POLITECNICO
MILANO 1863

Politecnico di Milano

School of Industrial and Information Engineering Master of Science in

Materials Engineering & Nanotechnology

Investigation of Intermolecular Interactions for Molecular Sensing of Drugs: the case of Carbamazepine interacting with Azahelicenes

Supervisor: Prof. M. Tommasini

Co-Supervisor: Dr. C. Zanchi

Master Thesis of:

Gianluca Acquadro

ID 920318

Academic Year 2020 – 2021

Ringraziamenti

Mi è doveroso dedicare questo spazio della mia tesi a tutti coloro che mi sono stati vicini in questo percorso di crescita personale e professionale.

In primis, un ringraziamento al mio relatore, prof. Matteo Tommasini per la sua infinita disponibilità e tempestività dimostrata in questi mesi, nonostante le svariate problematiche dovute al COVID-19 ed al lockdown.

Un ringraziamento speciale anche alla Dr. Chiara Zanchi il cui aiuto e sostegno è stato fondamentale nelle fasi iniziali di stesura di questo elaborato.

Infine, ringrazio di cuore la mia famiglia, che mi hanno sempre sostenuto durante il mio percorso di studi.

Index

List of figures

List of tables

Abstract

The subject of this thesis is the evaluation of the interaction between Carbamazepine and two Azahelicenes (Aza[5]helicene, Aza[6]helicene) using Molecular Mechanics and DFT methods. The context of this research are the nanostructured gold substrates functionalized with Azahelicene monolayers that have been recently obtained and characterized (Zanchi 2018). This thesis examines the interaction of such Azahelicenes with target drug molecules (mainly Carbamazepine – CBZ) aiming, in the long term, at Therapeutic Drug Monitoring (TDM) and chiral recognition by exploiting Surface Enhanced Raman Spectroscopy (SERS) or plasmon-enhanced electronic spectroscopies.

After introducing the methods upon which the thesis is based (Molecular Mechanics, Raman spectroscopy and SERS), the thesis deals with the modeling of molecules using the Avogadro program and the simulation of interacting dimers through the OBSolv program. Matlab scripts have been implemented and extensively used for all the data analysis reported in this thesis.

In the last part of the thesis DFT was used to simulate the Raman and UV-Vis spectra of relevant Azahelicene-CBZ dimers determined in the previous part. Spectroscopic signatures of intermolecular interactions that could be useful for molecular sensing have been evidenced. This work, and the procedure described here, provides an effective theoretical approach for studying intermolecular interactions that can be straightforwardly extended to other conditions and molecular dimers.

Riassunto

Oggetto di questa tesi è la valutazione dell'interazione tra Carbamazepina e due Azaeliceni (Aza[5]elicene, Aza[6]elicene) mediante i metodi della Meccanica Molecolare e della teoria DFT. Il contesto di questa ricerca sono i substrati d'oro nanostrutturati funzionalizzati con monostrati di Azaelicene che sono stati recentemente ottenuti e caratterizzati da C. Zanchi e co-autori (Zanchi 2018). Questa tesi esamina l'interazione di tali Azaeliceni con molecole di farmaci di interesse analitico (principalmente Carbamazepina - CBZ) mirando, a lungo termine, al monitoraggio terapeutico dei farmaci (TDM) e al riconoscimento chirale sfruttando la spettroscopia SERS o spettroscopie elettroniche intensificate da effetti plasmonici.

Dopo aver introdotto i metodi su cui si basa la tesi (Meccanica Molecolare, Spettroscopia Raman e SERS), la tesi affronta la modellazione di molecole utilizzando il programma Avogadro e la simulazione di dimeri interagenti tramite il programma OBSolv. Gli script Matlab sono stati implementati ad hoc per questo lavoro di tesi ed ampiamente utilizzati per tutte le analisi dati qui riportate.

Nell'ultima parte della tesi è stato utilizzato il metodo Density Functional Theory (DFT) per simulare gli spettri Raman e UV-Vis di dimeri Azaelicene-CBZ determinati nella parte precedente. Sono state evidenziate le firme spettroscopiche delle interazioni intermolecolari che potrebbero essere utili nello sviluppo di sensori molecolari. Questo lavoro di tesi, e la procedura con cui è stato svolto, forniscono un approccio teorico efficace per lo studio delle interazioni intermolecolari che può essere esteso ad altre condizioni e dimeri molecolari.

