles that are based on simplifying functions and searching for solutions to differential equations.

Noteworthy in this type of method is the calculation known as Hartree-Fock (HF), which is based on the central field approximation. [2]

This type of method has an advantage since it allows to simplify the Schrödinger equation in numerous simple-to-solve equations of an electron. In this way, a wave function of a single electron called orbital <sup>1</sup> and an energy, called orbital energy, is produced.

Furthermore, the described wave function must be expressed by a mathematical function such as the linear combination of Gaussian-type orbitals, since it is only known for a few systems of an electron.

On the other hand, semi-empirical calculations are those that have the same structure as an HF calculation, having a Hamiltonian function and a wave function, but certain data are approximated or directly omitted.

However, omitting part of the calculation leads to errors, therefore, the method must be parameterized using different parameters obtained from experimental data or *ab initio* calculations. So, despite being faster calculations than the previous method, the results obtained are less reliable by inducing erroneous results. One of the oldest and best known semi-empirical methods is the well-known Hückel method.

In recent decades, density functional theory (DFT) has played an important role as it is a method with fewer computational needs but with similar precisions. This method differs in determining the energy of a molecule, as it is obtained from electron density rather than from a wave function. [2]

---

<sup>1</sup> The orbital describes the behavior of an electron in the net field of all other electrons.



After obtaining the Schrödinger equation from the development of quantum mechanics, due to the application of this equation to chemistry by Hückel, the new concept known as quantum chemistry was created. [3]

Basically, quantum mechanics shows that energy is quantized, that is, it is absorbed and emitted by discrete packets of magnitude  $h\nu$  where  $h$  corresponds to Planck's constant and  $\nu$  to the frequency associated with energy.

Due to the shortcomings of Maxwell's theory, the physicist Bohr proposed the idea that an electron would orbit stably as long as its angular momentum was a multiple of  $h/2\pi$ . However, this model could only be applied to the hydrogen atom, so Schrödinger overcame the shortcomings of the previous method by a combination of classical wave theory and the de Broglie postulate,  $\lambda = h/p$ .

Because this model could not be applied to samples more complex than the hydrogen atom, the physicist-chemist Huckel succeeded in applying quantum mechanics to samples with some complexity.

### 2.1.1.3. Hybrid methods (QM/MM)

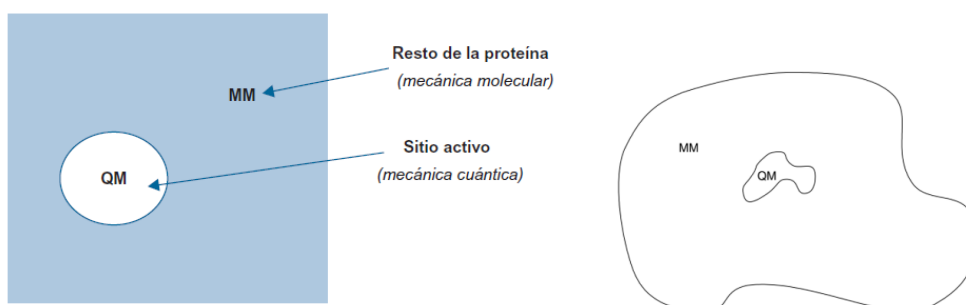
The development of theoretical models to study enzyme systems has made it possible to carry out highly complex studies in the field of computational enzymology.

The evolution of theoretical models has spread its application to describe biomolecular systems, design active principles or enzymes with a desired specific activity, among others.

In this way, by combining the two methodologies described above, that is, MM and QM, the hybrid models of QM/MM emerge.

The QM/MM model is characterized by the separation of the system into different regions. One of them, the "QM region", corresponds to the chemically significant part, that is, the active center, which is treated using a methodology based on quantum mechanics. On the other hand, there is the "MM region",

which comprises the rest of the system and is described by a methodology based on molecular mechanics.



**FIG 3.** Representation of the hybrid QM/MM system. The active site of the protein is described by quantum mechanics, while the protein environment and the solvent is described by a force field<sup>[5]</sup>

The appearance of this new hybrid methodology makes it possible to calculate based on how the Hamiltonian combination has been produced, whether by mechanical or electronic coupling, and by the treatment of the interface formed between both regions. <sup>[5]</sup>

$$H_{complete} = H_{QM} + H_{MM} + H_{QM/MM} \text{ [6] (10)}$$

#### 2.1.1.4. Molecular Dynamics (MD)

Molecular dynamics can be described as a simulation of the behavior of a molecular system dependent on time. In addition, this needs a way to calculate the energy of the system, being the vast majority of times by applying a calculation of molecular mechanics.

By means of the energy expression obtained, the value of the force exerted on the atoms for any given geometry is obtained. <sup>[2]</sup>

Ultimately, it is responsible for studying the distribution of atoms in the analyzed system as well as the different arrangements that they acquire as a function of time.

Atomic displacement can be described by applying classical mechanics using well-known Newton's laws of motion, from the following equations: [1]

$$F = m \cdot a \quad [1] \quad (11)$$

Where "F" represents the force exerted on an atom of mass "m" that experiences a proportional acceleration "a".

In the same way that this equation has been applied to an atom, it can be applied to a set of N atoms. In this case, F corresponds to a 3N-dimensional vector that collects the total amount of force applied to each atom, just as acceleration is a vector that groups together all the corresponding acceleration terms. However, the mass remains the mass of the atoms.

In addition, the force vector "F" is directly related to the potential energy surface represented by V (R) as shown below:

$$F = \frac{-\partial V(R)}{\partial R} \quad [1] \quad (12)$$

In principle, one might think that for simple potential energy surfaces Newton's equation can be solved in a simple way. However, in most cases they can only be solved roughly numerically.

So, for this type of case, the corresponding variables of initial position (R<sub>0</sub>) and its initial velocity are assigned to each atom:

$$V(0) = \frac{dR_0}{dt} \quad [1] \quad (13)$$

In order to be able to precede the new positions and speeds described in a short period of time  $\Delta t$ .

One way to express velocity and acceleration as a function of time using derivatives of position and velocity respectively:

$$v(t) = \frac{\partial R(t)}{\partial t} = \lim_{\Delta t \rightarrow 0} \frac{R(t+\Delta t) - R(t)}{\Delta t} \quad [1] \quad (14)$$

$$a(t) = \frac{\partial^2 R(t)}{\partial t^2} = \frac{\partial v(t)}{\partial t} = \lim_{\Delta t \rightarrow 0} \frac{v(t+\Delta t) - v(t)}{\Delta t} = \frac{1}{m} \times \frac{-\partial V(R)}{\partial R} \quad [1] \quad (15)$$

However, if instead of using small infinitesimal time steps, finite steps in time were used, the so-called Euler equations would be obtained.

$$R(t + \Delta t) = R(t) + v(t)\Delta t \quad [1] \quad (16)$$

$$v(t + \Delta t) = v(t) + a(t)\Delta t \quad [1] \quad (17)$$

Both velocity and acceleration vary over time, so that these equations are not really of much use in practice by misleading the position and velocity variables. Because the time step is small, the associated error is also small. However, if the number of steps increases to reach a certain time, the error accumulates in each of the steps:

$$n_{step} = \tau/\Delta t \quad [1] \quad (18)$$

For this reason, mathematical procedures have been developed that induce minor errors, in addition to allowing longer time steps.

An example of this type of procedure is the integration method proposed by the physicist Verlet to simulate molecular dynamics, known as the Velocity Verlet method.

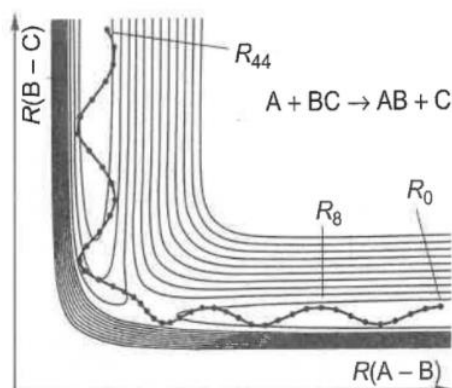
$$R(t + \Delta t) = R(t) + v(t)\Delta t + \frac{1}{2}a(t)\Delta t^2 \quad [1] \quad (19)$$

$$v(t + \Delta t) = v(t) + \frac{1}{2}\{a(t) + a(t + \Delta t)\}\Delta t \quad [1] \quad (20)$$

Calculations of dynamic trajectories using the previous method can be useful in various contexts.

As for example, in the detailed study of the dynamics of chemical reactions. In this case, if the reaction described in the following figure is considered, by applying the surface of simple potential energy described by the previous method, it will be possible to know if the reaction occurs more easily when the reactants come together until they collide, thus producing a high energy or, conversely, when an energy equal to the vibrational energy of the BC bond is provided in order to break the established bond.

This can be seen graphically by the following representation of the trajectory:

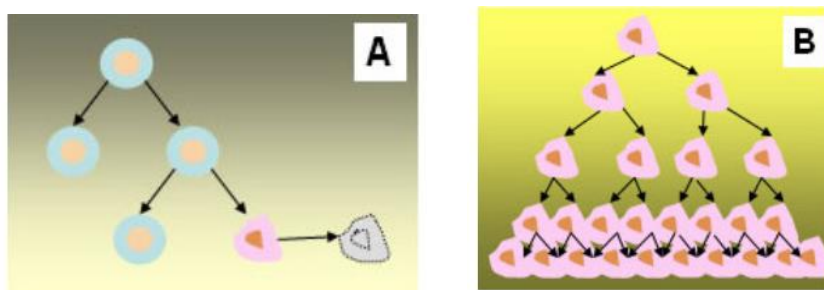


**FIG 4.** Classical trajectory from integration of Newton's law on a reactive potential energy surface. Starting from position  $R_0$  corresponding to reactants  $A+BC$ , the trajectory moves in 44 steps to  $AB+C$ . Such trajectories can be used to predict the conversion of the collision energy into vibrational energy. <sup>[1]</sup>

## 2.1.2 System studied

Nowadays, the term cancer is widely known for the number of diseases it encompasses. It can be described as the development of cells that reproduce uncontrollably in different parts of the body.

While normal cells replicate and die in a controlled way, cancer cells behave abnormally, as they reproduce uncontrollably, destroying or replacing the tissues formed by normal cells. <sup>[7]</sup>

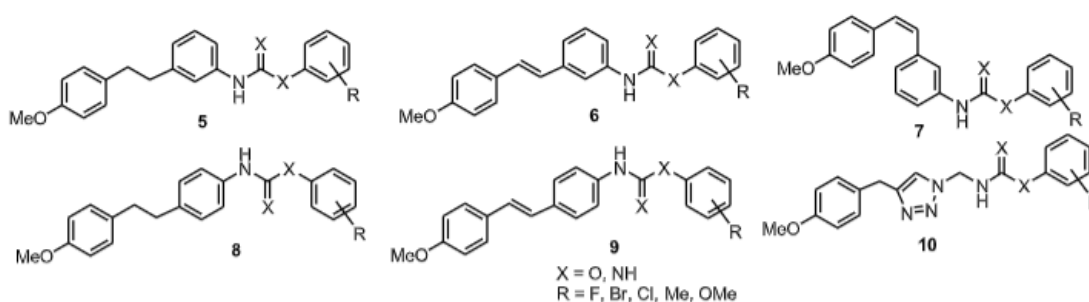


**FIG 5.** Comparison of the growth of a normal cell (A) and a cancer cell (B) <sup>[7]</sup>

In recent years, a large amount of research has been carried out to find new drugs capable of blocking these processes and fighting different diseases.

The progression of cancer cells through the body is related to the physiological process of angiogenesis, which consists of the growth of new blood vessels from previously existing blood vessels by previous processes. This process is regulated due to VEGF through the union produced to the corresponding receptors. [9]

Specifically, this work is based on previous studies carried out by the "Joining Medicine and Chemistry" group at Universitat Jaume I led by Eva Falomir. In which, different inhibitors capable of exerting their function on the VEGFR-2 protein were studied:



**FIG 6.** Possible structures of the inhibitors proposed by the research group "Joining Medicine and Chemistry" group at Universitat Jaume I

Finally, after carrying out the previous study, it was determined that the best candidate is inhibitor 6 with a chlorine in para position. In this way, the main objective of this study is to analyze the effect of placing the different proposed substituents in ortho, meta and para positions. In addition to, thus evaluating the different molecules with anticancer activity (inhibitors) capable of preventing the protein called "vascular endothelial growth factor" (VEGF) from carrying out its function. [8]

This factor, responsible for the growth and survival of endothelial cells, binds to different receptors known as VEGF receptors (VEGFR): VEGFR1, VEGFR2 and VEGFR3. The difference between these receptors is that VEGFR-1 is applied to endothelial cells and other types of cells while VEGFR-2 is applied only to endothelial cells. [9]

Specifically, in this study only VEGFR-2 has been evaluated, which is the one with a more important role due to has become an attractive cancer therapeutic target.

VEGFR2 is a type III transmembrane kinase receptor that includes three domains:

- Extracellular domain
- Transmembrane domain
- Cytoplasmic tyrosine kinase domain.

However, the vast majority of VEGFR2 inhibitors target the tyrosine kinase domain. <sup>[10]</sup>

Through several studies in animals it has been shown that inhibition of VEGF can paralyze tumor growth. Based on these investigations, treatments dedicated to angiogenesis have been implemented, mainly to block this factor.

However, despite the large amount of promising data in animal experiments, VEGF blocking by anti-VEGFR monotherapy is ineffective in advanced cases in the medical setting. <sup>[9]</sup>

### 2.1.3. Methodology

#### **LINUX**

Despite the fact that Windows is the most common operating system, more and more users are using Linux due to its great importance in certain aspects.

Linux is an open source operating system that, unlike other systems, allows access to the platform for free as it is not owned by a specific company but by a group of people who contribute to its development. <sup>[11]</sup>

Among the main features it offers are:

- *Free*: It is one of the main reasons why users use this computing tool.
- *Open Source*: Being an open source system allows anyone to develop new capabilities or functions and make them available to everyone.
- *Sure*: Few are interested in creating a virus that harms the system as it is a free system. However, it consists of a structure that prevents the presence of viruses or malware or, on the contrary, is capable of easily eliminating it.
- *Multi-user*: A large number of users can use the services it offers simultaneously.

- *Customizable*: Allows free modifications.
- *High device control*: Unlike other systems it offers the possibility of thoroughly controlling each device in order to act on any problem.
- *Independent*: It allows free distribution and modification without the need for prior permissions.
- *Stable*: Being one of the most robust and stable systems, it can be used in electronic devices that have to remain switched on without interruptions.
- *Scalable*: Being a system that has open source code and can thus be modified by any of the users, it has the ability to adapt to any change without producing any change.

#### VIM

Configurable text editor created to develop and modify texts efficiently. [12]

#### NAMD

Parallel molecular dynamics code developed to carry out the simulation and optimization of high-performance biomolecular systems. [13]

#### VMD (Visual Molecular Dynamics)

Molecular visualization program used to visualize and analyze large biomolecular systems using 3D representations, such as the one studied in this research.

This allows you to observe the dynamics in different time intervals and thus be able to analyze the results obtained. In addition to being able to observe the different interactions between the protein and the different inhibitors through the exact position of each of the atoms, and the distance, angle and dihedrals formed between them. [14]

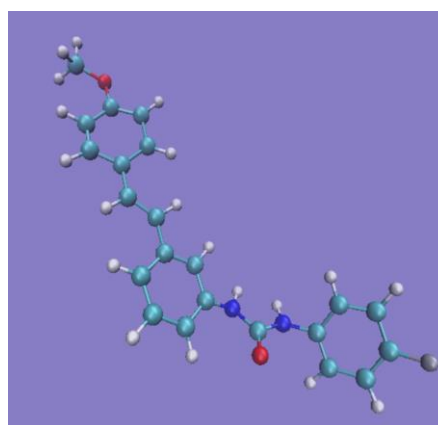


FIG 7. visualization of I6FP inhibitor in VMD



## DYNAMO

Visual programming tool that provides the ability to define behavior visually with scripts, define custom logic and scripts using various textual programming languages.<sup>[15]</sup>

## PuTTY

Type of open source software that makes use of a set of lines of text with the steps that the computer must follow to execute.<sup>[16]</sup>

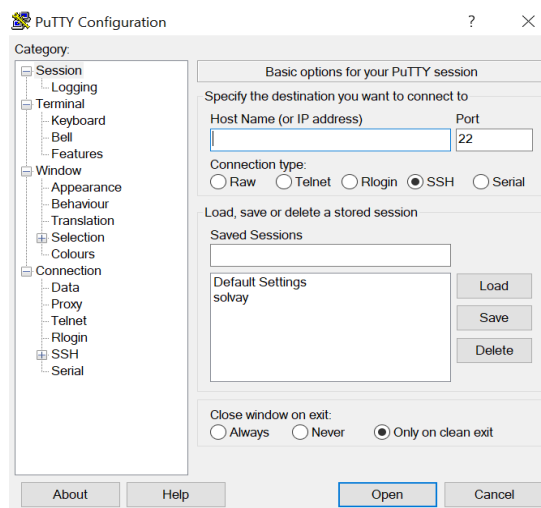


FIG 8. PuTTY configuration

## FileZilla

Computer platform that allows you to connect computer devices to manage, access and share files instantly. For this, it is necessary to have the required data such as server address, username and password.<sup>[17]</sup>

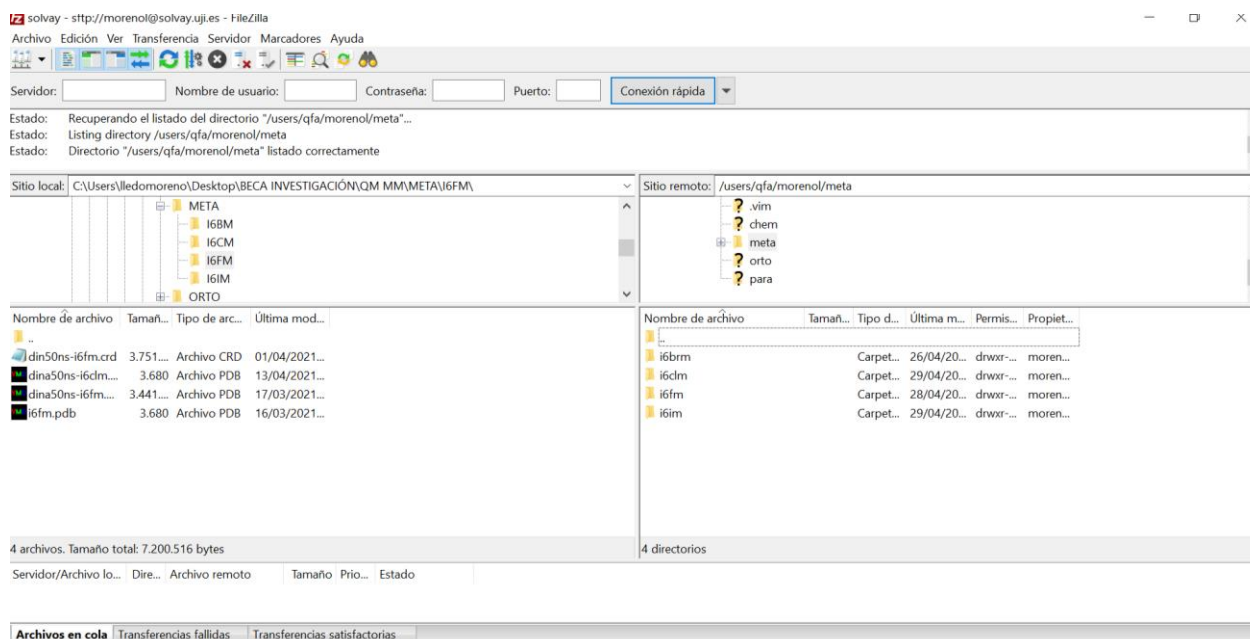


FIG 9. FileZilla main window

### **GaussView**

Graphical interface used together with software known as Gaussian, to build or develop molecular structures of interest. In addition to, execute, control and visualize Gaussian calculations graphically.

This program was developed to be able to carry out investigations of systems of great chemical interest in a simple way.

Therefore, it is a useful tool in the research carried out in this study to be able to construct the different inhibitors. <sup>[18]</sup>

### **Molden**

Program that beyond being able to do optimization work, has a Z matrix editor that gives you control over the geometry as well as being able to build molecules. <sup>[19]</sup>

# ***Objectives***



After completing a research scholarship in the area of physical chemistry tutored by professor Raquel Castillo, the amount of knowledge acquired on computational chemistry has increased throughout all these months.

Although the objectives during the course of this time have been different, these can be summarized below.

On the one hand, to acquire new knowledge necessary to be able to make use of different computer programs until now unknown such as PuTTY, Filezilla or also visualization programs such as VMD. All of them have been necessary to carry out this study.

Although most of the time has been dedicated to understanding carefully, the preparation of the Degree final project, other additional interests have been proposed.

Hence, the objectives to be carried out are:

- Introduction to computational chemistry
- Acquire new knowledge about quantum chemistry
- Learn to use the various quantum and visualization programs
- Study the interaction between inhibitors and the VEGFR-2 protein
- Determine the optimal inhibitor for VEGFR-2 protein
- Visualize atomic interactions between enzyme and inhibitor



# ***Results and analysis***





## 4.1. System setup

The chosen system in order to carry out molecular dynamics was the molecule determined by Eva Falomir group, which corresponds to the i6 inhibitor whose structure is shown on the right:

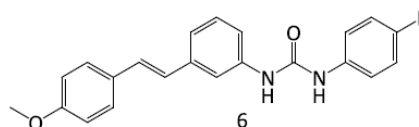


FIG 10. Structure of the inhibitor I6

Changes have been made to its structure using Molden program to prepare the proposed inhibitors in this study. In this way, different halogens, specifically fluorine, chlorine, bromine and iodine, have been exchanged in ortho, meta and para positions.

The identification of each one of them is assigned in three groups depending on the position occupied by the halogen through the following nomenclature:

### PARA

**i6fp**: Inhibitor with the **fluorine** atom in the **para** position

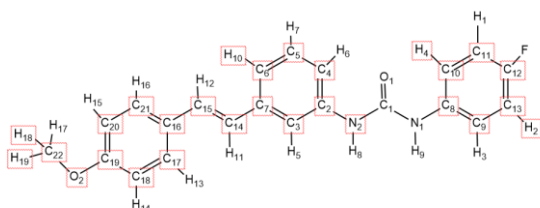


FIG 11. Visualization in VMD of I6FP

**i6bp**: Inhibitor with the **bromine** atom in the **para** position

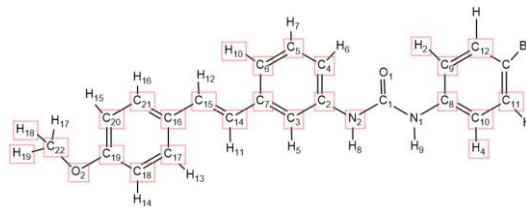


FIG 12. Visualization in VMD of I6BP

**i6cp**: Inhibitor with the **chlorine** atom in the **para** position

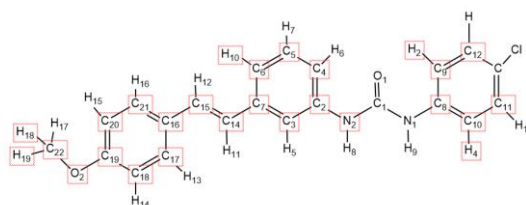


FIG 13. Visualization in VMD of I6CP

**i6ip**: Inhibitor with the **iodine** atom in the **para** position

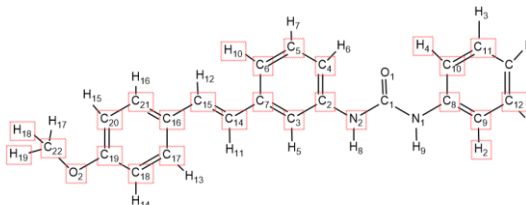


FIG 14. Visualization in VMD of I6IP

## META

**i6fm:** Inhibitor with the **fluorine** atom in the **meta** position

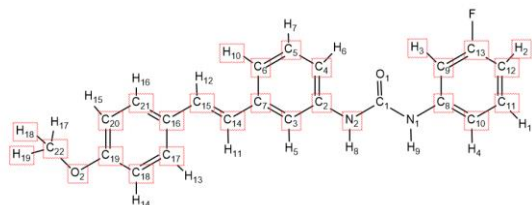


FIG 15. Visualization in VMD of I6FM

**i6bm:** Inhibitor with the **bromine** atom in the **meta** position

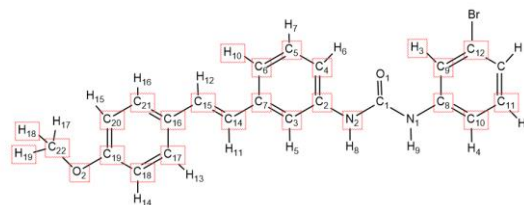


FIG 16. Visualization in VMD of I6BM

**i6cm:** Inhibitor with the **chlorine** atom in the **meta** position

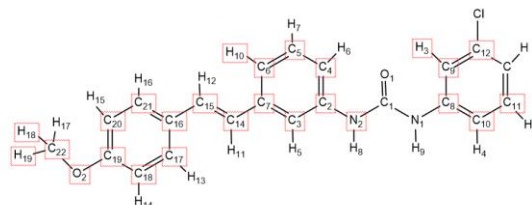


FIG 17. Visualization in VMD of I6CM

**i6im:** Inhibitor with the **iodine** atom in the **meta** position

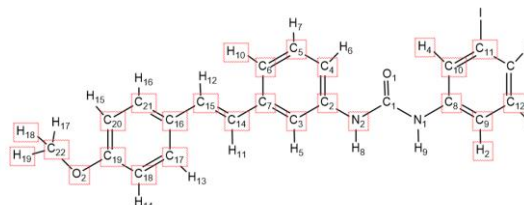


FIG 18. Visualization in VMD of I6IM

## ORTO

**i6fo:** Inhibitor with the **fluorine** atom in the **ortho** position

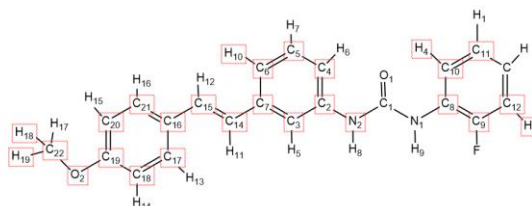


FIG 19. Visualization in VMD of I6FO

**i6bo:** Inhibitor with the **bromine** atom in the **ortho** position

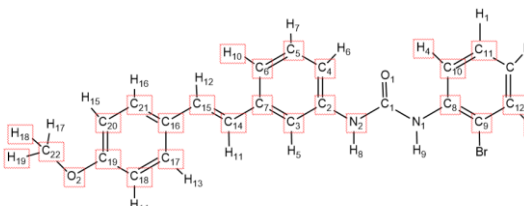
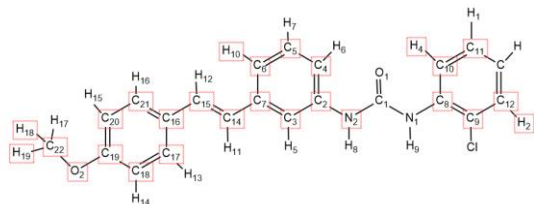


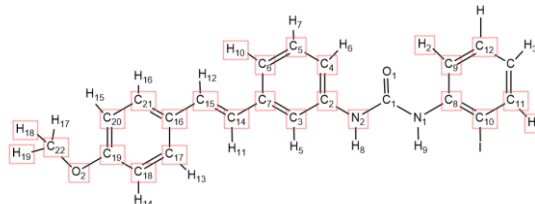
FIG 20. Visualization in VMD of I6BO

**i6co**: Inhibitor with the **chlorine** atom in the **ortho** position



**FIG 21.** Visualization in VMD of I6CO

**i6io**: Inhibitor with the **iodine** atom in the **ortho** position



**FIG 22.** Visualization in VMD of I6IO

Once the structures of the studied inhibitors have been determined, each of the changes carried out from the initial structure (i6fp) are visualized using the GaussView program and saved together with the protein in ".pdb" format.

Next, the steps to follow are carried out to finally execute the molecular dynamics. Which are based on:

- Previous study
- Perform minimization
- Heat

### 4.1.1. Previous study

Thus, the first step is to obtain the PSF file (Protein Structure File) from the previously created protein file together with the inhibitor, which contains the system parameters. This file contains the information necessary to apply the force field, which consists of data of interest such as the type of atom and its parameters, the position of the atoms, bonds, angles and dihedrals, among others. Once the PSF is obtained, a box of water molecules is built and the protein with the possible inhibitors is introduced into it. In addition to adding the counterions (three chlorine atoms) necessary to obtain a neutral system.

Required files to obtain psf file:

```
psfgen      top_all27_prot_na.prm
            i6fp.charmm.rtf
            prt.pdb
            i6fp.pdb
            ion.pdb
            box1.pdb
            box2.pdb
```

## 4.1.2. Minimization

Subsequently, the minimization process is carried out to ensure that the atoms do not overlap. For this, the structure of the system is optimized using the NAMD program, in order to obtain the structure with the minimum energy to start the molecular dynamics.

Required files to minimize the system:

	i6fp.psf
	proti6fbox.pdb
minimization.inp	par_all27_prot_na.prm
	i6fp.prm

## 4.1.3. Heat

Finally, the previous step to molecular dynamics consists of subjecting the system to an increase in temperature until reaching a temperature of 300K.

Required files to heat the system:

	i6fp.psf
	minimization-i6fp.coor
heat.inp	par_all27_prot_na.prm
	i6fp.prm

## 4.2. Molecular dynamics (MD)

Once the necessary files to execute the molecular dynamics have been obtained, which are:

- i6fp.psf: File that contains the information about the initial positions of the atoms of the system
- heat-i6fp.coor: File that contains the information about the final coordinates of the atoms once they have been heated
- par\_all27\_prot\_na.prm: File that contains the information of all the parameters of the amino acids that make up the protein
- i6fp.prm: File that contains the information of all the parameters of the atoms that compose the inhibitor

They are used to generate the system input "dina50ns-i6fp.inp" and make the modifications required to execute the order through the NAMD program and thus start the process, which takes approximately eighty hours as it is a dynamic with a number of steps equal to 50ns.

```

structure i6fp.psf
coordinates heat-i6fp.coor

paraTypeCharmm on
parameters par_all27_prot_na.prm
parameters i6fp.prm

#fixedatoms on
#extrabonds on
#extrabondsfile const_file

cellBasisVector1 82.3 0.0 0.0
cellBasisVector2 0.0 81.7 0.0
cellBasisVector3 0.0 0.0 64.4

#constantForce on
#constantForce fiyo.pdb
#consForceScaling 10.

exclude scaled1-4
l-4scaling 1.0
switching on
switchdist 12.0
cutoff 14.0
pairlistdist 16.0
wrapAll on

temperature 300
#firsttimestep 10000000
timestep 1
numsteps 5000000
nonbondedFreq 2
fullElectFrequency 4
stepspercycle 20

#minimization on

#binvelocities dina10ns.vel
    
```

**FIG 23.** Input used to execute the dynamics of the i6fp inhibitor

Once the calculation has finished, the data file obtained after the molecular dynamics ".dcd" contains the information about the trajectory described every picosecond. In this way, from this it is possible to know all information about the studied system, namely, inhibitors and protein.

The distances between the inhibitor and the most relevant amino acid at the beginning of the trajectory (0ns) and at the end of the trajectory (50ns) are observed through the VMD program and recorded in a table as a summary:

- 0 ns: heat-i6fp.coor
- 50 ns: dina50ns-i6fp.coor

i6FP	
0 ns	50 ns
OW1759-H8	OW9585-F
HE2Met93-C22	OW2884-H1
HG1Met49-O2	OW5244-C8
OGln27-C21	HW5244-O1
OW9501-C3	OW5244-O1
OW1759-N2	OW5244-C10
OAsp176-C4	HW5244-C10
Nlle29-C18	
OW9351-N1	
OW7163-C13	
OW2239-F	
OE2Glu65-H9	
HNAsp176-O1	
HNAsp176-C1	
HB1Cys175-O1	

**FIG 24.** Distances between inhibitor and selected amino acid

## 4.2.1. Results of MD simulations

Once this point is reached, the RMSD representations of all the systems are made and the distances that occur between each inhibitor and protein are analyzed, in order to interpret the most favorable. In this way, the most suitable inhibitor among those studied is determined later.

### 4.2.1.1. RMSD

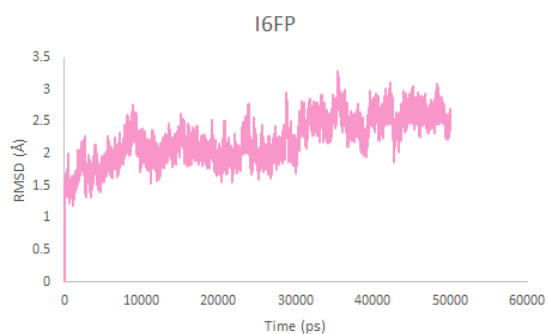
RMSD whose initials refer to Root-Mean-Square Deviation, refers to the average of the average distances between the atoms of different structures. Therefore, this representation allows to graphically observe the stability of the different interactions over the previously established time.

The result of the dynamics can be seen graphically as at the beginning of the same, the RMSD is zero since the movement of the structure is initially null. After a certain time, that is, once the dynamics has started, the RMSD values are different from zero.

Next, two different types of cases occur. On the one hand, it is possible that the dynamics remain constant for a period of time and thus say that the system is stable. On the other hand, it is possible that with the number of steps selected the dynamics do not remain constant and oscillate during the 50ns.

For this reason, the RMSD representations of each of the inhibitors studied are carried out. For this, it is necessary to use some commands of AmberTools, specifically "cpptraj".

In this way, the corresponding file has been obtained for each inhibitor. This contains all the information necessary to represent the graph of RMSD as a function of time.



**FIG 25.** Representation of the RMSD of inhibitor I6FP

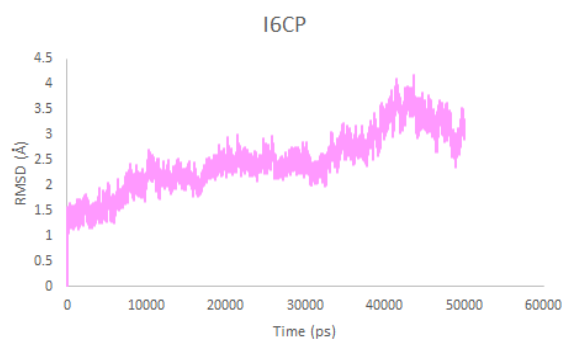
The RMSD graphs of each of the systems are shown below:

Figures 23 to 25 show the RMSD of para inhibitors.

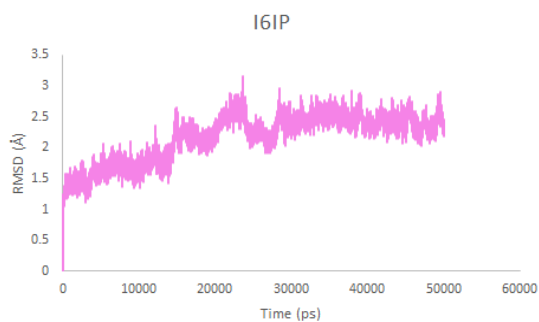
PARA



**FIG 26.** Representation of the RMSD of inhibitor I6BP



**FIG 27.** Representation of the RMSD of inhibitor I6CP

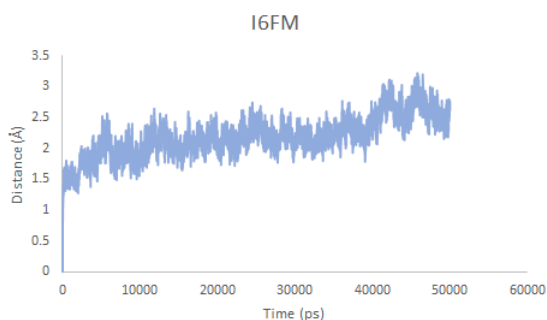


**FIG 28.** Representation of the RMSD of inhibitor I6IP

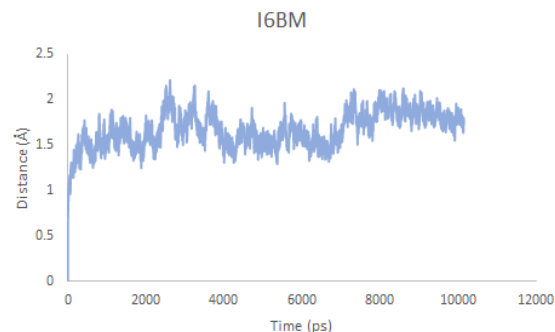
In the case of inhibitors that place the halogen atoms (F, Cl, Br and I) in the para position, it can be seen that most of them show some stability. However, in the case of chlorine, the signal oscillates at the end of the dynamic, which may indicate that it needs a longer amount of time to stabilize.

Figures 26 to 29 show the RMSD of meta inhibitors.

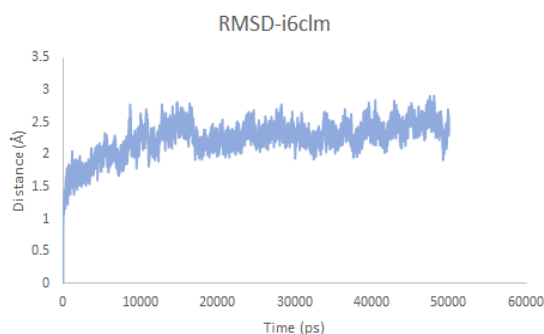
**META**



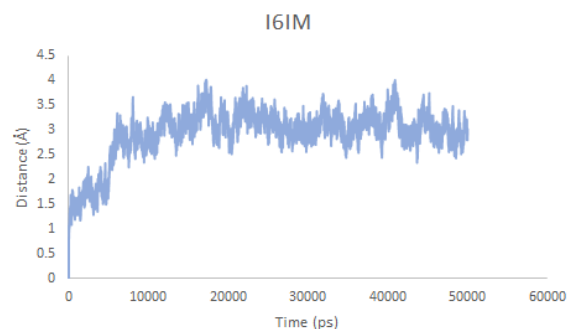
**FIG 29.** Representation of the RMSD of inhibitor I6FM



**FIG 30.** Representation of the RMSD of inhibitor I6BM



**FIG 31.** Representation of the RMSD of inhibitor I6FCM



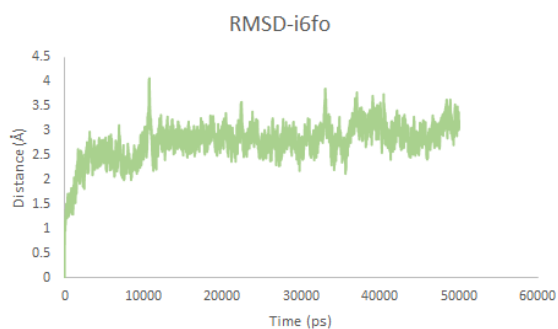
**FIG 32.** Representation of the RMSD of inhibitor I6IM

In the case of inhibitors that place the halogen atoms (F, Cl, Br and I) in a meta position, it can be seen how all of them show a stable trajectory with the number of steps selected. Therefore, it could be thought that it is the most favorable position to place the atoms.

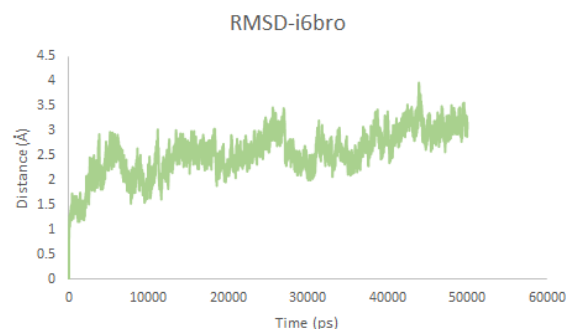


Figures 30 to 33 show the RMSD of ortho inhibitors.

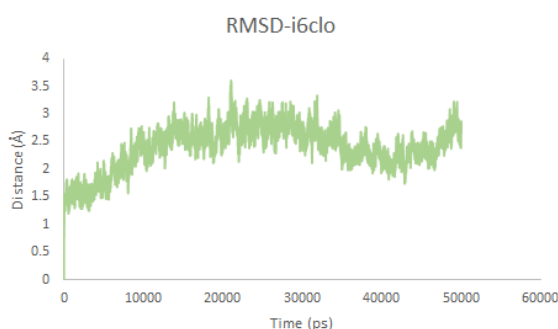
## ORTO



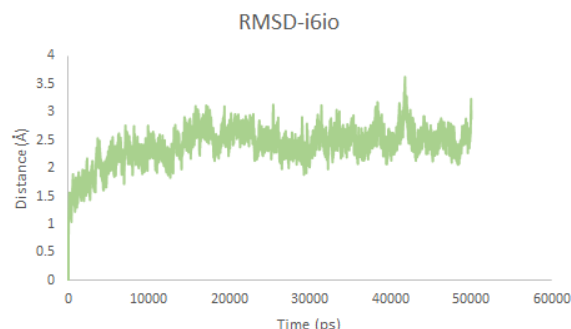
**FIG 33.** Representation of the RMSD of inhibitor I6FO



**FIG 34.** Representation of the RMSD of inhibitor I6BO



**FIG 35.** Representation of the RMSD of inhibitor I6CO

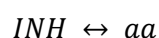


**FIG 36.** Representation of the RMSD of inhibitor I6IO

In the case of inhibitors that place the halogen atoms (F, Cl, Br and I) in ortho position, it can be seen how all of them show a stable trajectory with the number of steps selected. However, it differs from the previous case by the RMSD value by showing higher values.

### 4.2.1.2. Interactions

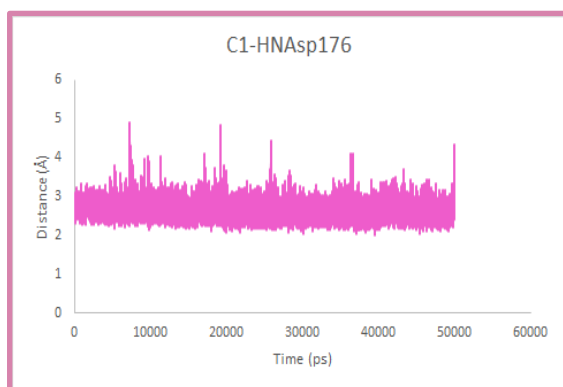
Within the file created "distance.in", each of the possible interactions between the inhibitor and the amino acids are obtained:



The analysis of the interactions allows to determine and classify the most suitable inhibitors along the dynamics described, by showing more favorable interactions with less energetic and therefore more stable systems.

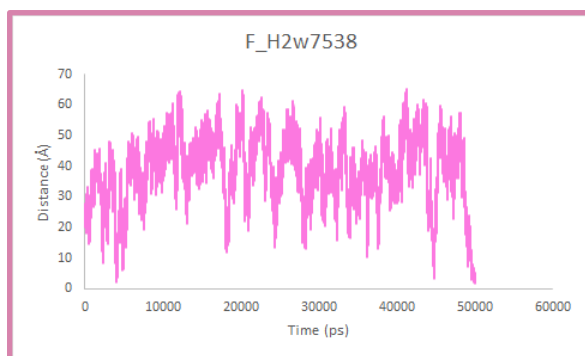
To carry out this analysis, it is necessary to visualize by means of the visualization program, VMD, the different possible distances between inhibitor-amino acids less than 3Å.

Some of the most relevant interactions are shown graphically below, and the rest are included in the annex section.



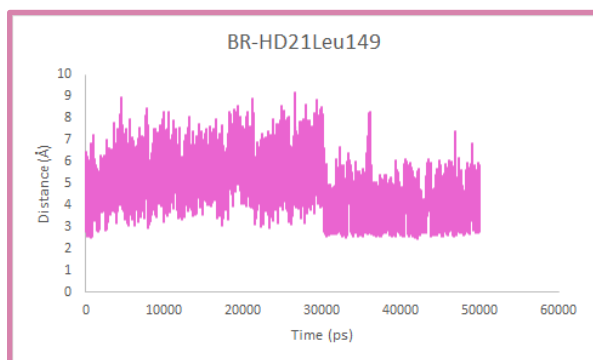
This interaction, which belongs to the bond between carbon 1 of the inhibitor and the hydrogen of the amino acid, Aspartate 176, turns out to be favorable since it is found throughout the stable trajectory with a distance around 2-3Å.

**FIG 37.** Representation of the distance between atom C1 of inhibitor I6FP and atom HN of Asp176 (amino acid)



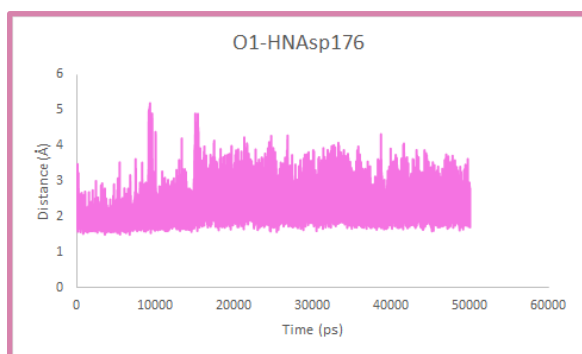
This interaction, which belongs to the bond between the fluorine atom of the inhibitor and the hydrogen atom of one solvent water 7538, turns out to be unfavorable since during the whole trajectory it is unstable with a distance very far from the previously admitted one (3Å).

**FIG 38.** Representation of the distance between atom F of inhibitor I6FP and atom H2 of water7538 (amino acid)



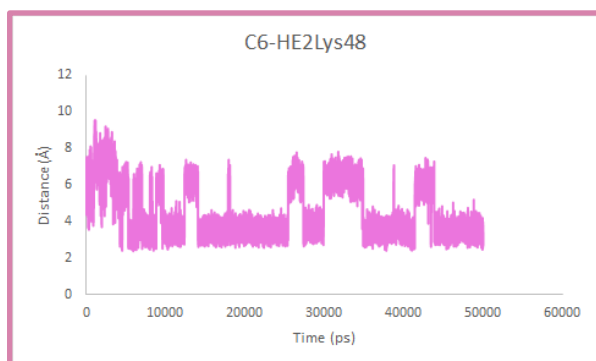
**FIG 39.** Representation of the distance between atom Br of inhibitor I6BP and atom HD21 of Leu149 (amino acid)

This interaction, which belongs to the bond between the bromine atom of the inhibitor and the hydrogen of Leucine 149, turns out to be stable by remaining constant during a wide interval of the trajectory. However, until the end of it it is not considered favorable, since previously the value of the distance is greater than 5 angstroms as it is a value greater than 2-3Å.



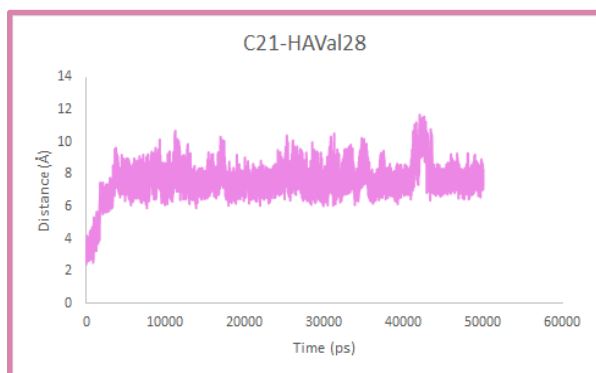
**FIG 40.** Representation of the distance between atom O1 of inhibitor I6CP and atom HN of Asp176 (amino acid)

This interaction, which belongs to the bond between the oxygen atom 1 of the inhibitor and the hydrogen of Aspartate 176, shows a stable and favorable trajectory over time.



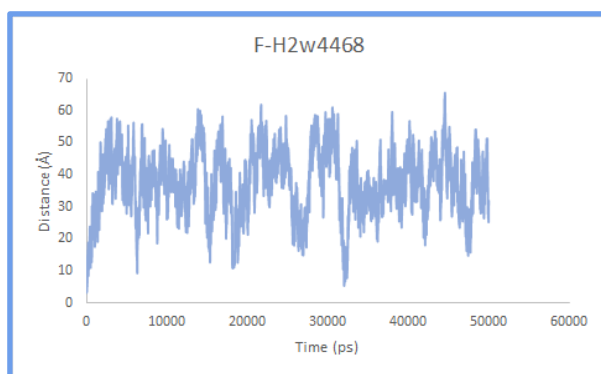
**FIG 41.** Representation of the distance between atom C6 of inhibitor I6IP and atom HE2 of Lys48 (amino acid)

This interaction, which belongs to the bond between the carbon atom 6 of the inhibitor and the hydrogen of Lysine 48, shows a strange trajectory over time. This is because on occasions, the distance turns out to be unfavorable because it exceeds the preset value. But, in time intervals it stabilizes and turns out to be favorable.



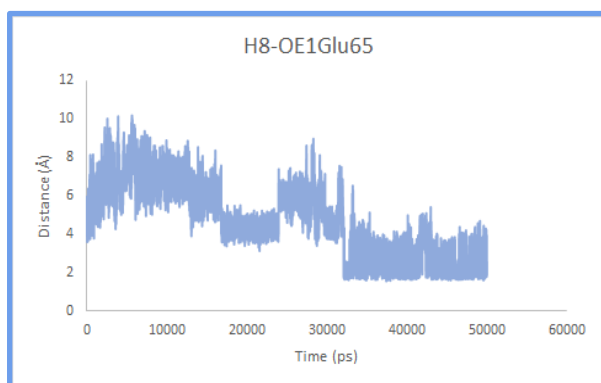
**FIG 42.** Representation of the distance between atom C21 of inhibitor I6IP and atom HA of Val28 (amino acid)

This interaction, which belongs to the bond between the carbon atom 21 of the inhibitor and the hydrogen of Valine 28, shows a stable but unfavorable trajectory over time. This is due to the fact that although it shows a constant trajectory with hardly any variations, the distance is too great to be considered an interaction.



**FIG 43.** Representation of the distance between atom F of inhibitor I6FM and atom H2 of water4468 (amino acid)

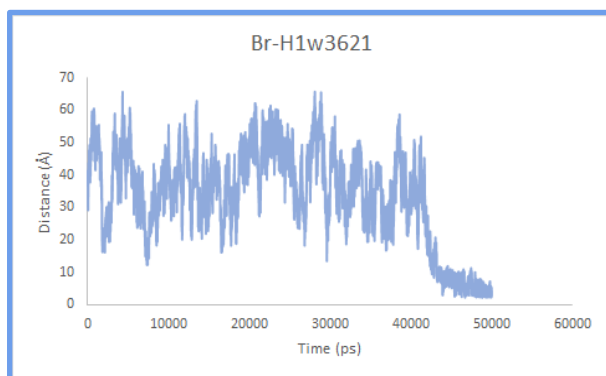
This interaction, which belongs to the bond between the fluorine atom of the inhibitor and the hydrogen of the water 4468, shows an unstable and unfavorable trajectory over time. This is due to the fact that it presents relevant oscillations between the distance of both molecules with values much higher than previously established.



**FIG 44.** Representation of the distance between atom H8 of inhibitor I6FM and atom OE1 of Glu65 (amino acid)

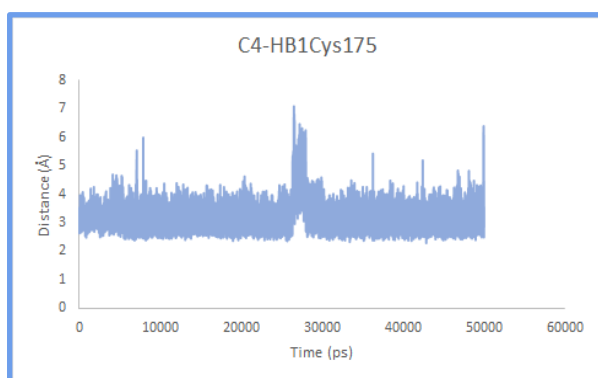
This interaction, which belongs to the bond between the hydrogen atom 8 of the inhibitor and the oxygen of Glutamate 65, shows an unstable but favorable trajectory over time.

This is due to the fact that although for more than half the time the distance is greater than the established one, finally, it is clear that a new interaction has been stabilized.



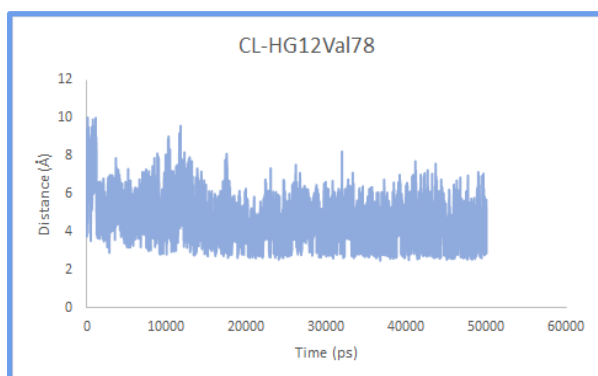
**FIG 45.** Representation of the distance between atom Br of inhibitor I6BM and atom H1 of water 3621 (amino acid)

This interaction, which belongs to the bond between the bromine atom of the inhibitor and the hydrogen of the water 3621, shows an unstable and unfavorable trajectory over time. This is due to the fact that it presents relevant oscillations between the distance of both molecules. However, at the end of the trajectory, it seems that it tends to stabilize, so it could be thought that with a longer time it could stabilize and present a distance around 3 Å.



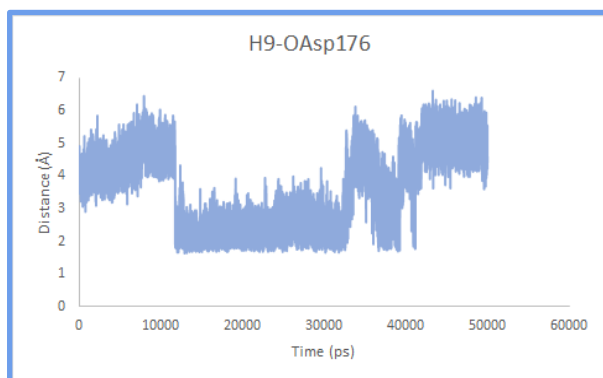
**FIG 46.** Representation of the distance between atom C4 of inhibitor I6BM and atom HB1 of Cys175 (amino acid)

This interaction, which belongs to the bond between carbon atom 4 of the inhibitor and the hydrogen of Cysteine 175, shows a stable and favorable trajectory over time by presenting a distance around the previously admitted one (3 Å).