1. Introduction

This thesis work belongs to a series of studies on the use of nanostructured substrates as sensors for Therapeutic Drug Monitoring (TDM) of AntiEpileptic Drugs (AEDs). TDM consists in the monitoring of a drug in the blood or serum sample of a patient in order to correlate the clinical response of the patient with the presence of the drug (Gross 1998). TDM gives information on the optimal dosage for the patient and allows to assess the drug efficacy and its safety range. This is of particular relevance when drugs are characterized by a narrow therapeutic index (NTI) (Kang 2009). In such conditions, a small variation of their dosage can result in toxic or ineffective treatment. Remarkably, for a given drug, the therapeutic index is not the same for all the patients. It depends on factors like sex, age, weight, and others (Kang 2009). TDM is needed for AEDs because the seizures occur at irregular interval, so it can be difficult to find the optimal dose on clinical grounds alone and signs of toxicity can also be difficult to reveal, mainly for patients who suffer additional neurological diseases (Johannessen 2003). Epilepsy is the most common, chronic, and serious neurological disease affecting 65 million people worldwide (Thurman 2011). In most cases, epilepsy is successfully treated, but the treatment gap is enormous, especially in developing countries (Quintas 2012), because AEDs are inaccessible or too expensive (Meyer 2010). Moreover, most patients experience side effects during the treatment such as hyperactivity, confusion, insomnia and depression (Beghi 2004). To reduce the occurrence of side effects, each patient should assume his own personalized dose of the drug. As cited above, this can be established by a reliable TDM (Mercolini 2010).

The main analytical technique used nowadays for TDM is high performance liquid chromatography (HPLC) coupled with mass spectrometry (MS). This is a relatively expensive and time-consuming technique that also requires sizeable blood sample volumes. For this reason, a complementary analytical technique has been studied in the recent years, namely Surface Enhanced Raman Spectroscopy (SERS). In a successful scenario, SERS would enable to quickly monitor drugs at low concentrations in small samples of clinical origin by using nanostructured noble metal surfaces as sensors. Such kind of surfaces can be produced with optimized plasmon resonance and uniform morphology by using pulsed laser deposition

(PLD). This is a physical vapor deposition (PVD) technique that can achieve the realization of cheap SERS sensors with costs limited to a few Euros each. In PLD the target material that has to be deposited is vaporized by a high-power pulsed laser beam. For our purposes the selected noble metal is gold, and the optimized SERS sensor shows a surface plasmon resonance peak with a maximum at about 780 nm (Pistaffa 2017). It is possible to control the film nanostructure by changing the deposition parameters and so the peak position, this means that the system can be tuned for different experimental conditions (Agarwal 2014). To date, different studies showed the good analytical potential of SERS (Jaworska 2016), mainly on simplified samples or solutions, *i.e.*, considering a scenario where the samples of clinical origin (e.g., blood plasma) are subject to solvent extraction procedures. Nowadays there are many different antiepileptic drugs on the market that can be used in different scenarios. Perampanel is a relatively new molecule that shows a good compatibility with contraceptive drugs (EMA 2012), and a prominent SERS signal that can be used for TDM purposes (Tommasini 2019). Unfortunately, not all of the AEDs exhibit a strong SERS response. For instance, this is the case of Carbamazepine (CBZ), which has been shown to weakly interact with nanostructured gold films used in SERS experiments (Zanchi 2016). Therefore, the aim of this thesis work is to understand whether the reported functionalization of gold substrates with Azahelicenes (Zanchi 2018) may improve the SERS signal for low interacting AEDs, such as CBZ. Furthermore, since Azahelicenes are chiral molecules, it is interesting to investigate possible chiral recognition phenomena that may be useful to increase the specificity against chiral drugs in TDM applications. In ref. (Zanchi 2018) it has been obtained a stable functionalization of a plasmonic gold substrate, produced by PLD, with a single molecular layer of Aza[5]helicene without losing the SERS characteristics of the sensor. In ref. (Zanchi 2018) it was also shown that these functionalized substrates are remarkably stable with time and withstand repeated washing with methanol. This means that the Azahelicene is strongly bonded to the gold surface (Zanchi 2018). Hence, such Azahelicene-functionalized SERS sensors can achieve new characteristics and capability due to the peculiar interaction between the AEDs molecules and the layer of functional molecules. For this purpose, the family of Azahelicenes is a good candidate. Indeed,

Azahelicenes have shown a strong SERS signal because of π-conjugation (Zanchi 2018) & (Zanchi 2019). Furthermore, it is also possible to speculate about stereospecific interactions with chiral molecular systems (drugs), which is one of the specific subjects of my thesis work.

CBZ is a widely studied AED whose SERS behavior has been investigated in several works (Zanchi 2016, Tommasini 2019). CBZ displays two iso-energetic isomeric forms that are the mirror image one of the other Figure 1.

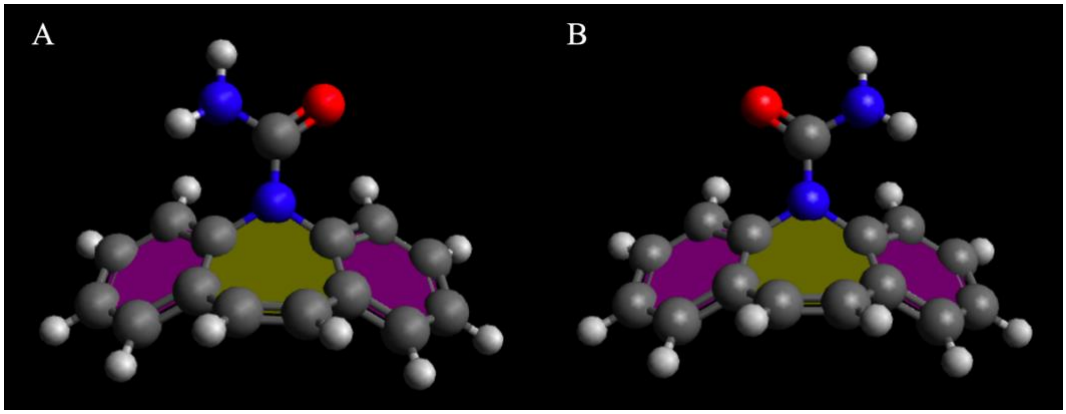

Figure 1: iso-energetic forms of carbamazepine (CBZ). On the left, the isomeric form named A. On the right, the isomeric form named B.