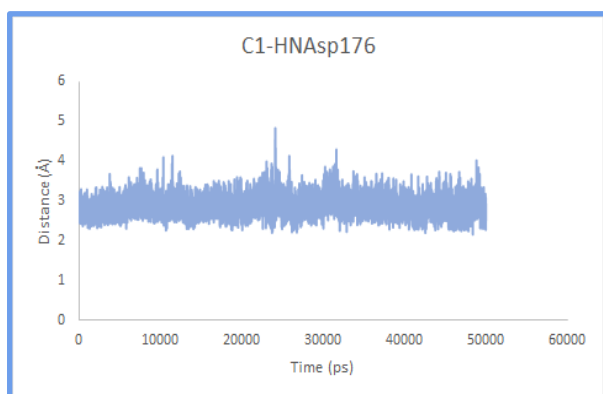
**FIG 47.** Representation of the distance between atom Cl of inhibitor I6CM and atom HG12 of Val78 (amino acid)

This interaction, which belongs to the bond between the chlorine atom of the inhibitor and the hydrogen of Valine 78, shows a stable and favorable trajectory over time. This is due to the fact that although it shows a constant trajectory with hardly any variations, the distance between both atoms is greater than the preset value but acceptable.



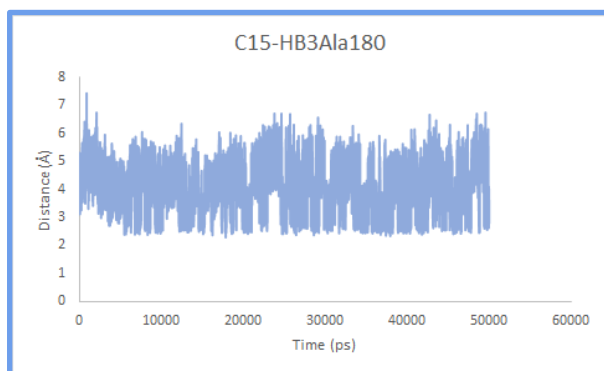
**FIG 48.** Representation of the distance between atom H9 of inhibitor I6CM and atom O of Asp176 (amino acid)

This interaction, which belongs to the bond between the hydrogen atom of the inhibitor and the oxygen of Aspartate 176, shows an unstable and unfavorable trajectory over time. Time-dependent stability until after 10 ns is unstable and after 10 to 30 ns approximation it stabilizes. This is due to the fact that it presents relevant oscillations between the distance of both molecules with values higher than previously established. However, it can be observed how in certain intervals it describes a stable and favorable trajectory by presenting a distance less than the selected one.



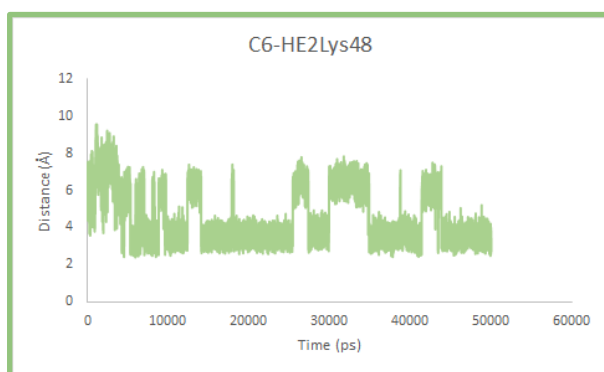
**FIG 49.** Representation of the distance between atom C1 of inhibitor I6IM and atom HN of Asp176 (amino acid)

This interaction, which belongs to the bond between the carbon 1 atom of the inhibitor and the hydrogen of Aspartate 176, shows a stable and favorable trajectory over time. This is because the distance between atoms remains constant throughout the trajectory with a value between 2-3Å.



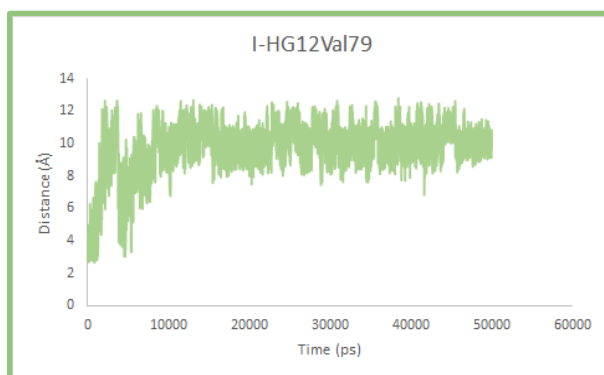
**FIG 50.** Representation of the distance between atom C15 of inhibitor I6IM and atom HB3 of Ala180 (amino acid)

This interaction, which belongs to the bond between the carbon 15 atom of the inhibitor and the hydrogen of Alanine 180, shows a stable but unfavorable trajectory over time. This is because even if the distance remains constant in a range, this turns out to be an interval outside the allowed one, that is, greater than 3Å.



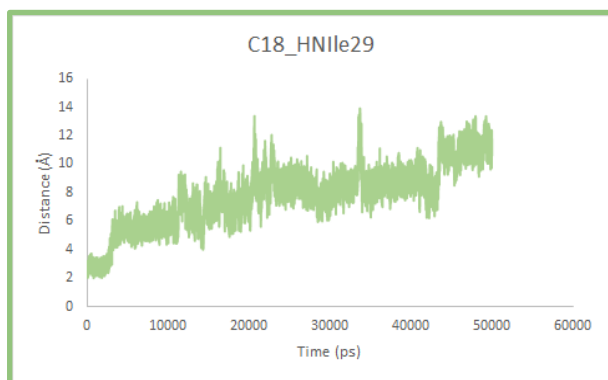
**FIG 51.** Representation of the distance between atom C6 of inhibitor I6FO and atom HE2 of Lys48 (amino acid)

This interaction, which belongs to the bond between the carbon 6 atom of the inhibitor and the hydrogen of Lysine 48, shows an unstable but favorable trajectory over time. This is due to the fact that the trajectory shows oscillations when presenting modifications in the distance values. However, it eventually stabilizes within the allowed range.



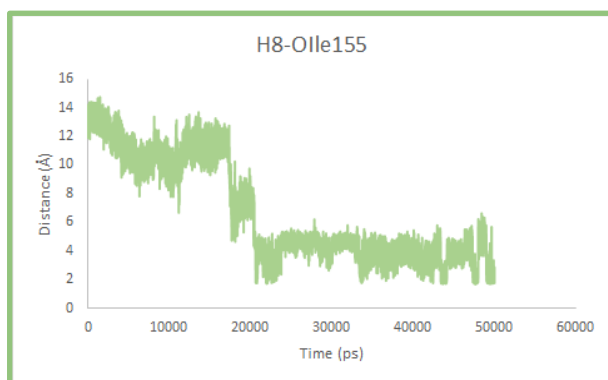
**FIG 52.** Representation of the distance between atom I of inhibitor I6FO and atom HG12 of Val79 of (amino acid)

This interaction, which belongs to the bond between the iodine atom of the inhibitor and the hydrogen of Valine 79, shows a stable but unfavorable trajectory over time. This is because even though the trajectory remains constant over a certain distance, it is outside the established limit, that is, it is greater than 3Å.



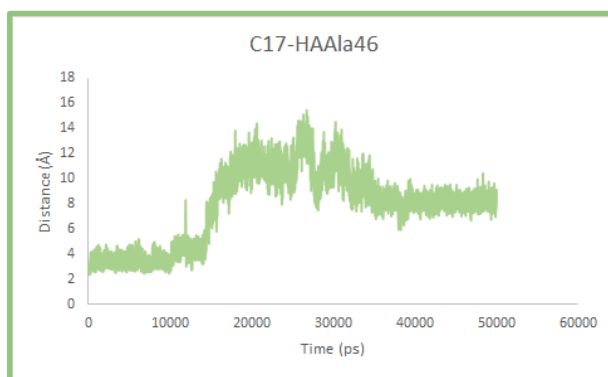
**FIG 53.** Representation of the distance between atom C18 of inhibitor I6BO and atom HN of Ile29 (amino acid)

This interaction, which belongs to the bond between the carbon 18 atom of the inhibitor and the hydrogen of Isoleucine 29, shows an unstable and unfavorable trajectory over time. This is because over time, the distance between the atoms increases.



**FIG 54.** Representation of the distance between atom H8 of inhibitor I6BO and atom O of Ile155 (amino acid)

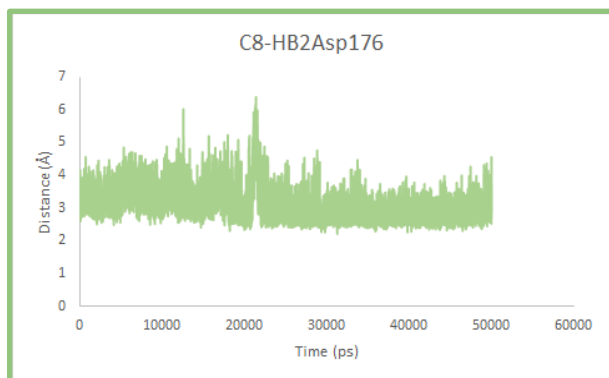
This interaction, which belongs to the bond between the hydrogen 8 atom of the inhibitor and the oxygen of Isoleucine 155, shows an unstable but favorable trajectory over time. Although it initially shows a very high distance between atoms, it finally stabilizes within the accepted range.



**FIG 55.** Representation of the distance between atom C17 of inhibitor I6CO and atom HA of Ala46 (amino acid)

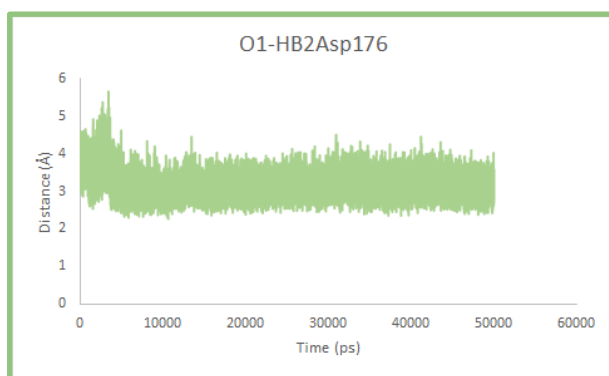
This interaction, which belongs to the bond between the carbon 17 atom of the inhibitor and the hydrogen of Alanine 46, shows an unstable and unfavorable trajectory over time. Although at the beginning it shows an adequate distance between atoms, finally this distance increases and stabilizes outside the accepted range.





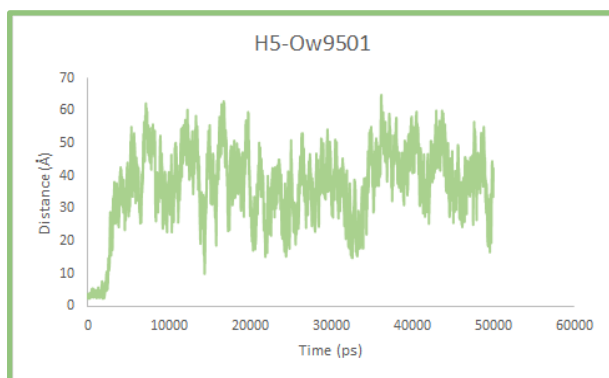
This interaction, which belongs to the bond between the carbon 8 atom of the inhibitor and the hydrogen of Aspartate 176, shows a stable and favorable trajectory over time by keeping the distance constant within the allowed range.

**FIG 56.** Representation of the distance between atom C8 of inhibitor I6CO and atom HB2 of Asp176 (amino acid)



This interaction, which belongs to the bond between the oxygen 1 atom of the inhibitor and the hydrogen of Aspartate 176, shows a stable and favorable trajectory over time by keeping the distance constant within the allowed range.

**FIG 57.** Representation of the distance between atom O1 of inhibitor I6IO and atom HB2 of Asp176 (amino acid)



This interaction, which belongs to the bond between the hydrogen atom 5 of the inhibitor and the oxygen of the water 9501, shows an unstable and unfavorable trajectory over time. This is due to the fact that the trajectory presents oscillations with distances outside the allowed range.

**FIG 58.** Representation of the distance between atom H5 of inhibitor I6IO and atom O of water9501 (amino acid)

Once the presentations of each of the distances between each inhibitor and the amino acids of the protein. All of them are classified as long as the distance of the trajectory remains constant between 2-3 Å, depending on the type of interaction: hydrogen bond, water bond or other type of interaction.

<b>Table 1:</b> Classification of distances between interactions <i>INH (I6FP)</i> ↔ <i>aa</i>		
<b>I6FP</b>		
<b>H bond</b>	<b>H<sub>2</sub>O bond</b>	<b>Other</b>
H9-OE1Glu65		C4-OAsp176
H8-OE1Glu65		H5-HZ1Lys48
H8-OE2Glu65		H5HZ2Lys48
H9-OE2Glu65		
O1-HNAsp176		
O2-HNGln27		
O1-HB1Cys175		
O2-HG1Met49		

<b>Table 2:</b> Classification of distances between interactions <i>INH (I6BP)</i> ↔ <i>aa</i>		
<b>I6BP</b>		
<b>H bond</b>	<b>H<sub>2</sub>O bond</b>	<b>Other</b>
H8-OE1Glu65		BR-HD21Leu149
H8-OE2Glu65		H8-HZ1Lys48
H9-OE1Glu65		C1-HNAsp176
H9-OE2Glu65		
O1-HACys175		
O1-HG12Val79		
O1-HNAsp176		

Table 3: Classification of distances between interactions <i>INH (I6CP)</i> ↔ <i>aa</i>		
<b>I6CP</b>		
<b>H bond</b>	<b>H<sub>2</sub>O bond</b>	<b>Other</b>
H8-OE1Glu65		H4-HNAsp176
H8-OE2Glu65		H8-HD1Lys48
O1-HNAsp176		C1-HNAsp176
O2-HNGly23		C9-HD21Leu69
		C10-HNAsp176
		C14-HB3Ala46
		C15-HG13Val28

Table 4: Classification of distances between interactions <i>INH (I6IP)</i> ↔ <i>aa</i>		
<b>I6IP</b>		
<b>H bond</b>	<b>H<sub>2</sub>O bond</b>	<b>Other</b>
O1-HB2Asp176		H5-HNAsp176
O1-HNAsp176		C6-HE2Lys48
		C17-HD23Leu179
		C20-HB3Ala180
		H15-CE1Phe177

Table 5: Classification of distances between interactions <i>INH (I6FM)</i> ↔ <i>aa</i>		
<b>I6FM</b>		
<b>H bond</b>	<b>H<sub>2</sub>O bond</b>	<b>Other</b>
H8-OE1Glu65		C12-HE1HSE156
H8-OE2Glu65		H16-CBAIa46
H9-OE2Glu65		

<b>Table 6:</b> Classification of distances between interactions <i>INH (I6BM)</i> ↔ <i>aa</i>		
<b>I6BM</b>		
<b>H bond</b>	<b>H<sub>2</sub>O bond</b>	<b>Other</b>
O1-HACys175		H6-HNAsp176
O1-HNAsp176		C4-HB1Cys175

<b>Table 7:</b> Classification of distances between interactions <i>INH (I6CM)</i> ↔ <i>aa</i>		
<b>I6CM</b>		
<b>H bond</b>	<b>H<sub>2</sub>O bond</b>	<b>Other</b>
H8-OE1Glu65		H9-HB2Asp176
H8-OE2Glu65		C15-HG2Cys48
H9-OE1Glu65		N1-HB2Asp176
H9-OE2Glu65		
H16-OGln27		
O1-HB1Cys175		
O1-HNAsp176		

<b>Table 8:</b> Classification of distances between interactions <i>INH (I6IM)</i> ↔ <i>aa</i>		
<b>I6IM</b>		
<b>H bond</b>	<b>H<sub>2</sub>O bond</b>	<b>Other</b>
O1-HNAsp176		C1-HNAsp176

<b>Table 9:</b> Classification of distances between interactions <i>INH (I6FM)</i> ↔ <i>aa</i>		
<b>I6FO</b>		
<b>H bond</b>	<b>H<sub>2</sub>O bond</b>	<b>Other</b>
O1-HNAsp176		H15-CE1Phe177
O1-HB2Asp176		H5-HNAsp176
		C6-HE2Lys48

**Table 10:** Classification of distances between interactions *INH (I6BM)* ↔ *aa*

<b>I6BO</b>		
<b>H bond</b>	<b>H<sub>2</sub>O bond</b>	<b>Other</b>
H8-OIle155		

**Table 11:** Classification of distances between interactions *INH (I6CM)* ↔ *aa*

<b>I6CO</b>		
<b>H bond</b>	<b>H<sub>2</sub>O bond</b>	<b>Other</b>
C1-HNAsp176		C8-HB2Asp176
H8-OAsp176		H9-HB2Asp176
H9-OAsp176		
H12-OVal47		
H14-OGln27		

**Table 12:** Classification of distances between interactions *INH (I6IM)* ↔ *aa*

<b>I6IO</b>		
<b>H bond</b>	<b>H<sub>2</sub>O bond</b>	<b>Other</b>
O1-HNAsp176		H5-HNAsp176
O1-HB2Asp176		C6-HE2Lys48
		H15-CE1Phe177



## ***Conclusions***





Once the corresponding representations have been made, those suitable inhibitors are determined to prevent the angiogenesis process.

While the RMSD representations provide information on the stability of the inhibitor with 50ns, thus indicating whether more or less dynamic time is needed to stabilize the system. The interaction analysis, represented in Tables 1-12, indicates that the more favorable interactions, the more stable the inhibitor will be with the system.

<b>Table 13: Number of favorable interactions for each of the inhibitors</b>				
<b>I6FP</b>	<b>I6BP</b>	<b>I6CP</b>	<b>I6IP</b>	total
11	10	11	7	<b>39</b>
<b>I6FM</b>	<b>I6BM</b>	<b>I6CM</b>	<b>I6IM</b>	
5	4	10	2	<b>21</b>
<b>I6FO</b>	<b>I6BO</b>	<b>I6CO</b>	<b>I6IO</b>	
5	1	7	5	<b>18</b>

In general, it can be observed by the number of favorable interactions that all halogens are most favored in para position, followed by meta position and lastly, ortho position.

However, its difficult to distinguish which of the halogens shows the most favorable interactions in para position by presenting a similar number of favorable interactions. Nevertheless, despite the fact that for fluorine, bromine and chlorine halogens the values are around a similar value, iodine halogen atom can be discarded, as it presents a lower value than the rest.

Therefore, it is analyzed which of the halogen atoms shows in the other positions a greater number of more favorable interactions.

On the one hand, in the meta position it can be seen through the table above how chlorine has a greater number of favorable interactions with respect to the rest of halogens.

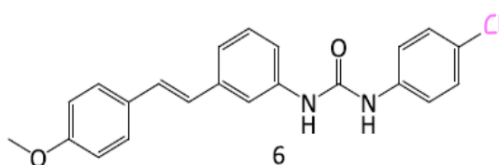
On the other hand, in the ortho position, the halogen atom with the highest number of favorable interactions is chlorine.

In this way, it can be concluded that the inhibitor with chlorine in the para position, I6CP, is the one that in principle should be selected to paralyze the angiogenesis process and thus be able to carry out its synthesis experimentally.

**i6cp:** Inhibitor with the **chlorine** atom in the **para** position



**FIG 59.** Visualization in VMD of I6CP



**FIG 60.** Structure of the inhibitor I6CP

## ***References***



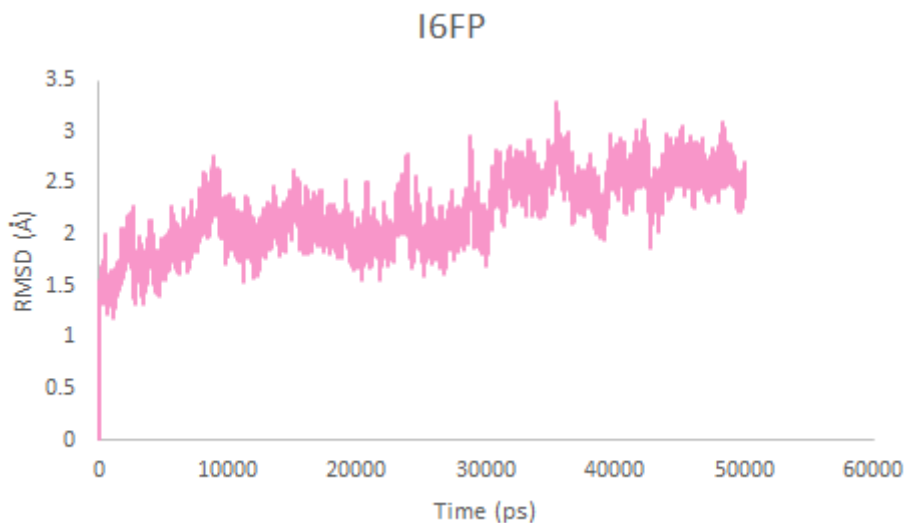
- [1] Harvey, J. (2018). *Computational Chemistry* (Illustrated ed.). Oxford University Press, USA.
- [2] Young, D. (2001). *Computational Chemistry: A Practical Guide for Applying Techniques to Real World Problems* (1.<sup>a</sup> ed.). Wiley-Interscience.
- [3] Lewars, E. G. (2004). *Computational Chemistry: Introduction to the Theory and Applications of Molecular and Quantum Mechanics* (1.a ed.). Kluwer Academic Publishers
- [4] Jensen, F. (2007). *Introduction to Computational Chemistry (English Edition)* (2.<sup>a</sup> ed.). Wiley.
- [5] Ardèvol, A. (2013a). Métodos de simulación QM/MM y de mecánica estadística: una aproximación teórica a la biología molecular. *Anales de la Real Sociedad Española de Química*, 4, 268–275. <https://dialnet.unirioja.es/servlet/articulo?codigo=4559185>
- [6] Cramer, C. J. (2004). *Essentials of Computational Chemistry: Theories and Models* (2nd ed.). Wiley.
- [7] Puente, J., & de Velasco, G. (2019, 16 diciembre). *¿Qué es el cáncer y cómo se desarrolla?* - SEOM: Sociedad Española de Oncología Médica © 2019. SEOM (Sociedad Española de Oncología Médica). <https://seom.org/informacion-sobre-el-cancer/que-es-el-cancer-y-como-se-desarrolla>
- [8] Conesa-Milián, L., Falomir, E., Murga, J., Carda, M., & Marco, J. A. (2019). Novel multitarget inhibitors with antiangiogenic and immunomodulator properties. *European Journal of Medicinal Chemistry*, 170, 87–98. <https://doi.org/10.1016/j.ejmech.2019.03.012>
- [9] Ciciola, P., Cascetta, P., Bianco, C., Formisano, L., & Bianco, R. (2020). Combining Immune Checkpoint Inhibitors with Anti-Angiogenic Agents. *Journal of Clinical Medicine*, 9(3), 675. <https://doi.org/10.3390/jcm9030675>
- [10] Wang, Y., Peng, C., Wang, G., Xu, Z., Luo, Y., Wang, J., & Zhu, W. (2019). Exploring binding mechanisms of VEGFR2 with three drugs lenvatinib, sorafenib, and sunitinib by molecular dynamics simulation and free energy calculation. *Chemical Biology & Drug Design*, 93(5), 934–948. <https://doi.org/10.1111/cbdd.13493>

- [11] Adeva, R. (2021, 4 junio). Todo sobre Linux, el sistema operativo de código abierto. ADSLZone. <https://www.adslzone.net/reportajes/software/que-es-linux/>
- [12] About Vim : vim online. (2007). Vim the Editor. <https://www.vim.org/about.php>
- [13] NAMD - Scalable Molecular Dynamics. (2006). Grupo de Biofísica Teórica y Computacional. <https://www.ks.uiuc.edu/Research/namd/>
- [14] VMD - Visual Molecular Dynamics. (2006). Theoretical and Computational Biophysics Group. <https://www.ks.uiuc.edu/Research/vmd/>
- [15] ¿What is dynamo? | Dynamo Primer. (2019). Dynamo. [https://primer.dynamobim.org/es/01 Introduction/1-2 what is dynamo.html](https://primer.dynamobim.org/es/01%20Introduction/1-2%20what%20is%20dynamo.html)
- [16] Download PuTTY - a free SSH and telnet client for Windows. (2021). PuTTY. <https://www.putty.org/>
- [17] FileZilla - The free FTP solution. (2006). FileZilla. <https://filezilla-project.org/index.php>
- [18] Using GaussView 6 | Gaussian.com. (2018, 25 octubre). Expanding the Limits of Computational Chemistry. <https://gaussian.com/gv6main/>
- [19] MOLDEN a visualization program of molecular and electronic structure. (2016). Molden. <https://www3.cmbi.umcn.nl/molden/>

# ***Annexes***

PARA

i6fp

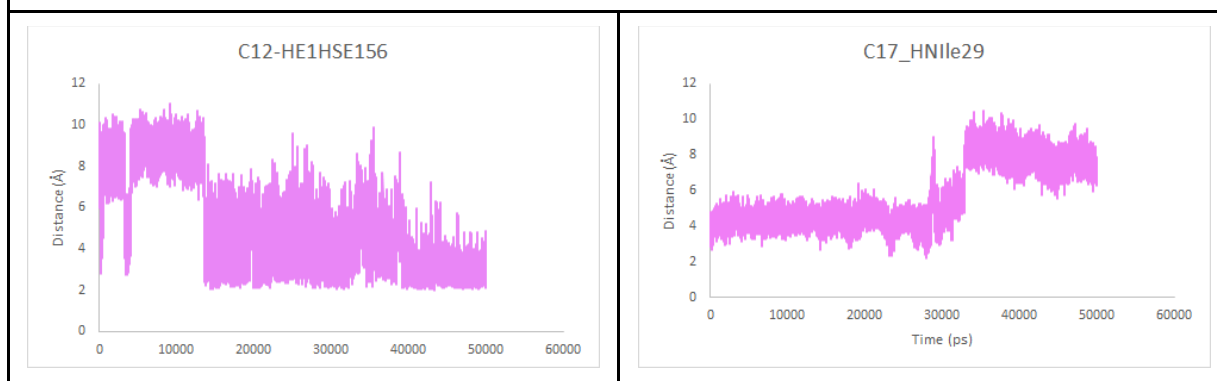


The trajectories described by each of the interactions observed by:

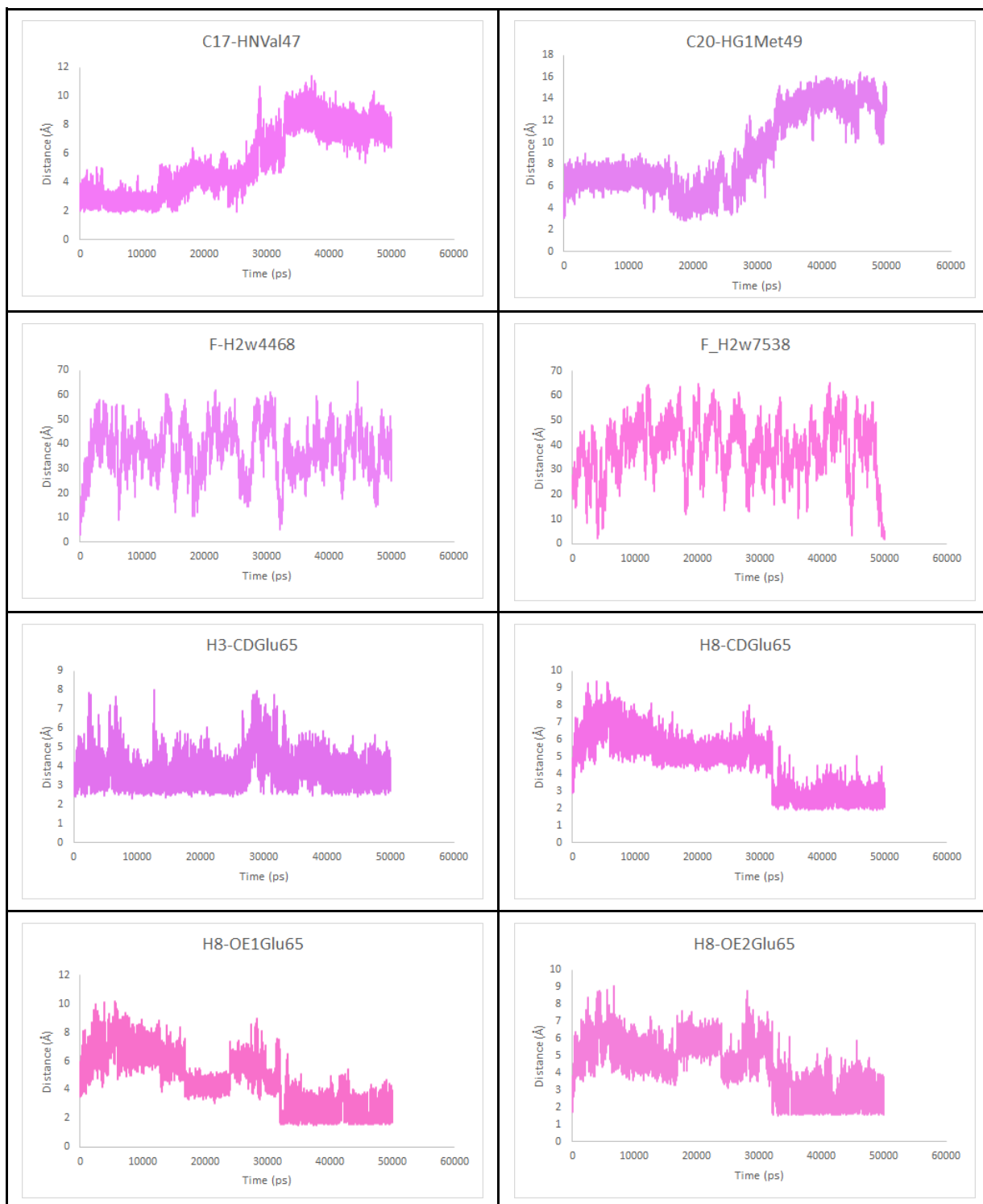
- The program "vmd" at 0ns and 50ns
- The files :  
i6fp\_solute.avg.dat  
solue\_i6fp.avg.dat

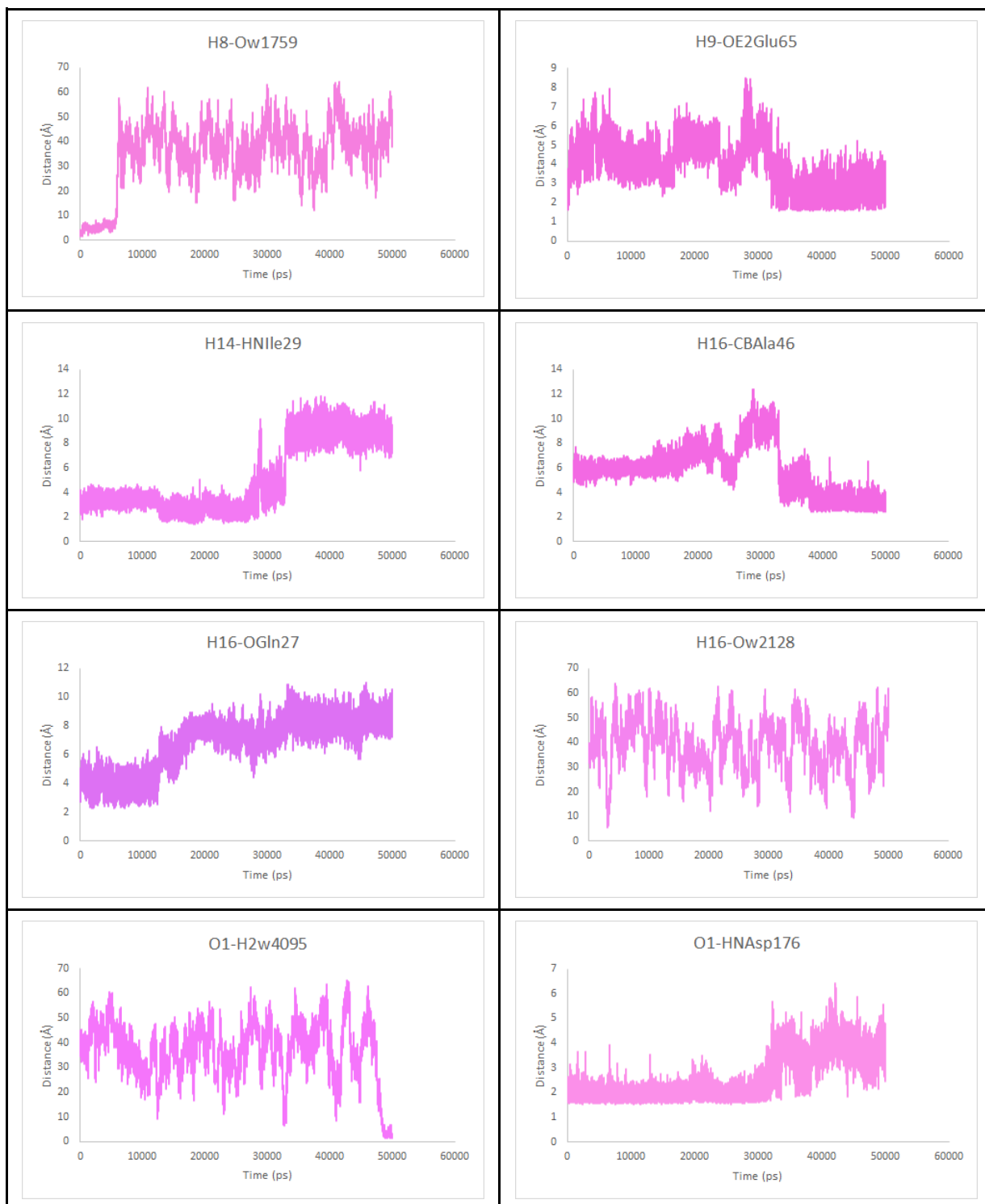
are represented as follows:

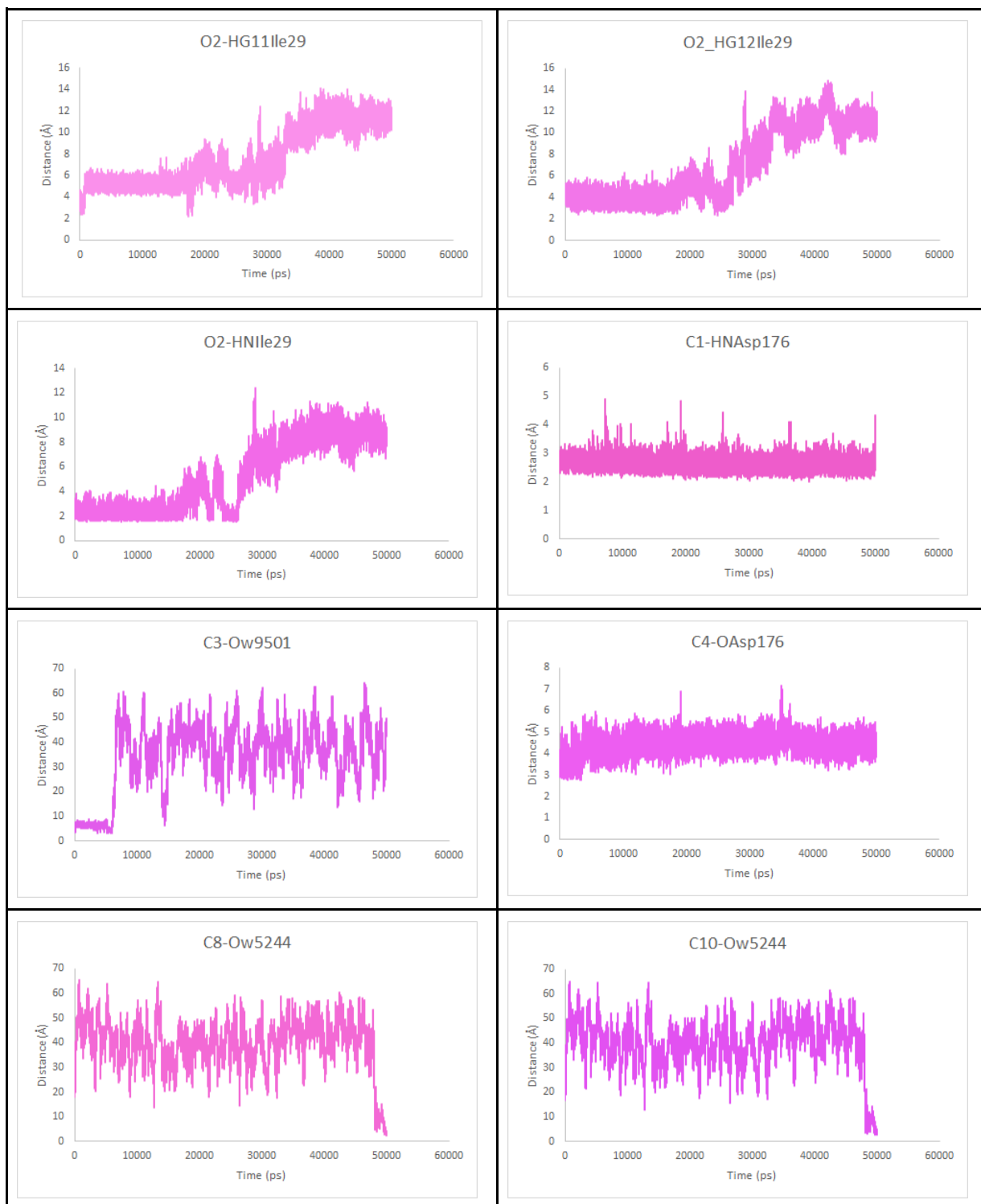
**Table 14:** Representation of the distances between interactions *INH (I6FP) ↔ aa*

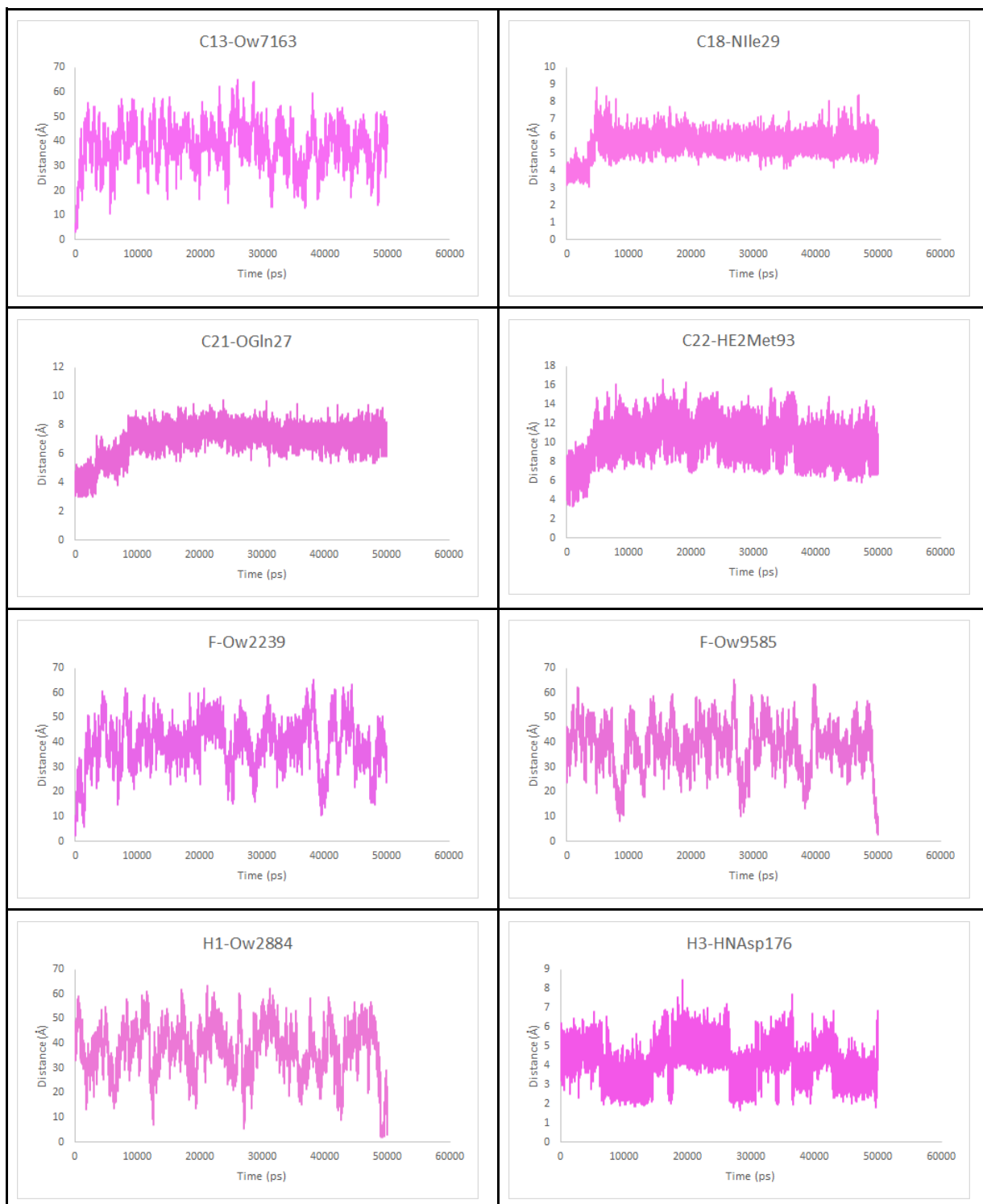


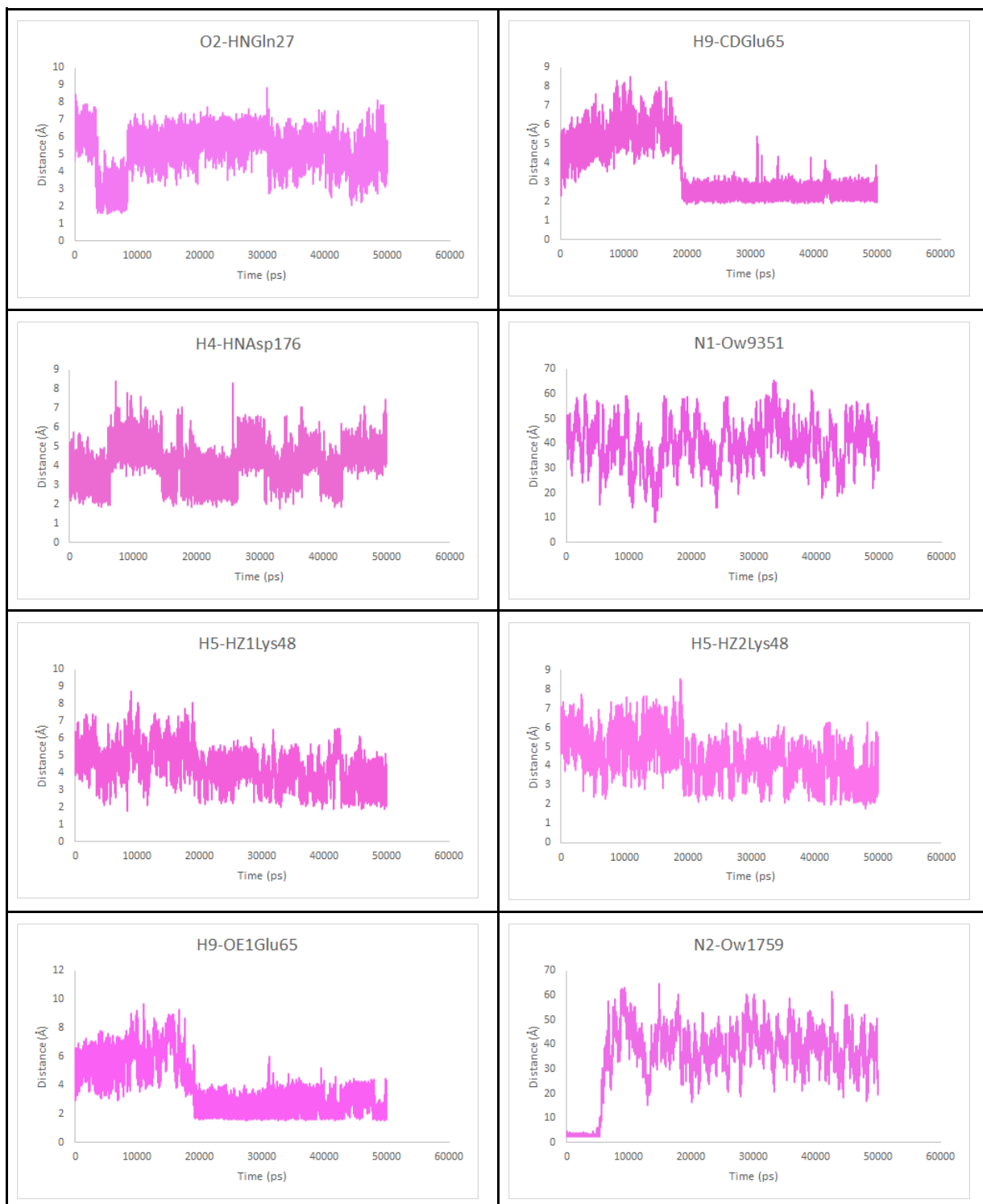


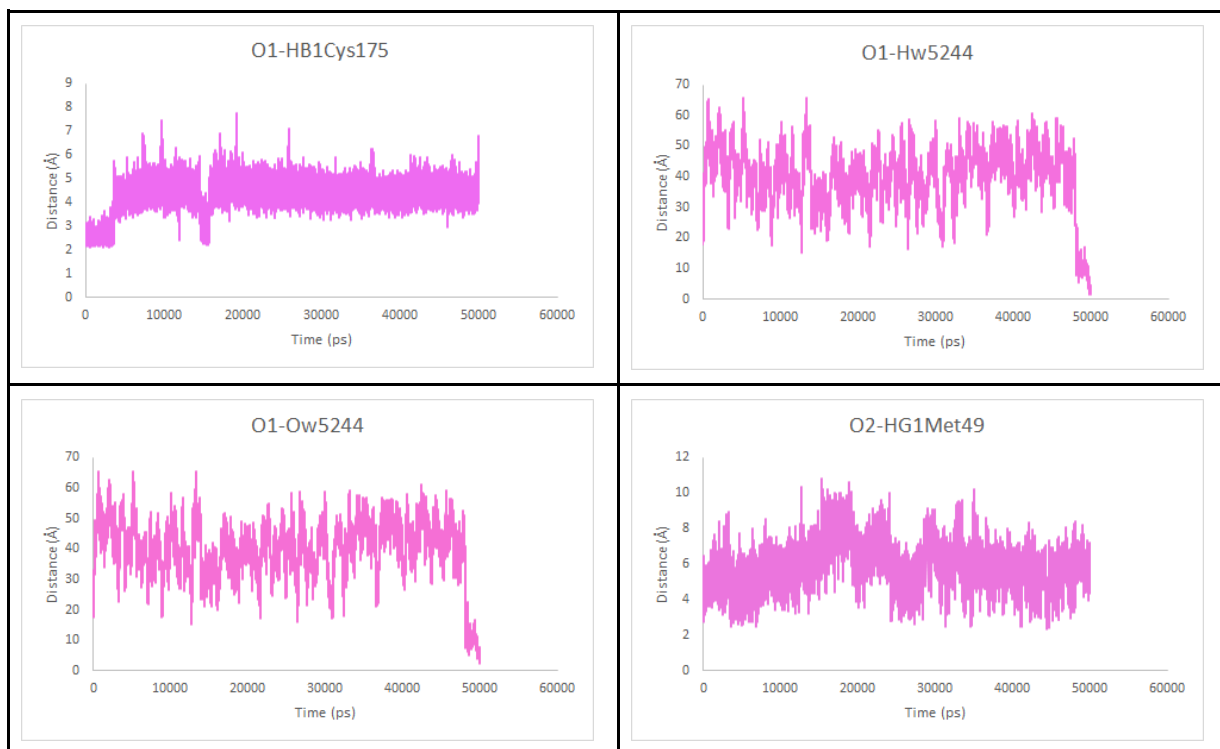




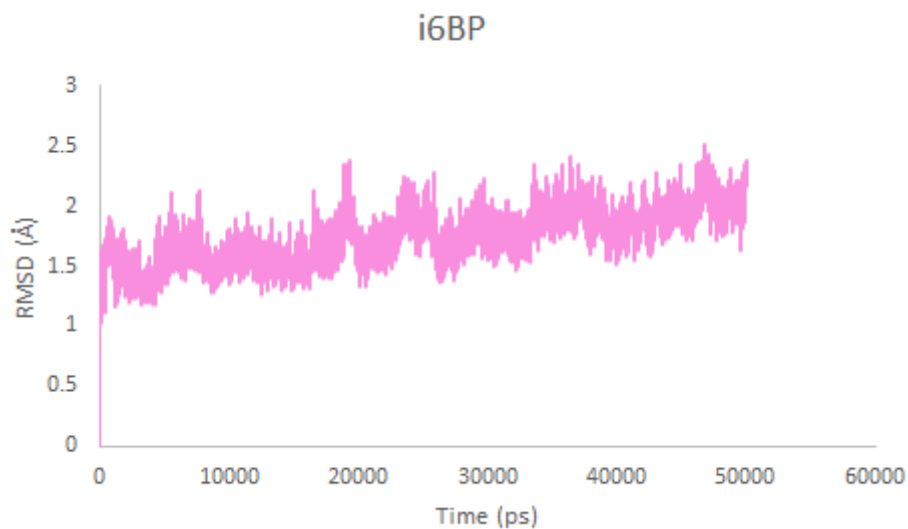








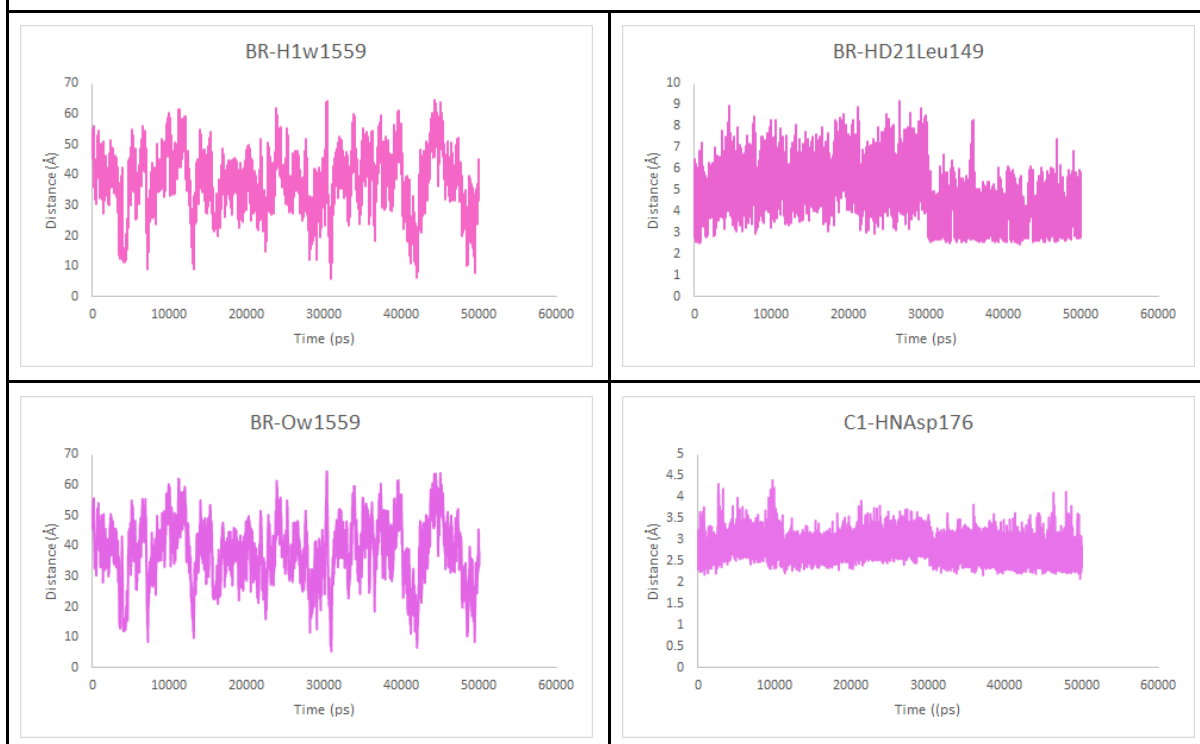
i6bp

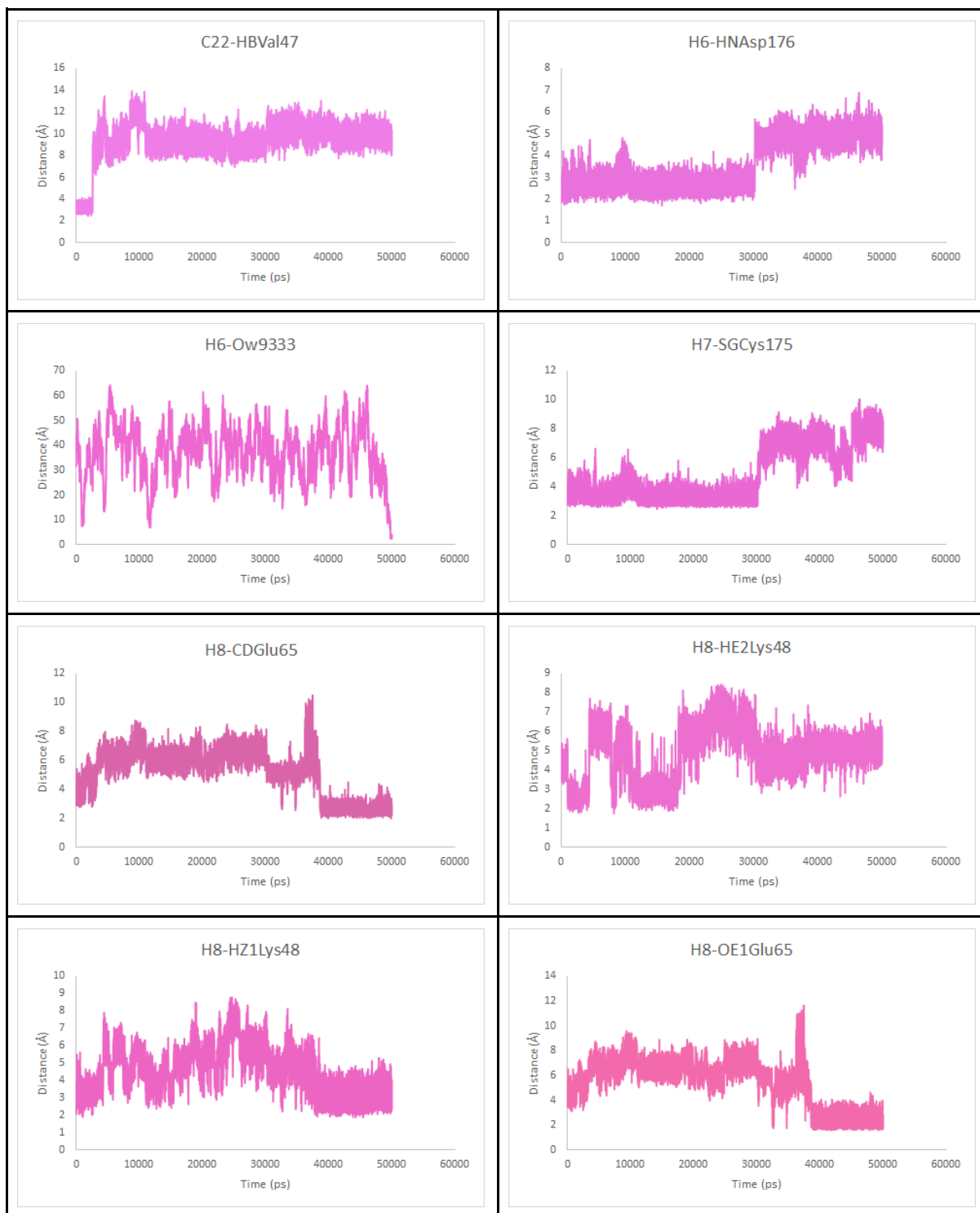


The trajectories described by each of the interactions observed by:

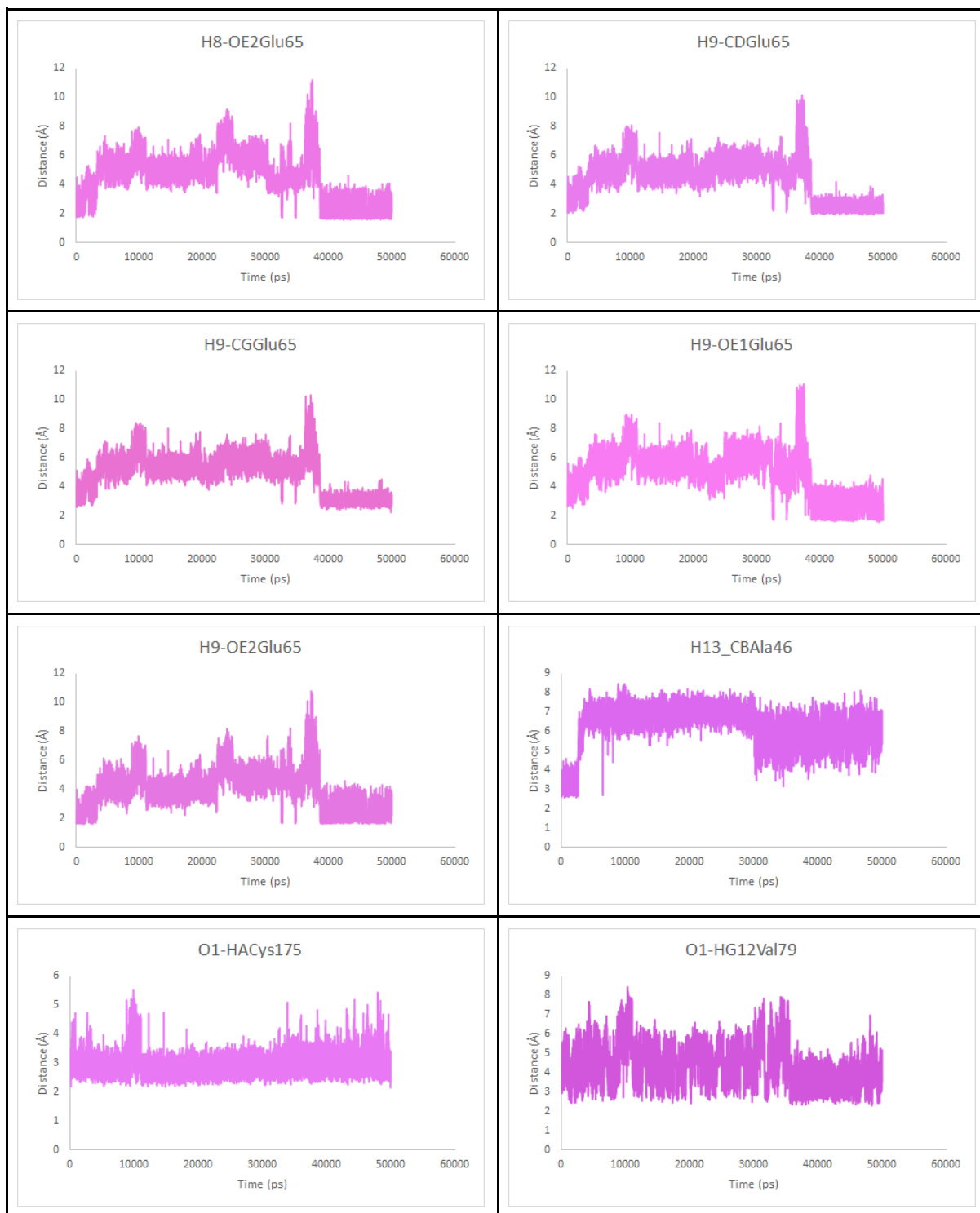
- The program "vmd" at 0ns and 50ns
- The files :
  - i6brp\_solute.avg.dat
  - solue\_i6brp.avg.dat

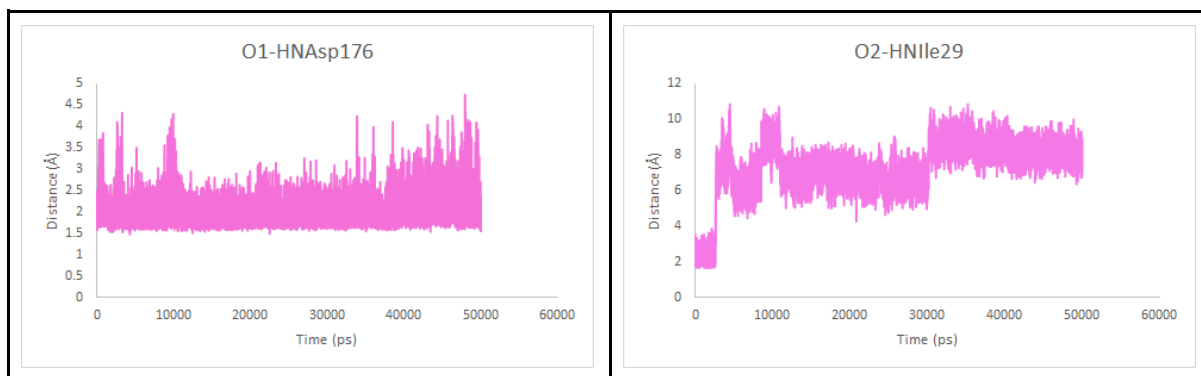
**Table 15:** Representation of the distances between interactions *INH (I6BP)* ↔ *aa*



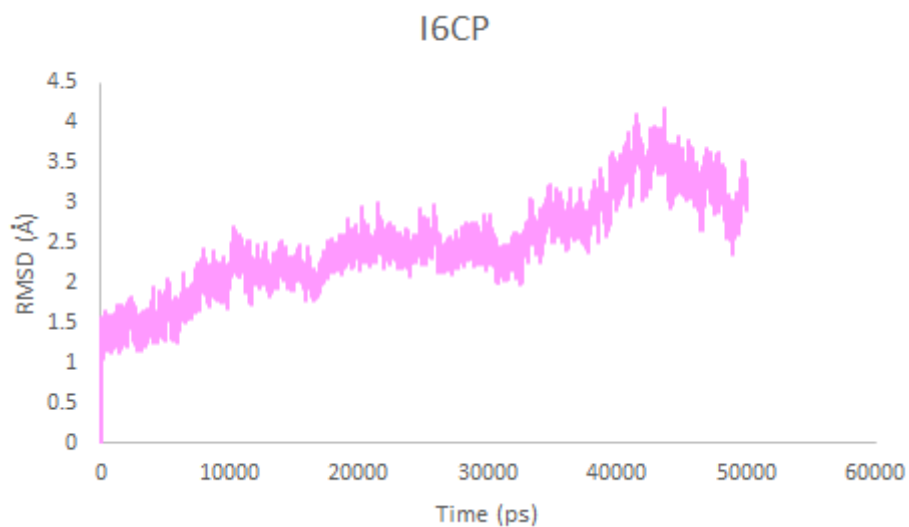








i6cp

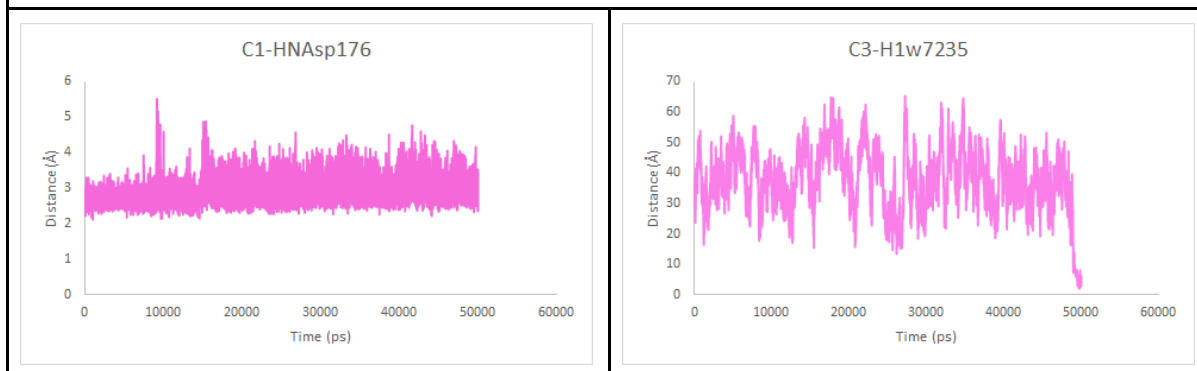


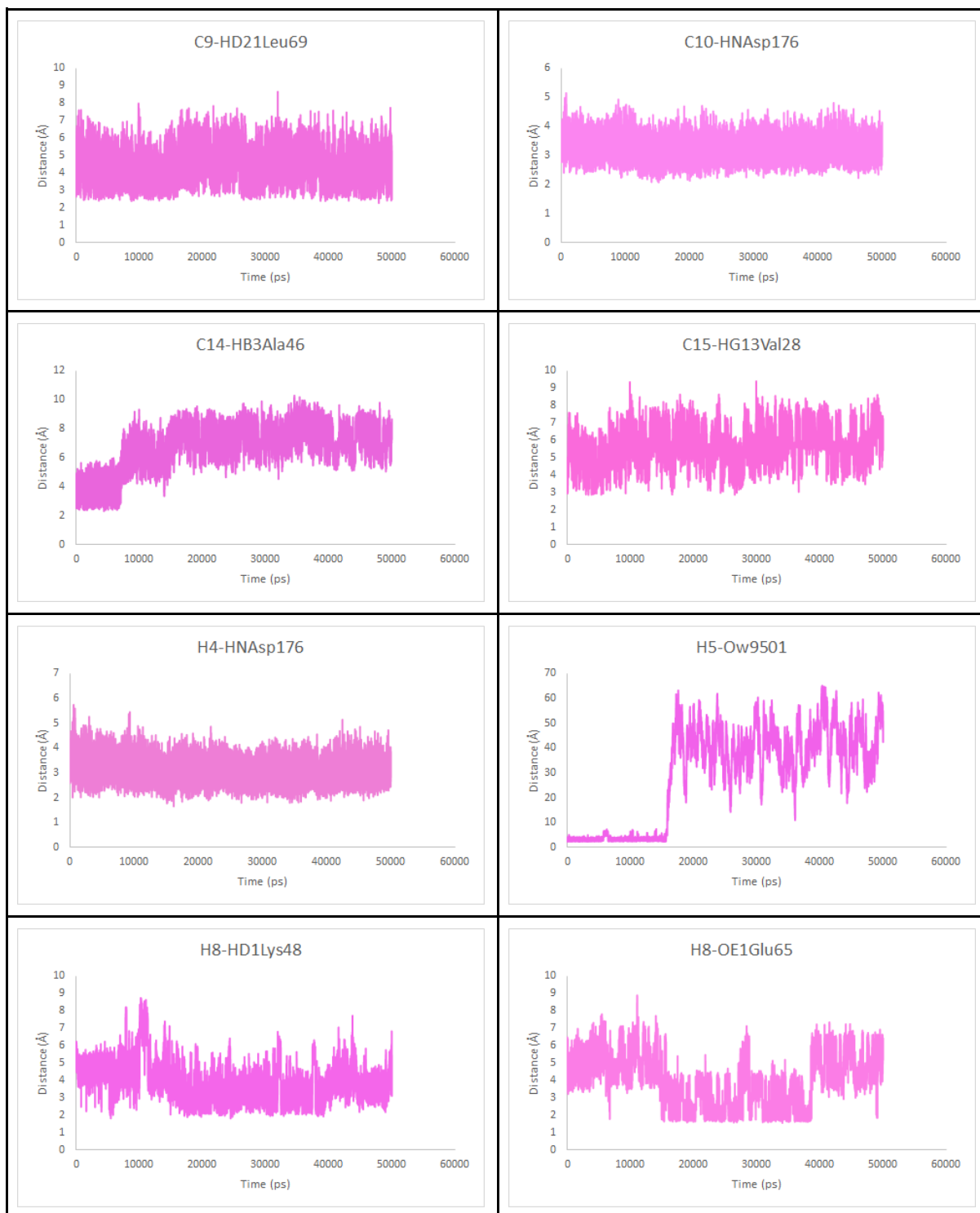
The trajectories described by each of the interactions observed by:

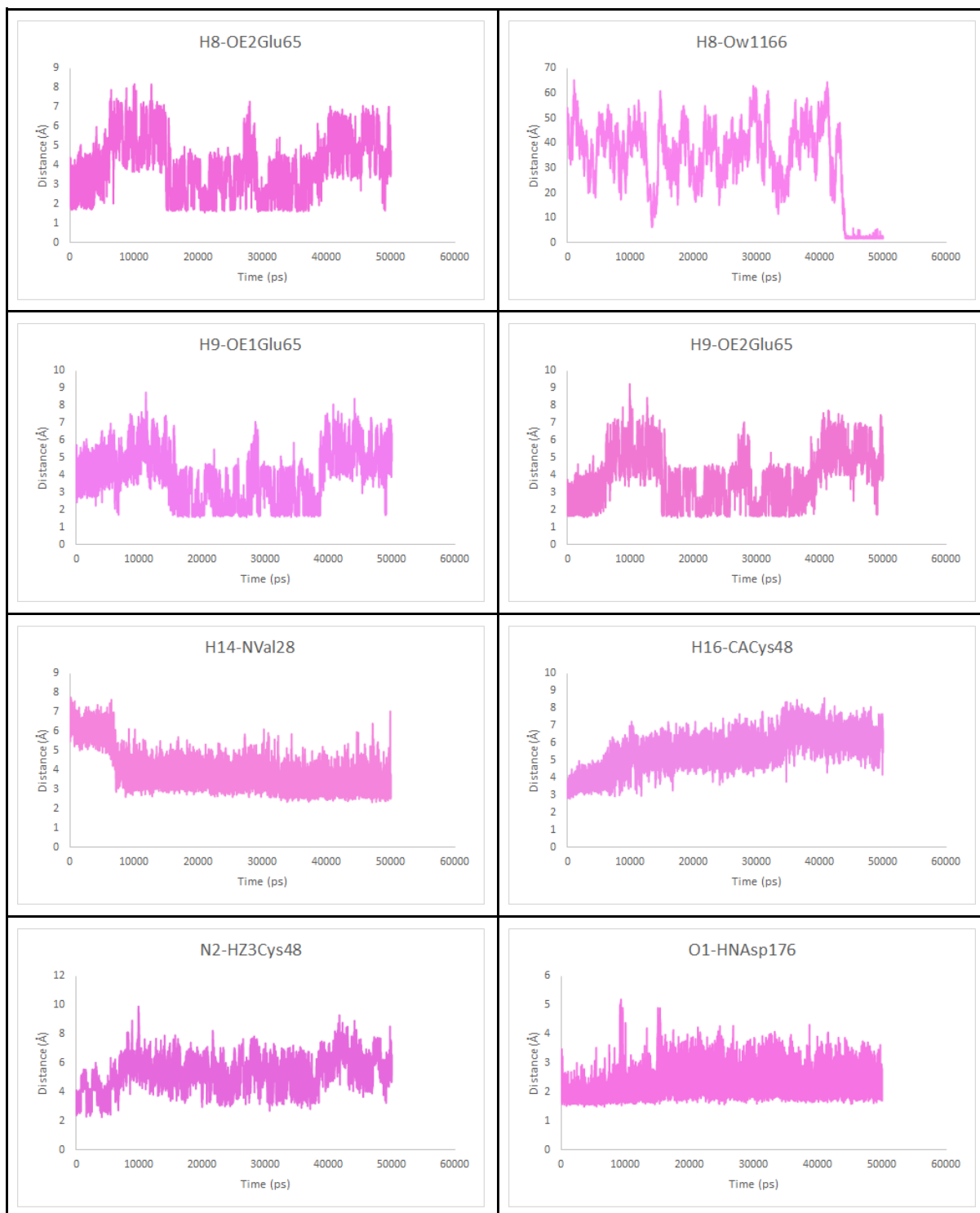
- The program "vmd" at 0ns and 50ns
- The files :
  - i6cp\_solute.avg.dat
  - solue\_i6cp.avg.dat

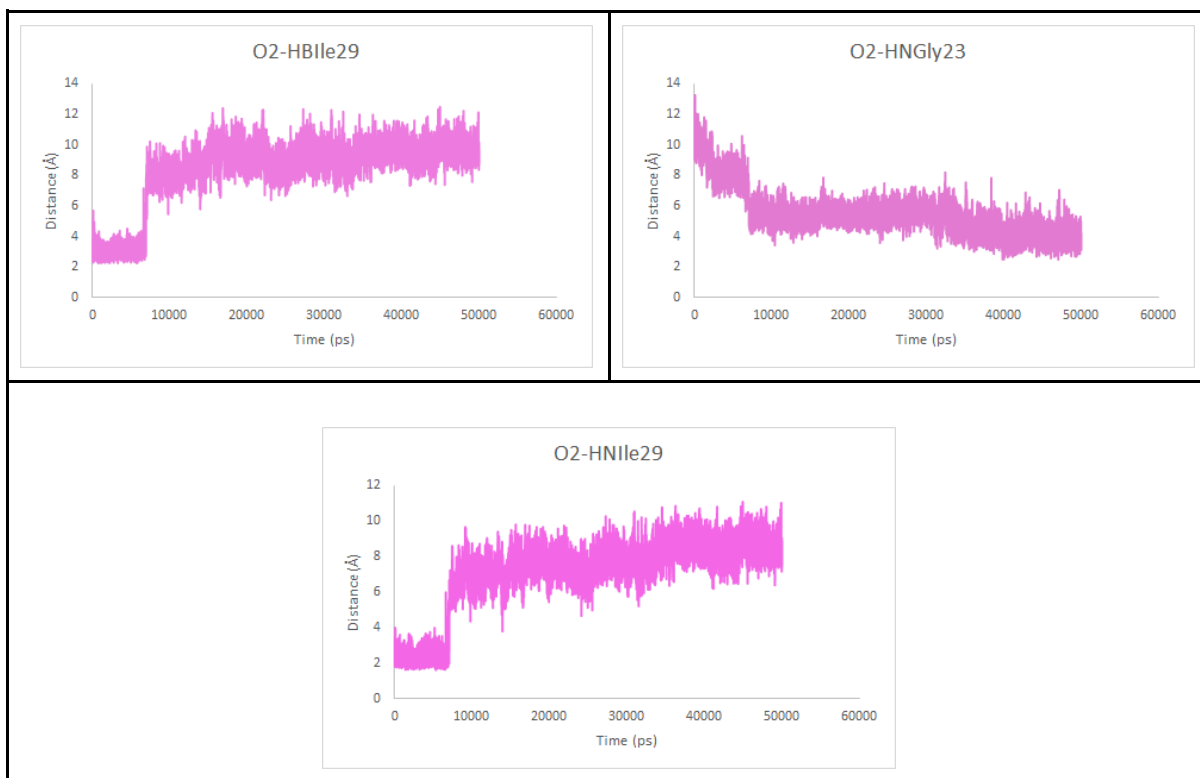
are represented as follows:

**Tabla 16:** Representation of the distances between interactions *INH (I6CP) ↔ aa*

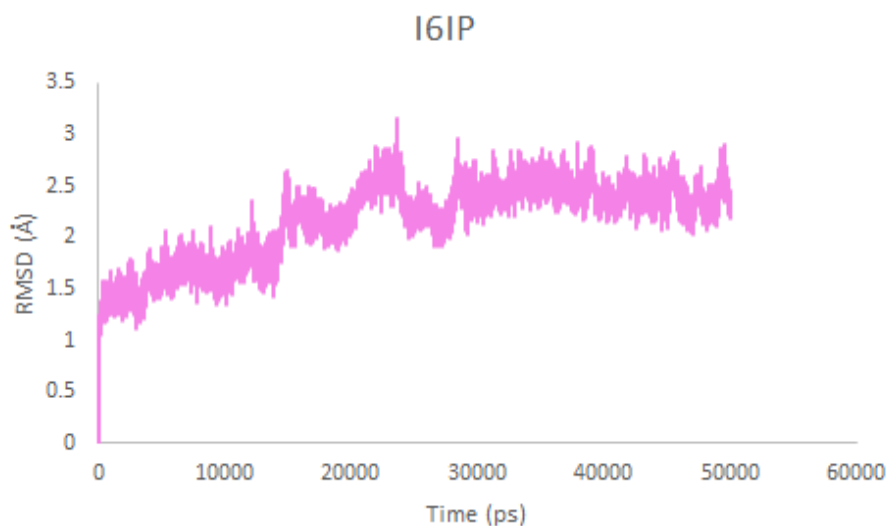








i6ip

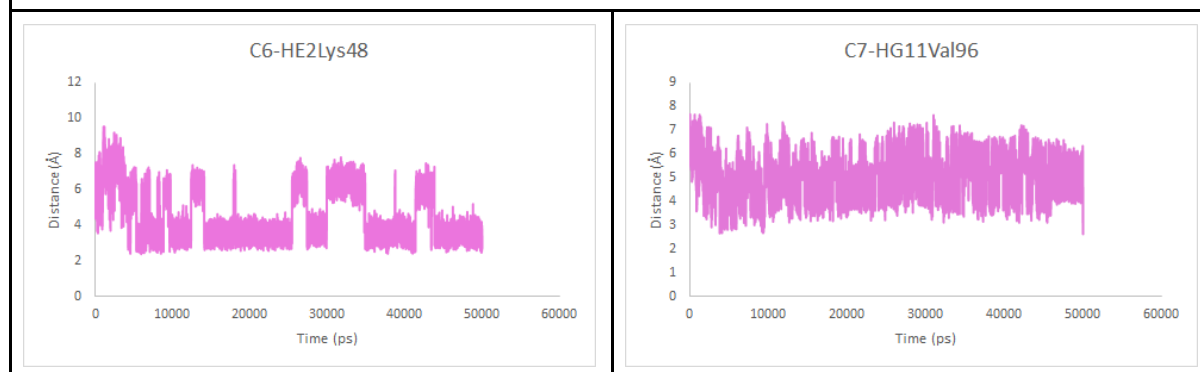


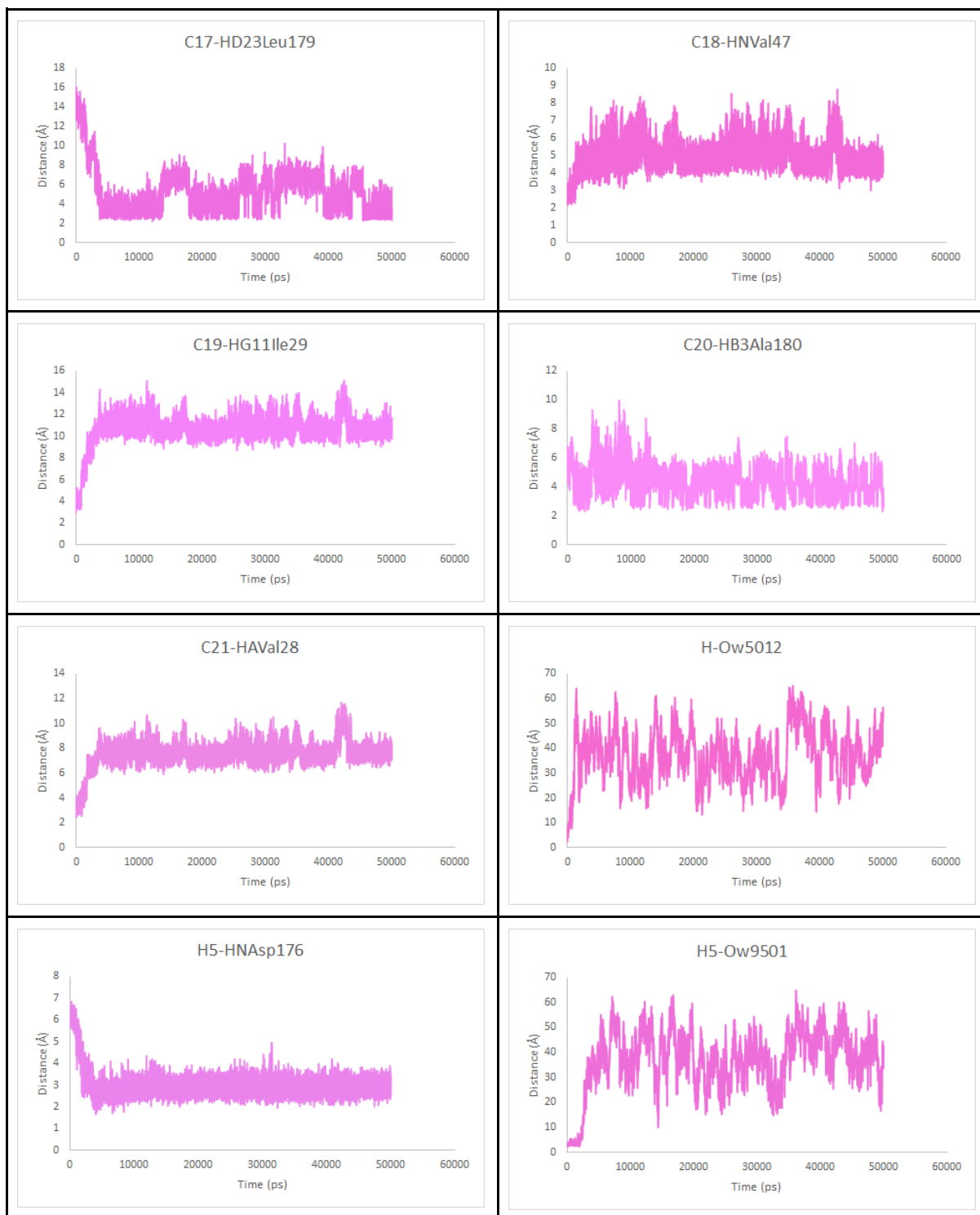
The trajectories described by each of the interactions observed by:

- The program "vmd" at 0ns and 50ns
- The files :
  - i6ip\_solute.avg.dat
  - solue\_i6ip.avg.dat

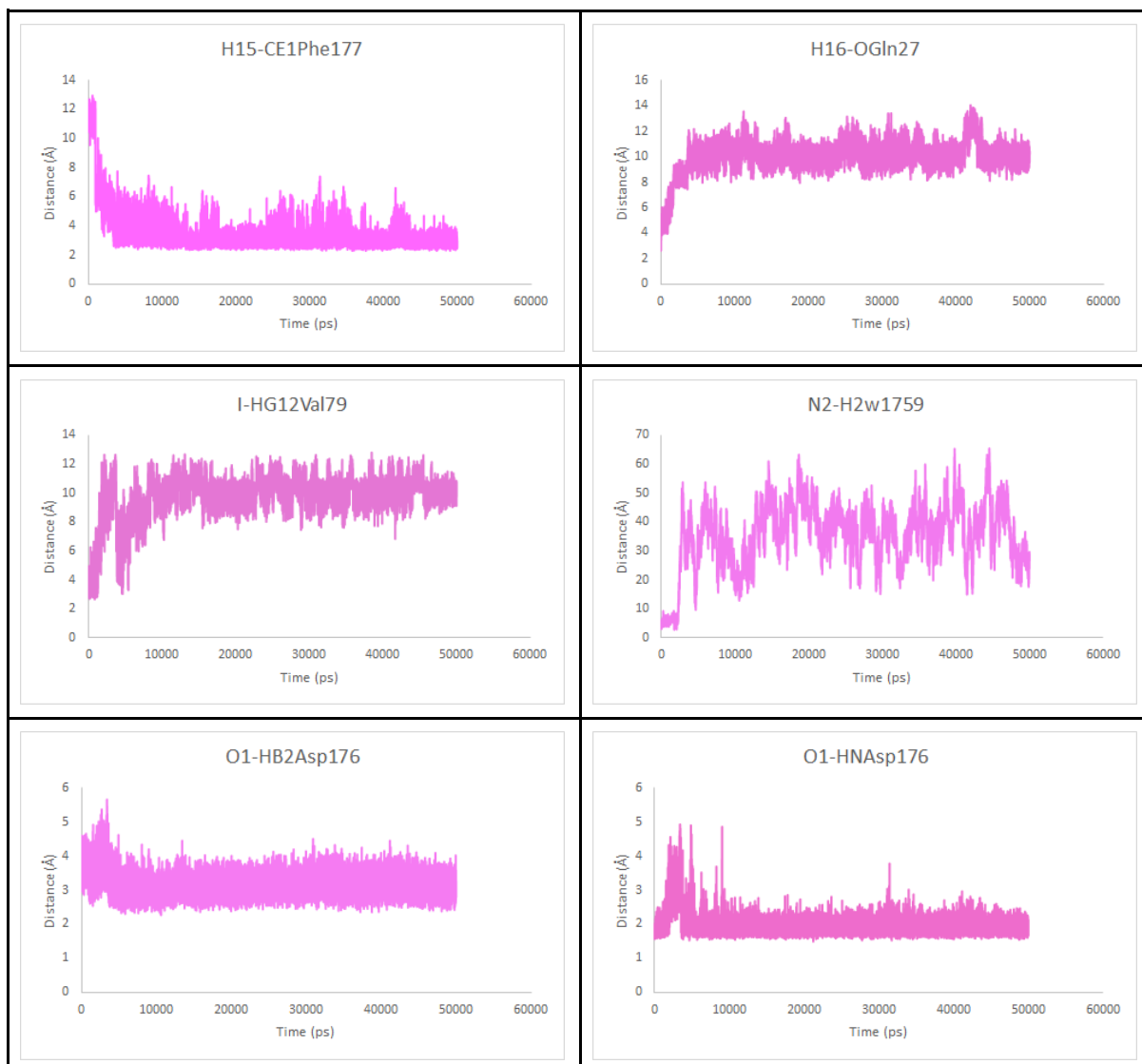
are represented as follows:

**Table 17:** Representation of the distances between interactions *INH (I6IP) ↔ aa*



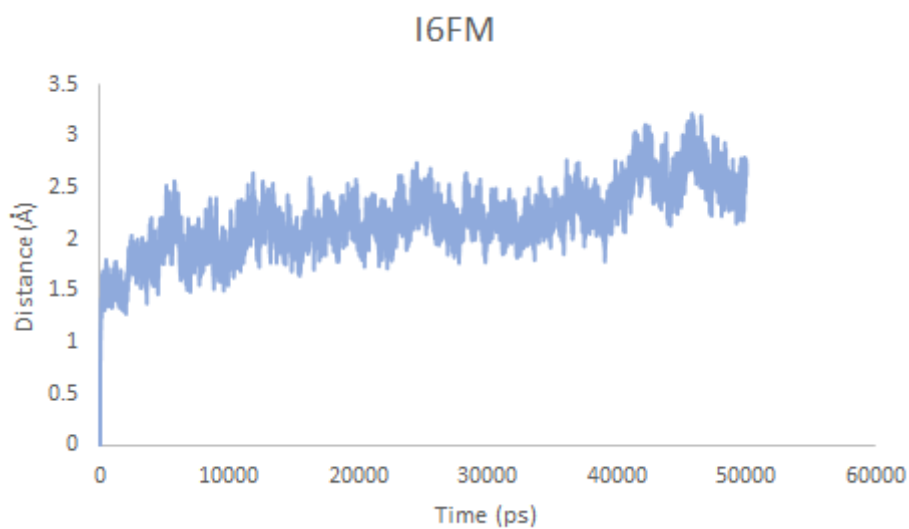






META

i6fm

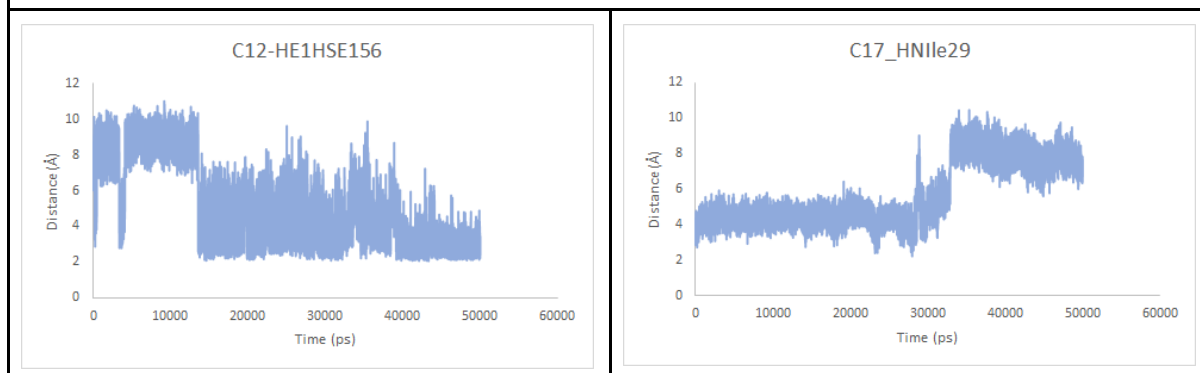


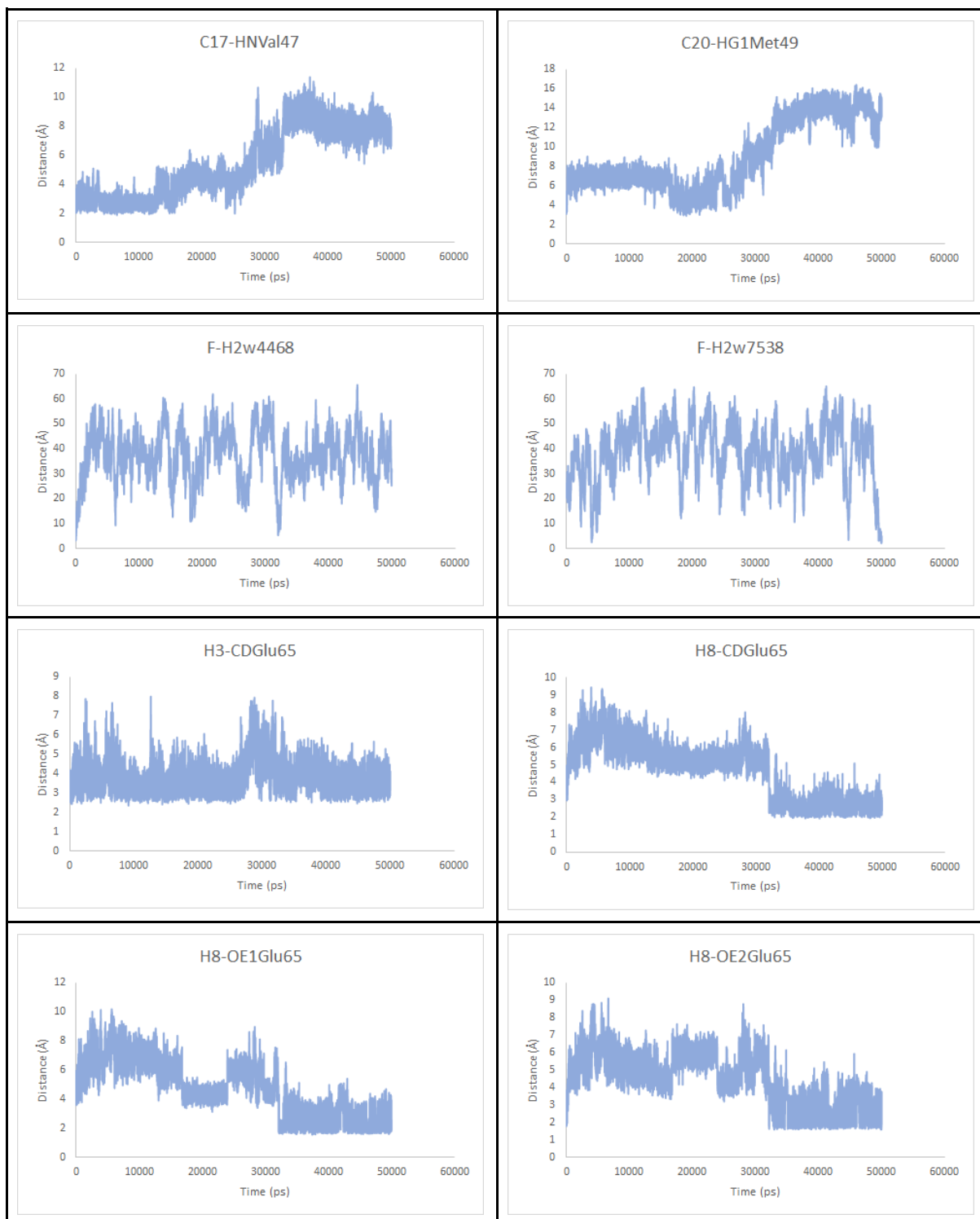
The trajectories described by each of the interactions observed by:

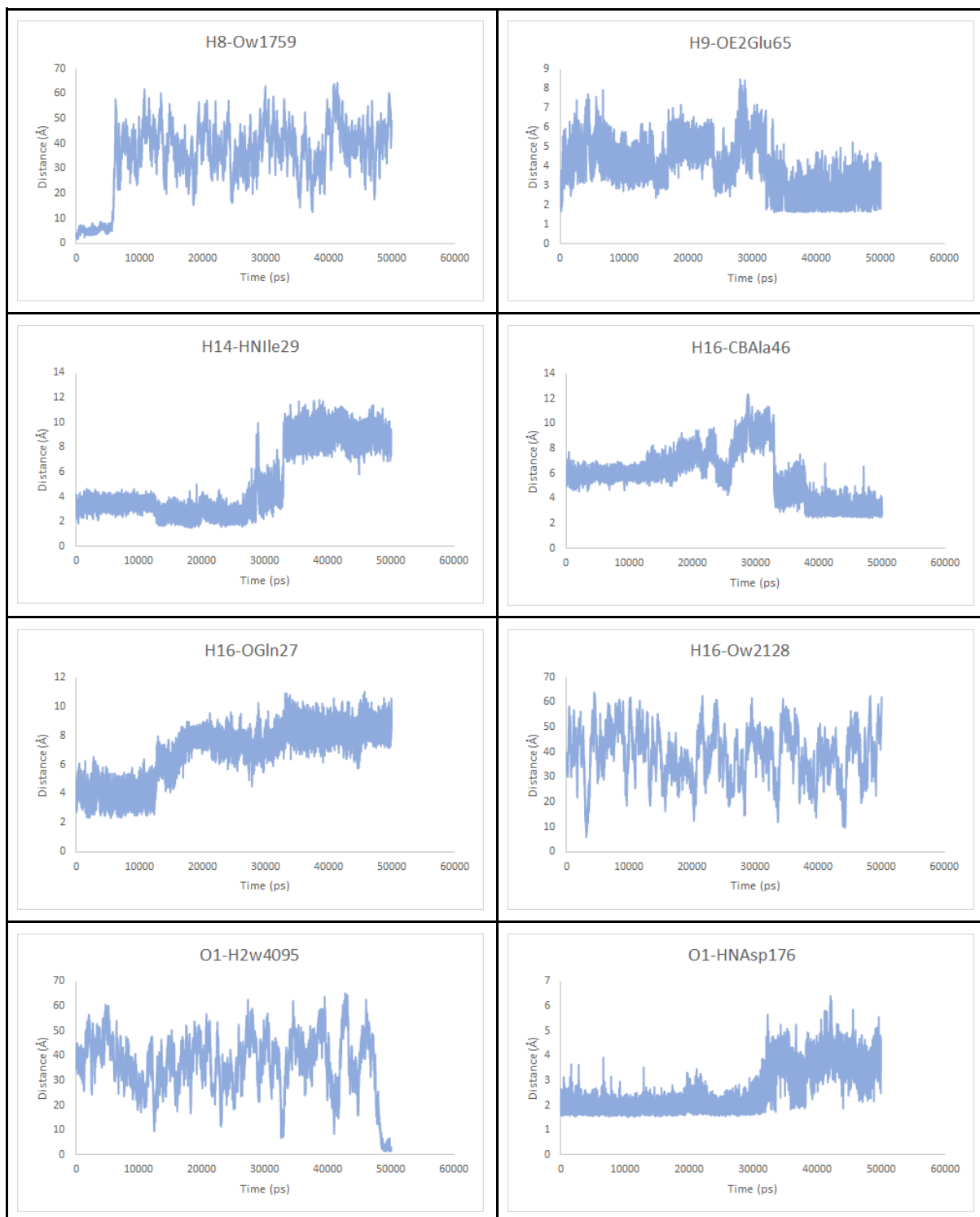
- The program "vmd" at 0ns and 50ns
- The files :
  - i6fm\_solute.avg.dat
  - solue\_i6fm.avg.dat

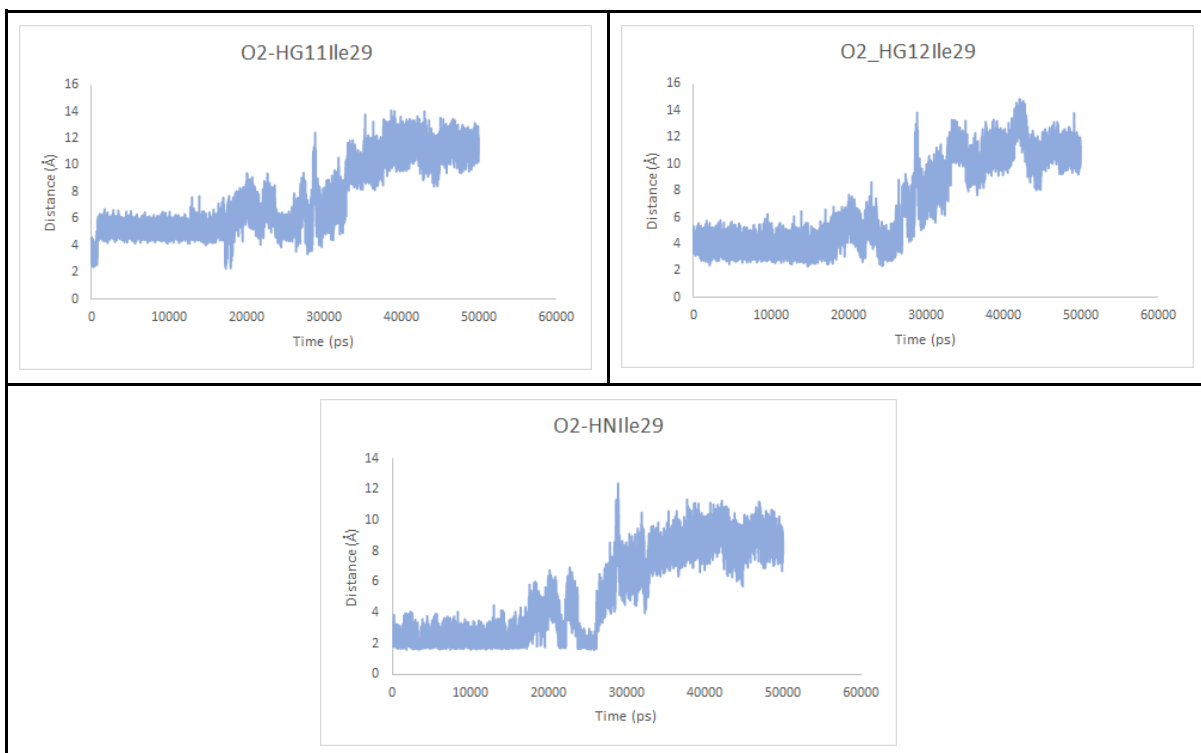
are represented as follows:

**Tabla 18:** Representation of the distances between interactions *INH (I6FM) ↔ aa*

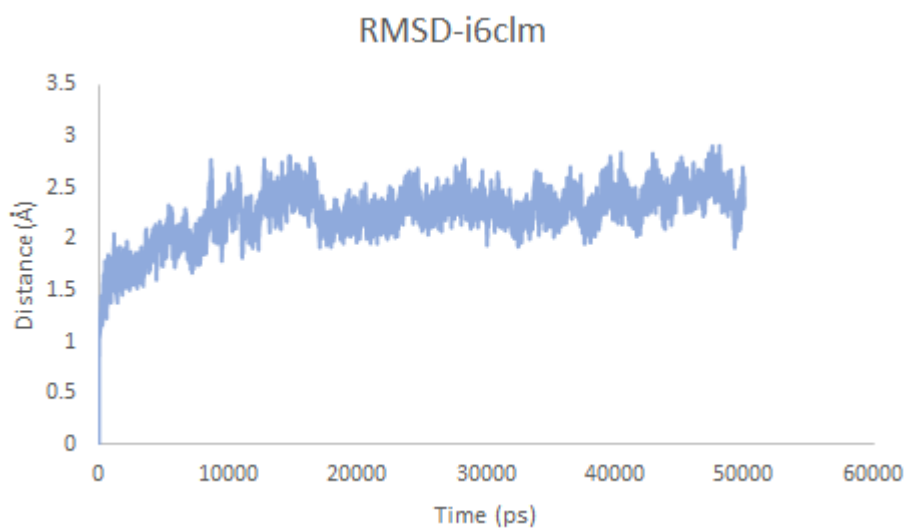








i6clm

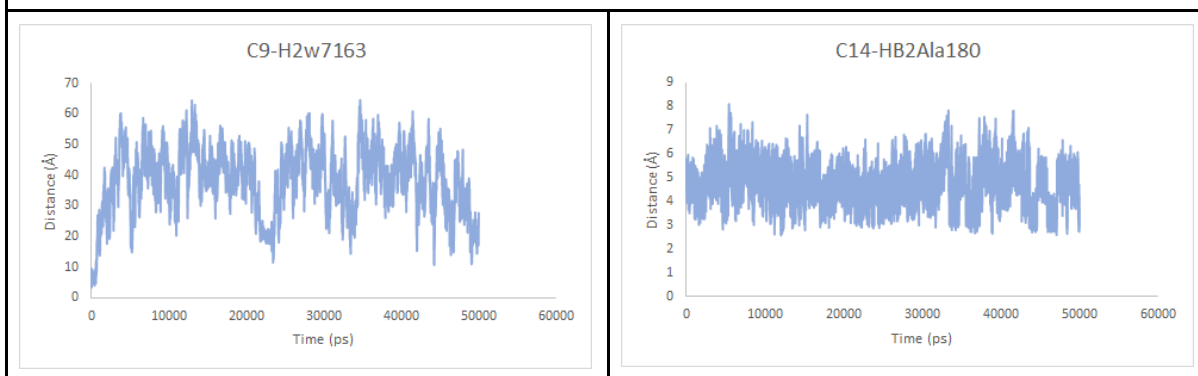


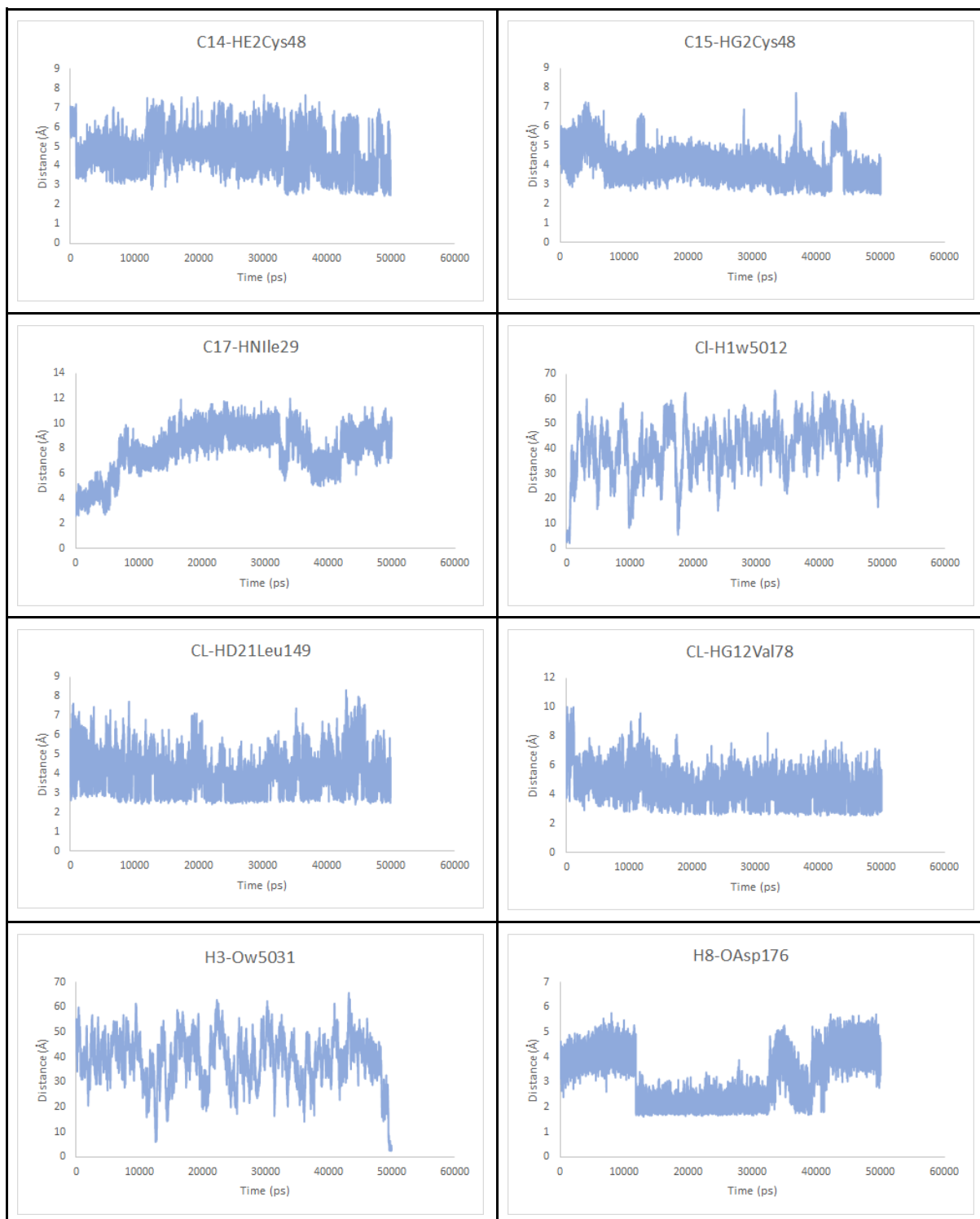
The trajectories described by each of the interactions observed by:

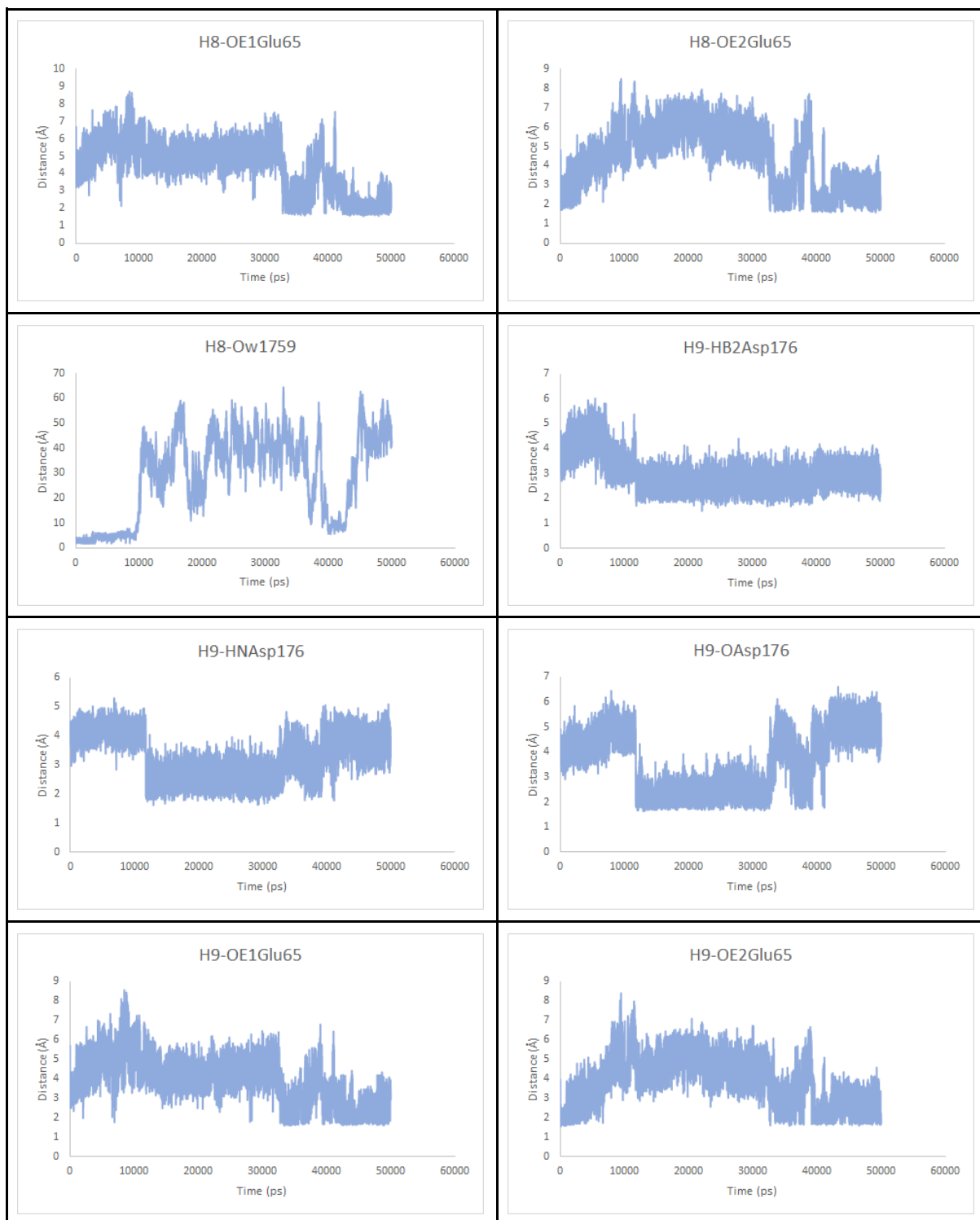
- The program "vmd" at 0ns and 50ns
- The files :
  - i6clm\_solute.avg.dat
  - solute\_i6clm.avg.dat

are represented as follows:

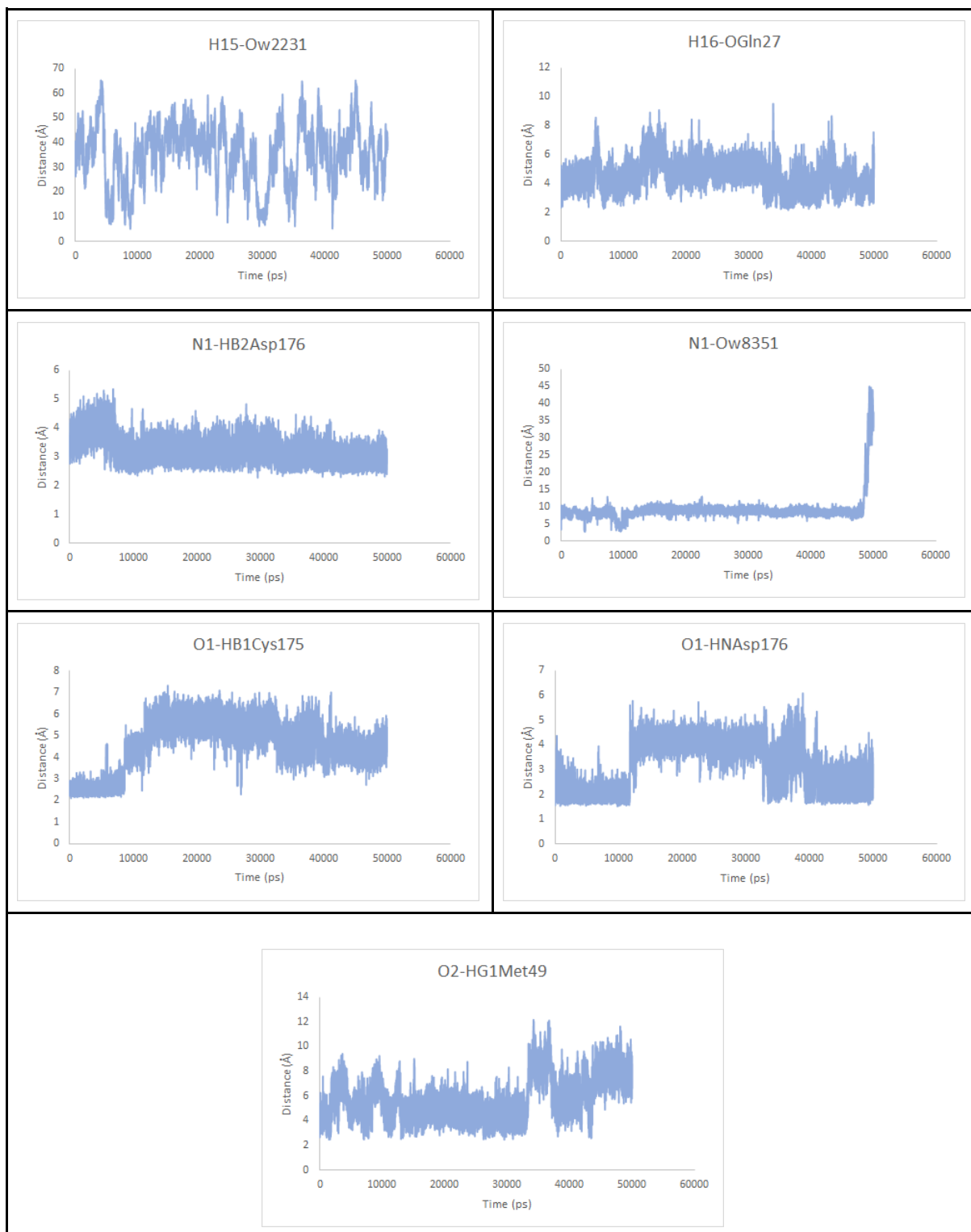
**Tabla 19:** Representation of the distances between interactions *INH (I6CM) ↔ aa*



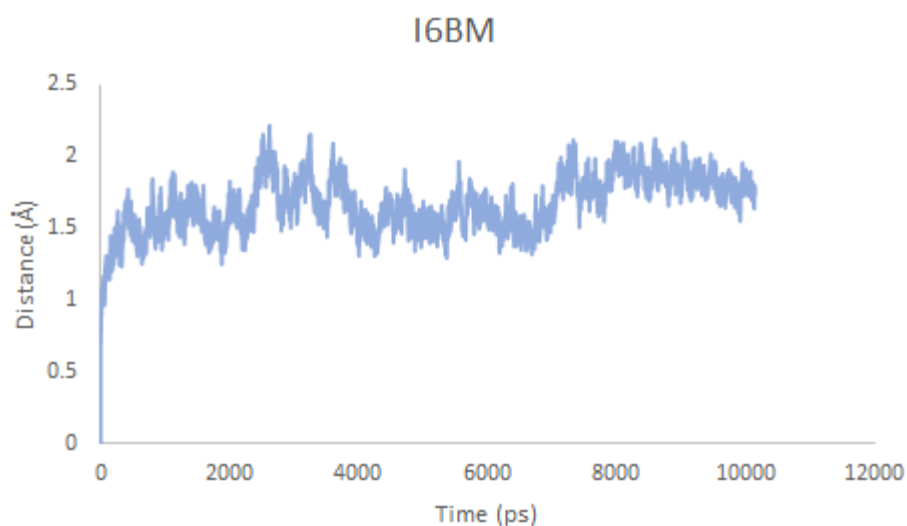








i6brm

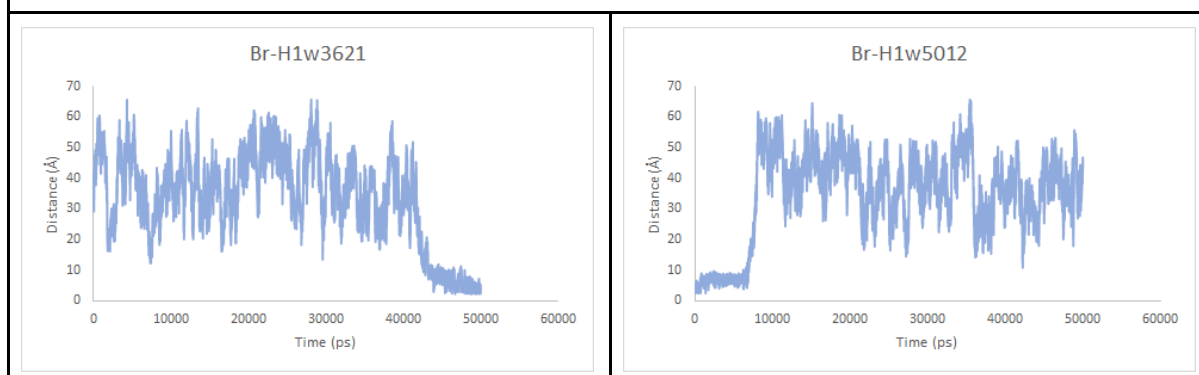


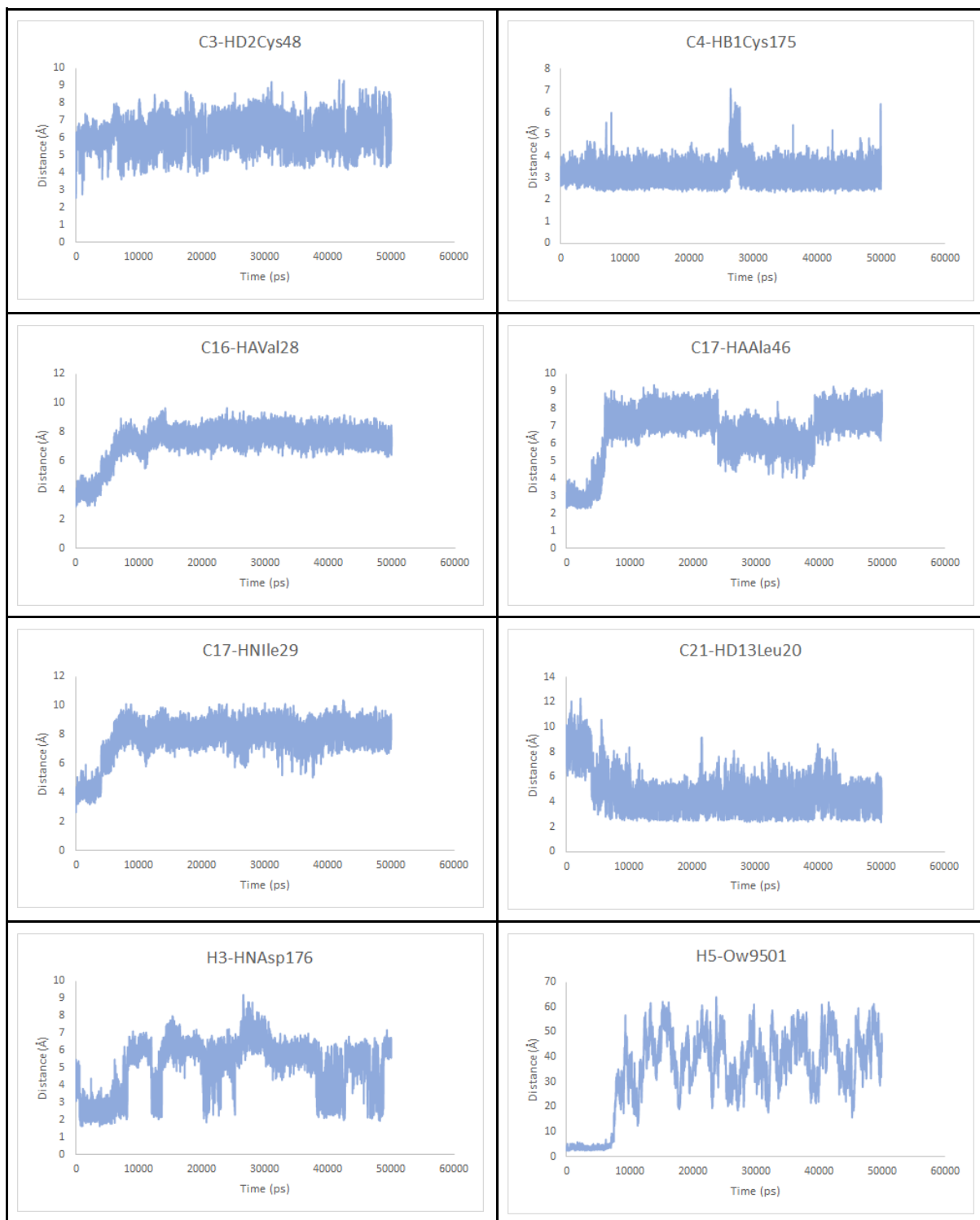
The trajectories described by each of the interactions observed by:

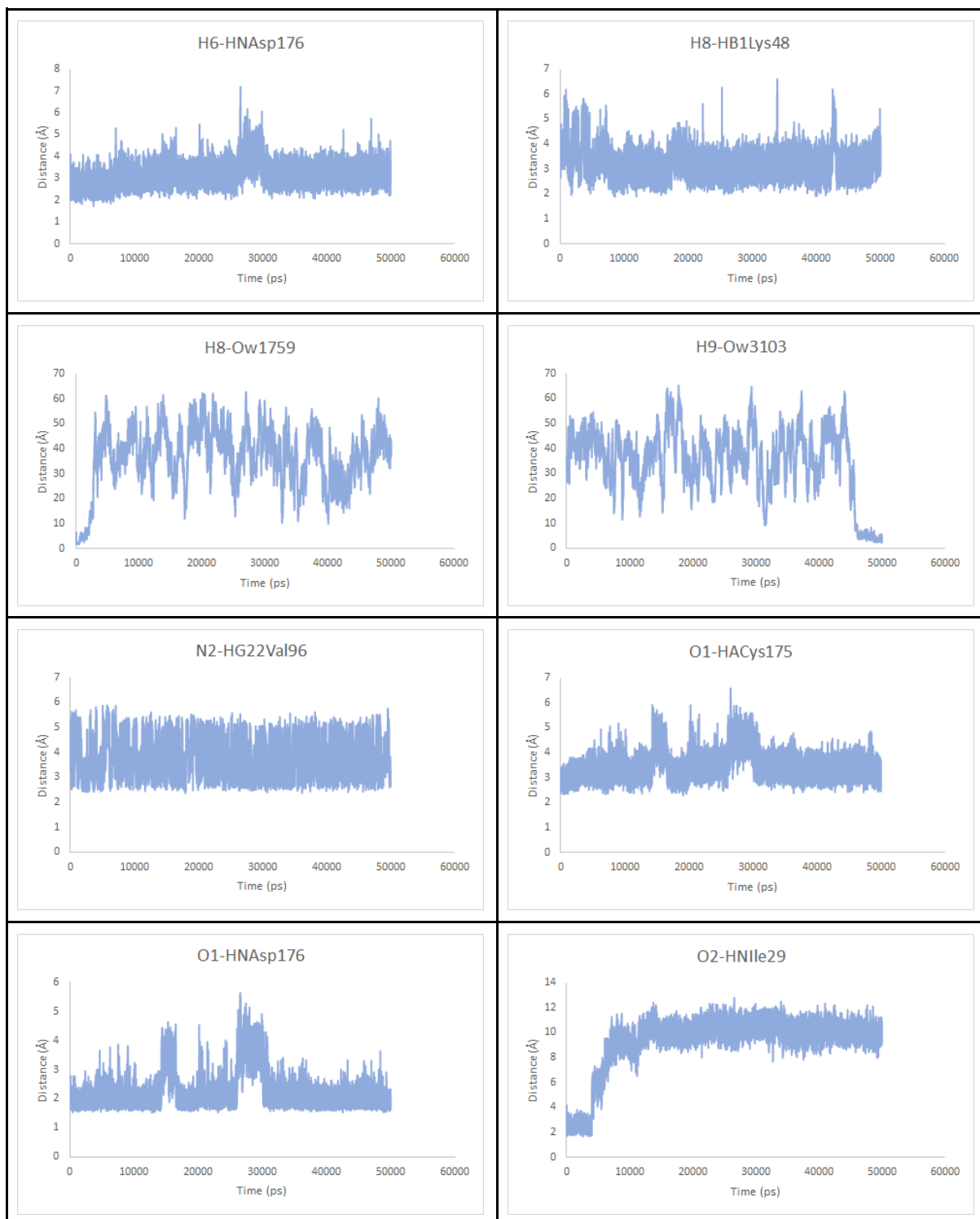
- The program "vmd" at 0ns and 50ns
- The files :
  - i6brm\_solute.avg.dat
  - solue\_i6brm.avg.dat

are represented as follows:

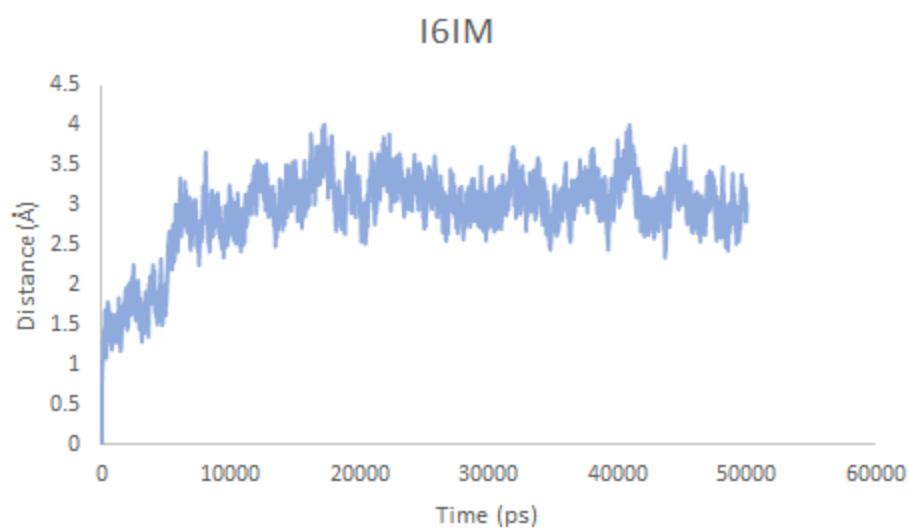
**Tabla 20:** Representation of the distances between interactions *INH (I6BM) ↔ aa*







i6im

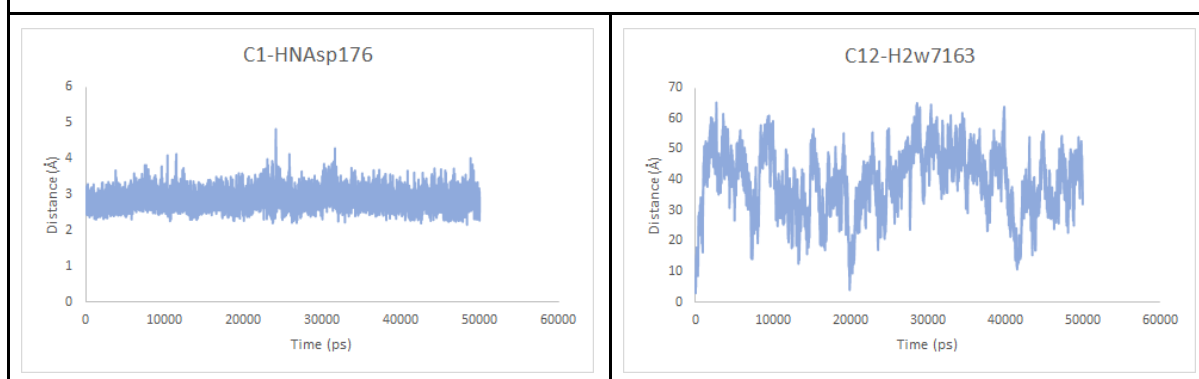


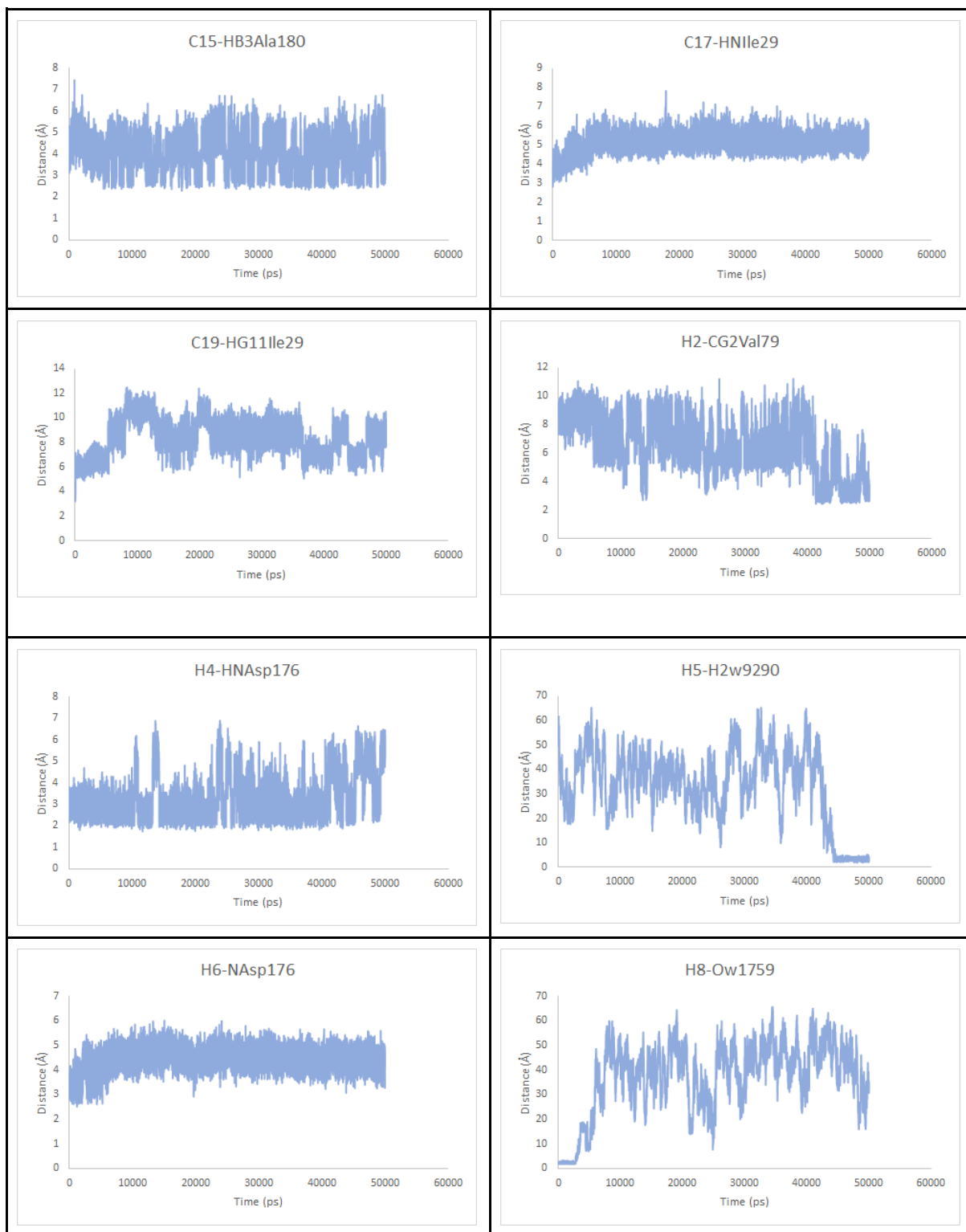
The trajectories described by each of the interactions observed by:

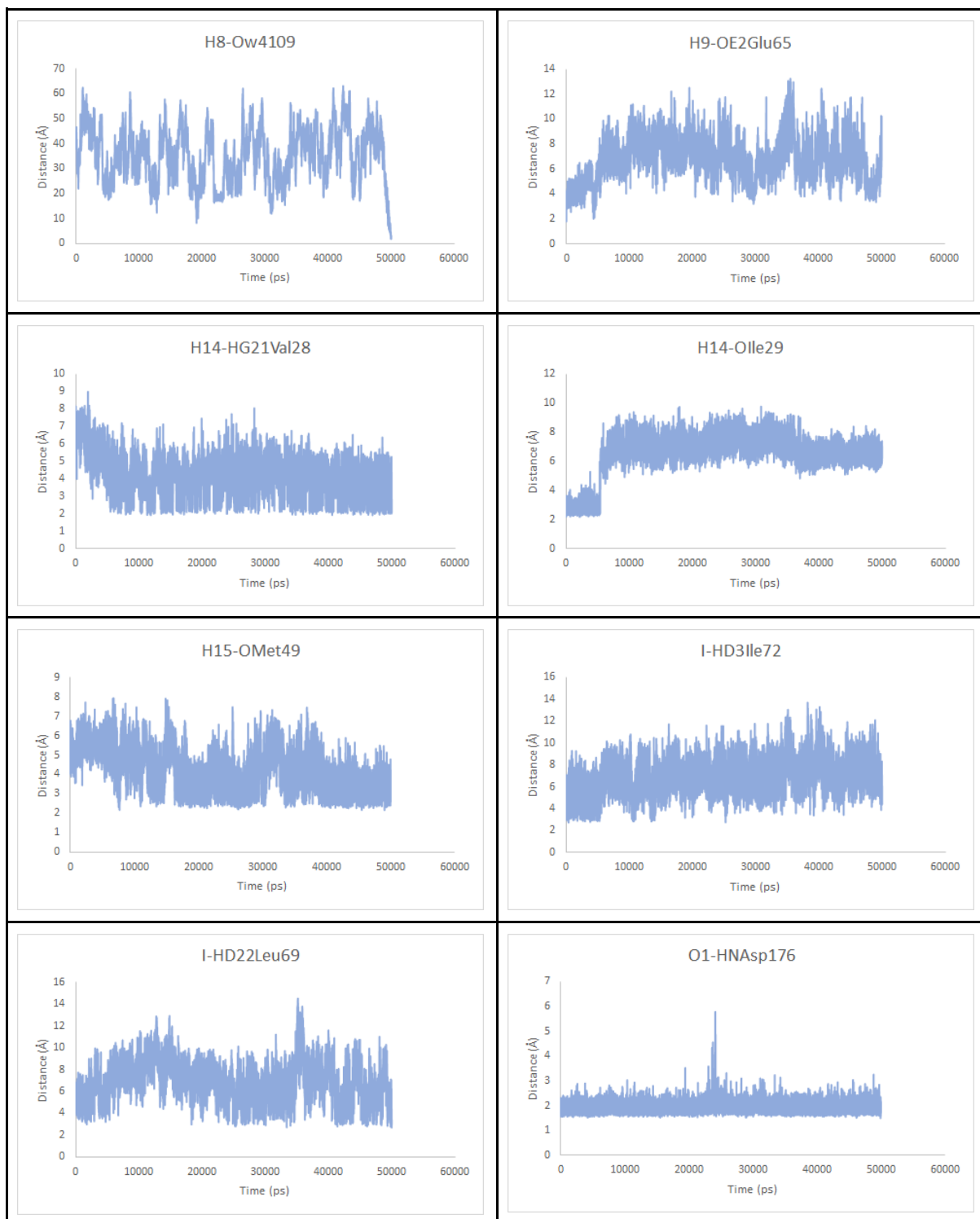
- The program "vmd" at 0ns and 50ns
- The files :
  - i6im\_solute.avg.dat
  - solue\_i6im.avg.dat

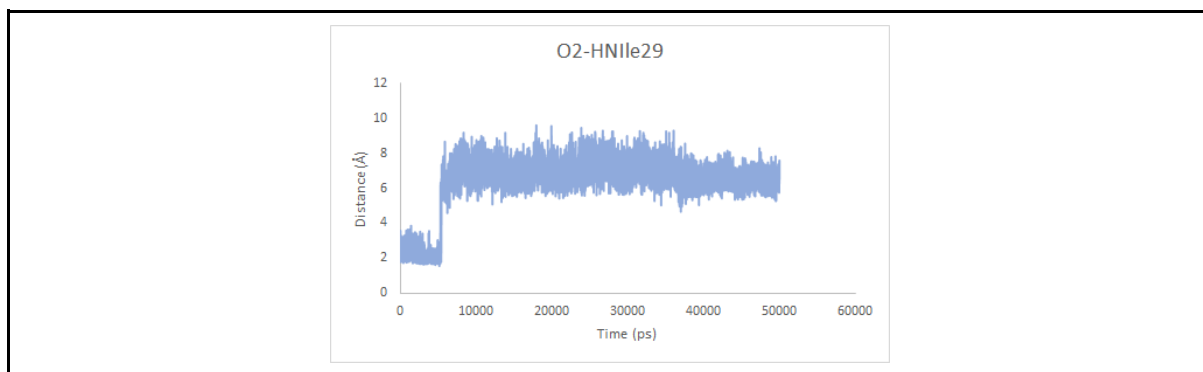
are represented as follows:

**Tabla 21:** Representation of the distances between interaction *INH (I6IM)* ↔ *aa*





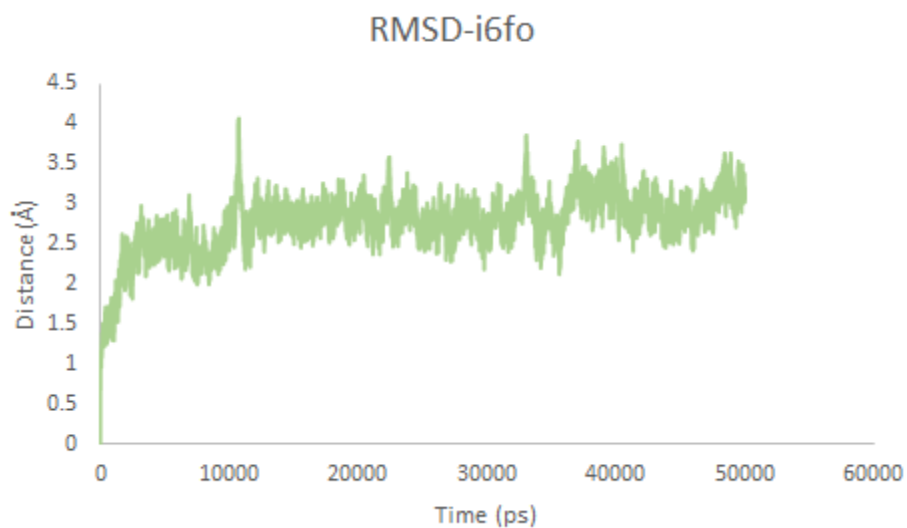






ORTO

i6fo

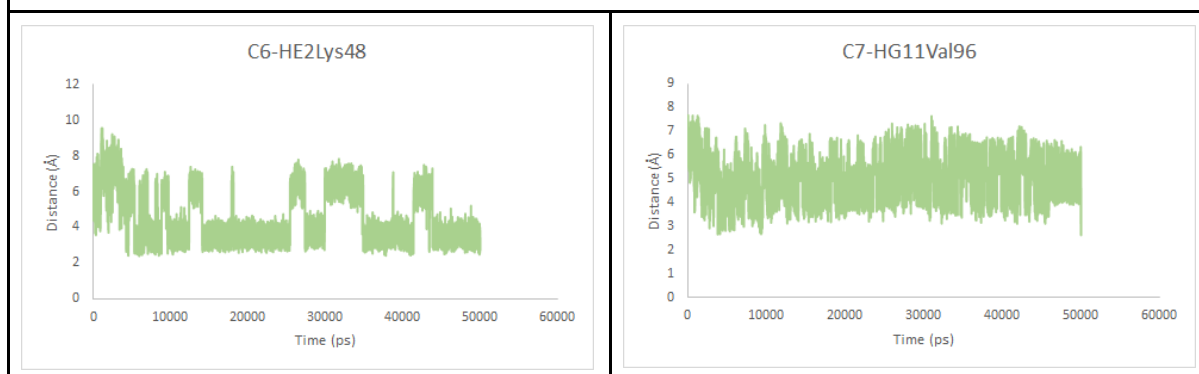


The trajectories described by each of the interactions observed by:

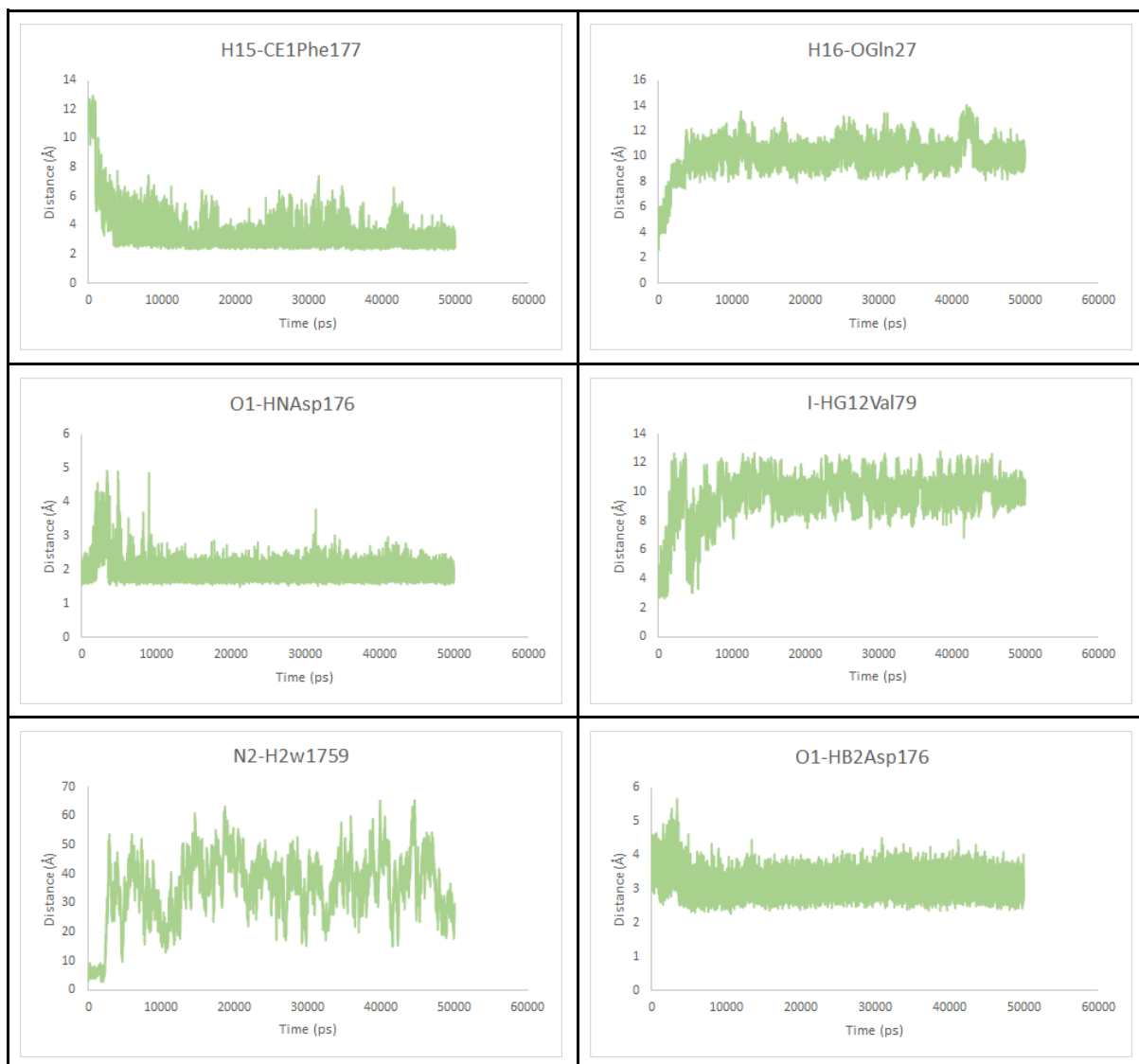
- The program "vmd" at 0ns and 50ns
- The files :  
    i6fo\_solute.avg.dat  
    solue\_i6fo.avg.dat

are represented as follows:

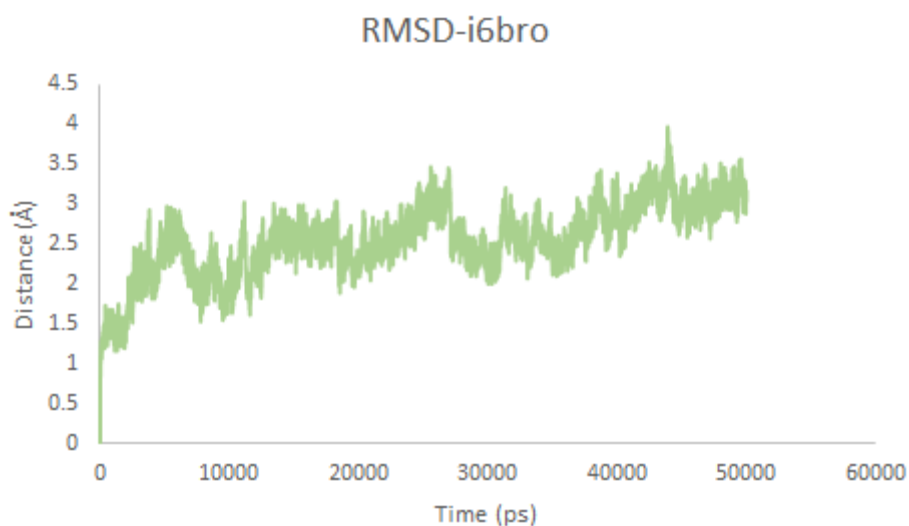
**Tabla 22:** Representation of the distances between interaction *INH (I6FO)* ↔ *aa*







i6bo

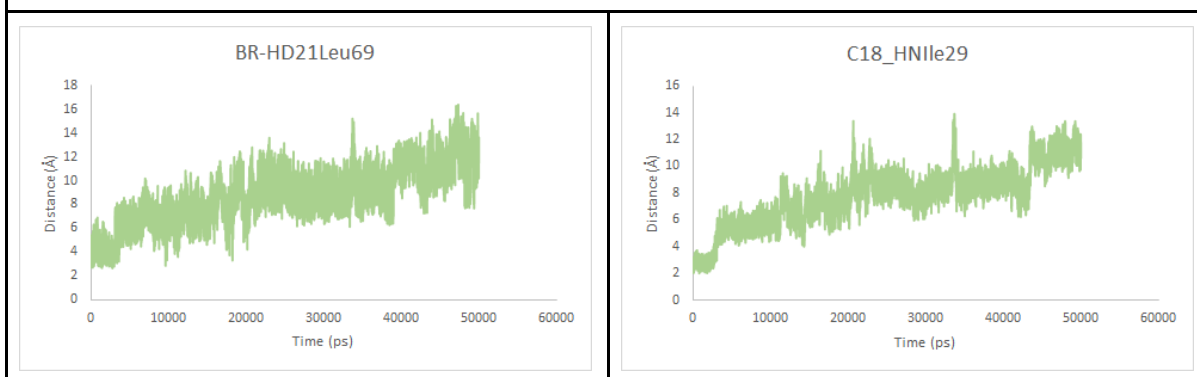


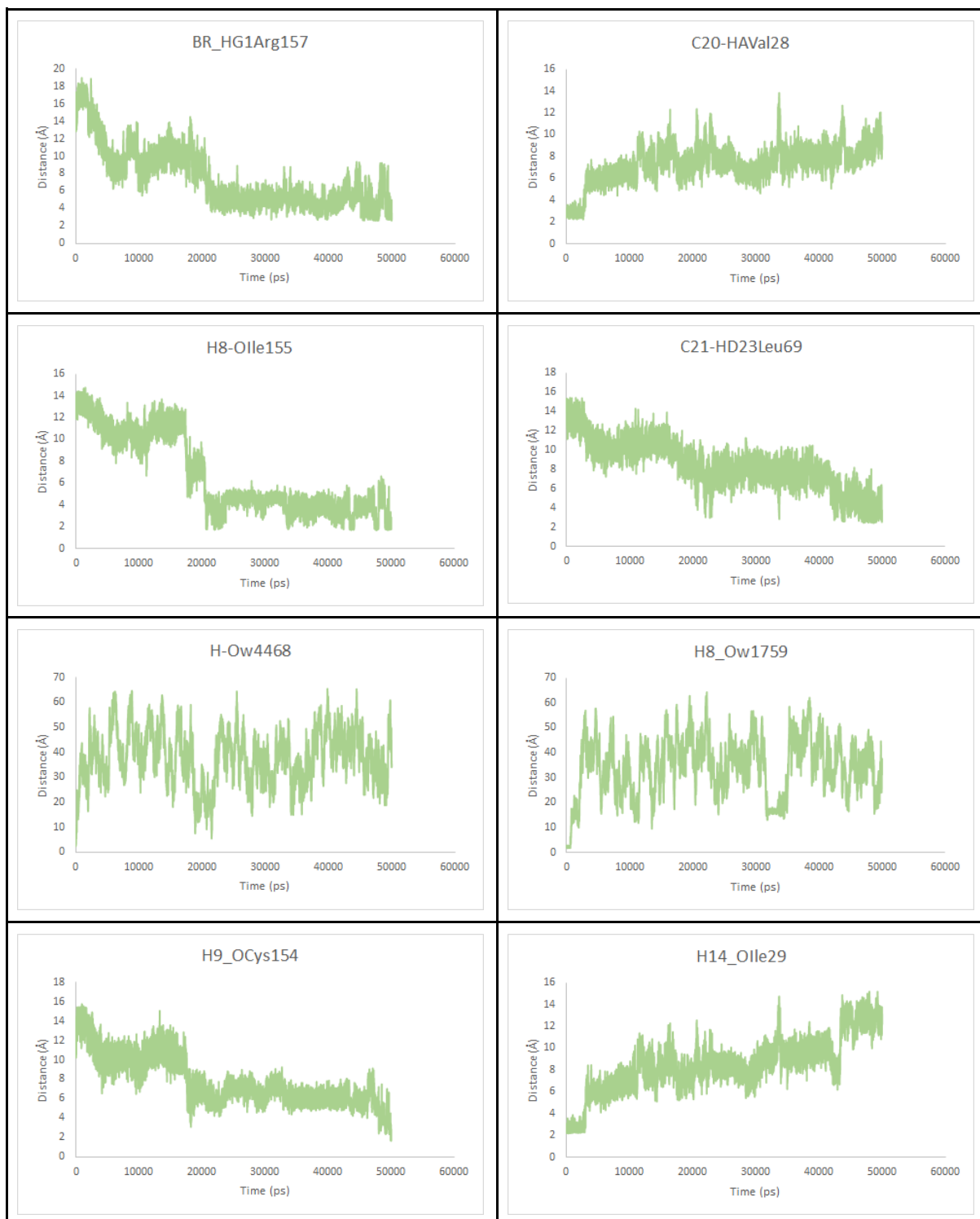
The trajectories described by each of the interactions observed by:

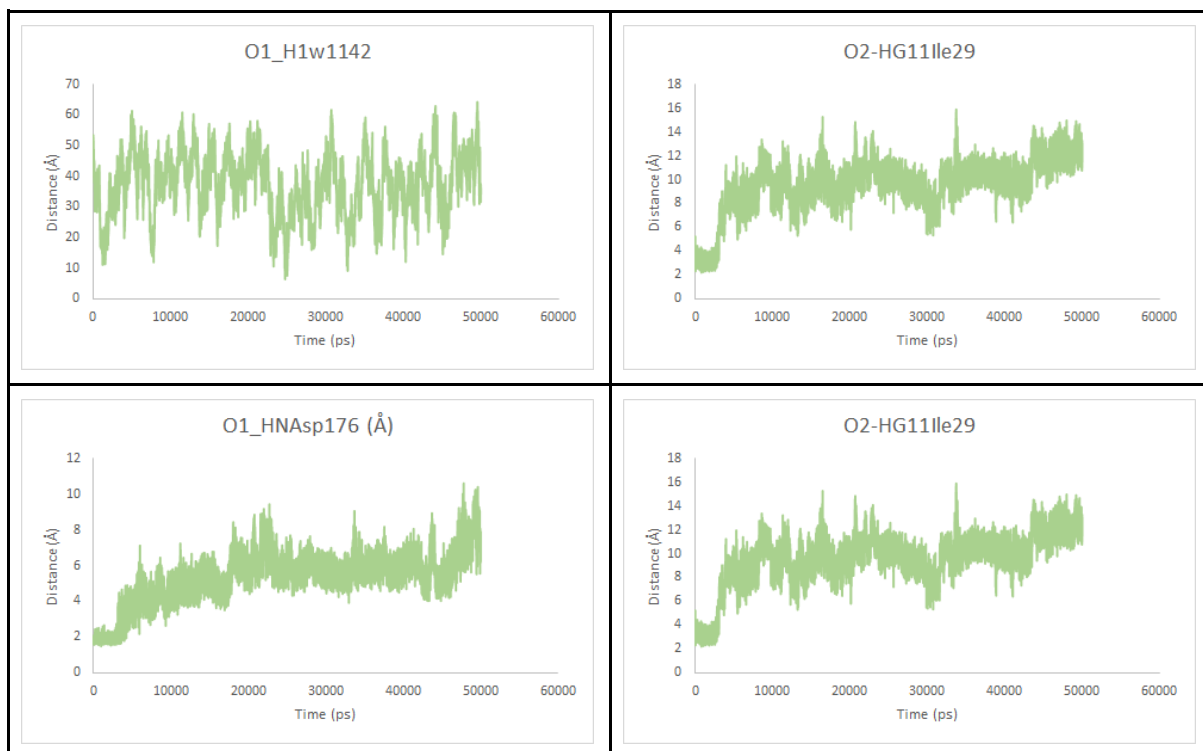
- The program "vmd" at 0ns and 50ns
- The files :
  - i6bro\_solute.avg.dat
  - solue\_i6bro.avg.dat

are represented as follows:

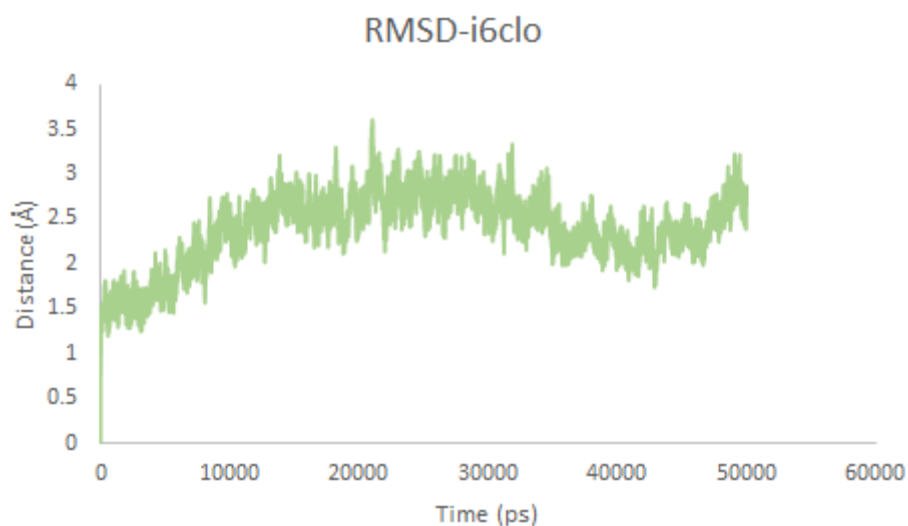
**Table 23:** Representation of the distances between interaction *INH (I6BO)* ↔ *aa*







i6co

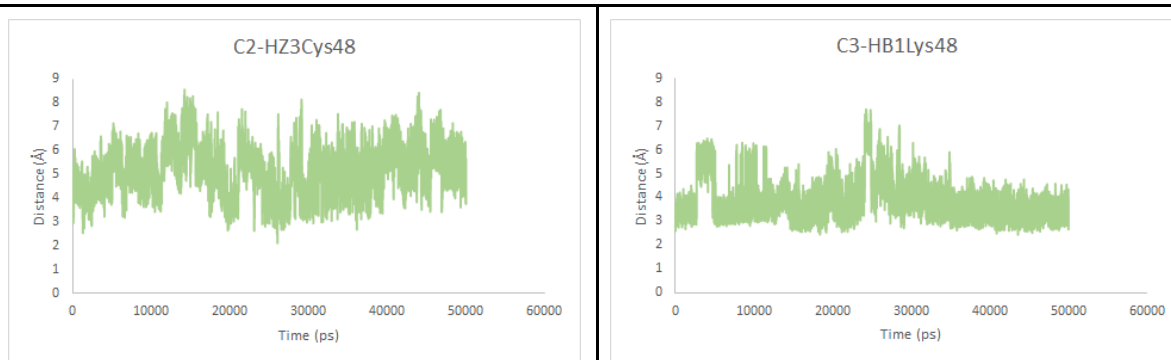


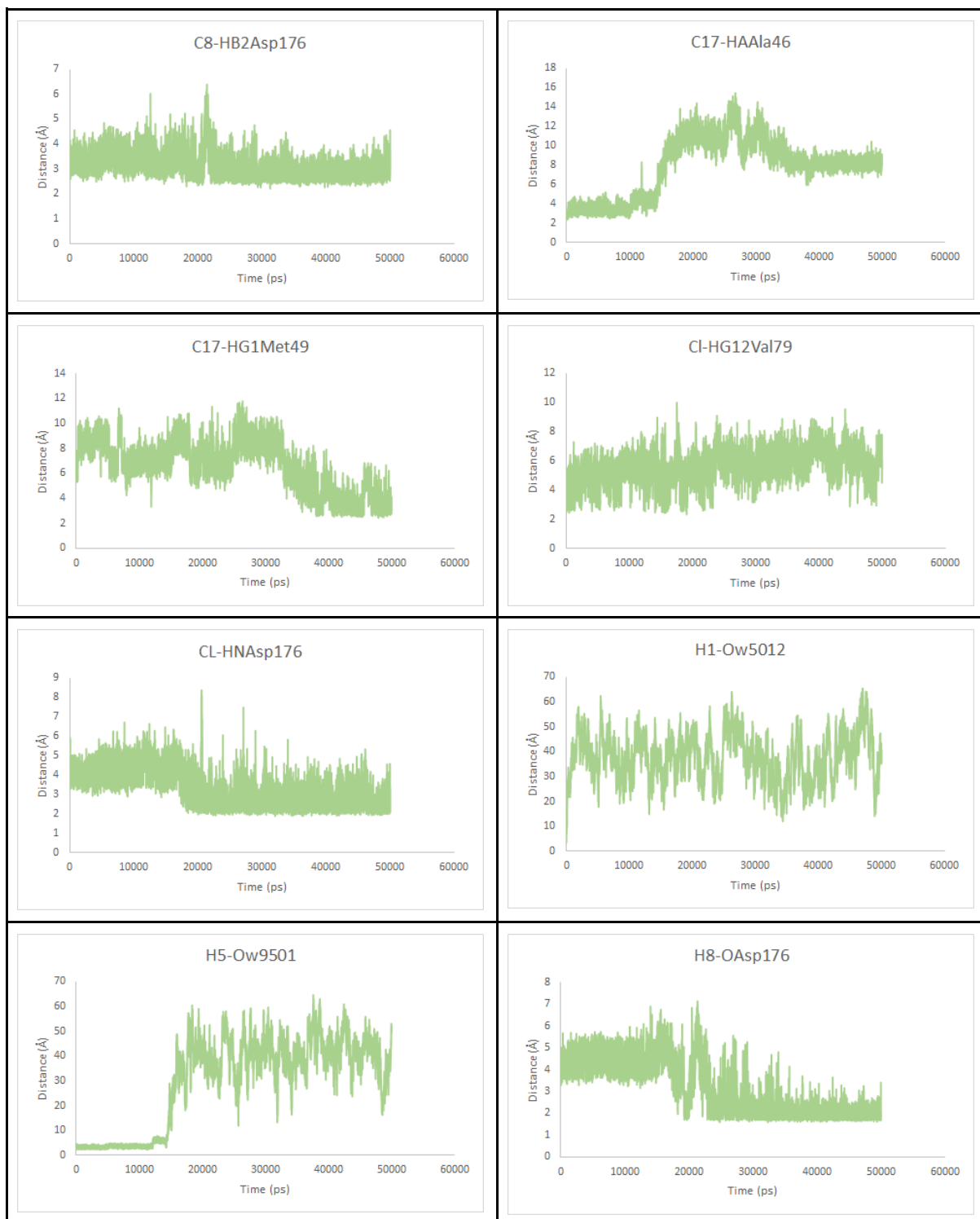
The trajectories described by each of the interactions observed by:

- The program "vmd" at 0ns and 50ns
- The files :
  - i6clo\_solute.avg.dat
  - solue\_i6clo.avg.dat

are represented as follows:

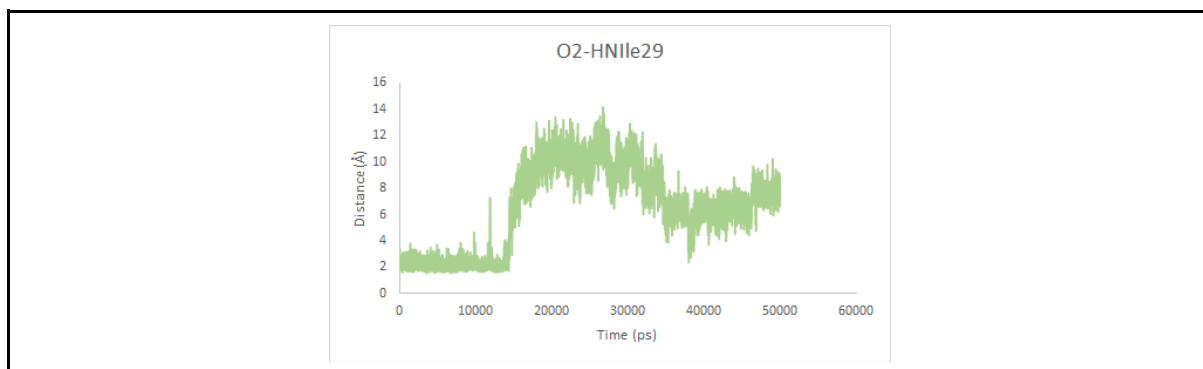
**Table 24:** Representation of the distances between interaction *INH (I6CO) ↔ aa*



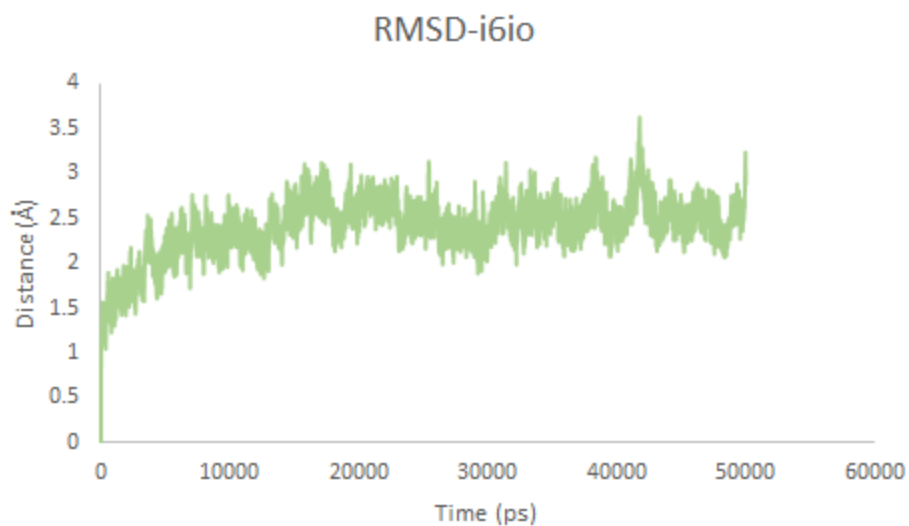








i6io



The trajectories described by each of the interactions observed by:

- The program "vmd" at 0ns and 50ns
- The files :
  - i6io\_solute.avg.dat
  - solue\_i6io.avg.dat

are represented as follows:

**Tabla 25:** Representation of the distances between interaction *INH (I6IO) ↔ aa*