Such two forms could be named enantiomers, but since they may interconvert at room temperature due to low potential barriers, CBZ is not a chiral drug. However, from the conceptual point of view CBZ may serve very well to benchmark chiral recognition effects in molecular interactions with Azahelicenes. Hence, in principle, by functionalizing plasmonic gold surfaces with enantiomerically pure Azahelicenes, one could achieve a SERS sensor capable of a chiral recognition behavior.

The approach used in my thesis work is to simulate the interaction between CBZ and Aza[5]helicene or Aza[6]helicene using an in-house program called OBSolv that is based on molecular mechanics methods and the OpenBabel library (O'Boyle 2011, The Open Babel Package 2011). The extensive data produced by OBSolv have been analyzed with Matlab. In summary my thesis aims at providing a

theoretical proof of concept for indirect sensing from functionalized plasmonic surfaces. Future research is required to implement such kind of approaches in experimental devices, along the lines sketched in the pioneering work by Zanchi et al. (Zanchi 2018).

This thesis is organized as follows.
In Chapter 2 I will introduce the concept of Molecular Mechanics, Force field and the OBSolv program that I use for our simulations. Here I will also explain how SERS work and his enhancement mechanisms (electromagnetic and chemical).
In Chapter 3 I will present our AEDs and the Azahelicenes and start with an early investigation of their interaction mechanisms.
In Chapter 4 there will be the first results of simulations carried out with OBSolv on the interaction between Carbamazepine (CBZ) and the two Azahelicenes chosen.
In Chapter 5 I will define the geometrical parameters associated to the geometries of interaction using MATLAB. Then I will analyze the data of CBZ-Aza[5]helicene and CBZ-Aza[6]helicene and lastly compare the different cases.
In Chapter 6 there will be the comparison between the DFT Raman and UV-Vis spectra of our cases and I will study the possible presence of shift, extra peak and change in intensity.
Chapter 7 will present the conclusions and possible future applications of this thesis work.

2. Methods

2.1. Molecular Mechanics

In molecular mechanics (MM) methods the potential energy is written as a parametric function of the nuclear coordinates. The basic unit of the theory is the atom and electrons are not explicitly considered. In MM methods is not necessary to solve the Schrödinger equation because the quantum details of the system are neglected. One of the main aims of MM methods is to study the stable geometries of molecules, finding the energy minima of the potential energy. The potential energy is the sum of the following terms.

$$V = V_r + V_\theta + V_\gamma + V_\tau + V_{vdw} + V_{el} + V_{cross}$$

V_r is the contribution of stretching between two bonded atoms, A-B; V_θ is the contribution for bending the angle between three consecutively bonded atoms, A-B-C. If the central atoms B in A-B-C is sp^2-hybridized, we need to consider also the out of plane contribution, V_γ. V_τ is the torsional contribution associated to the relative rotation of two molecular fragments around a bond that connects them, it is in function of the associated torsional (dihedral) angle τ. All these are the bonded contributions to the potential energy. The interactions among non-bonded atoms are taken into account with the following two terms:

- the Van der Waals interaction V_{vdw} describes the repulsion and the interaction between non directly bonded atoms.
- the electrostatic interaction V_{el} describes the non-bonded interactions due to charge distribution of the molecule that creates slightly negative (δ^-) and positive (δ^+) parts in different groups (atoms) of the molecule itself.

Lastly, the cross contribution V_{cross} takes into account the coupling between all these fundamental (diagonal) terms.

All these terms can be expressed with different expressions, based on the type of molecule we want to model. Furthermore, the parameters used in those expressions are chosen on the basis of the atoms type and their relative position and coordination (Jensen 2006).

It is important to underline that the absolute value of the potential energy (V) has no peculiar physical meaning, as it may be related to a given hypothetical reference molecule. However, the energy differences between different conformers of the same molecule are physically meaningful quantities, that can be used, for instance, to compute the Boltzmann distribution of the conformers.

Force Fields (FFs) are used in MM to model the molecular systems and there are several classes of them.
Class 1 FFs also called "harmonic" are used to treat large systems, four this purpose the FF must be as simple as possible, so it doesn't take in account the cross terms (V_{cross}) and it uses the harmonic approximation to treat the other terms (V_r , V_θ…) and the Lennard Jones potential for the VdW interactions.
Class 2 FFs are used to study small size molecules with a high degree of accuracy. In this case the cross terms (V_{cross}) are considered, and the other terms (V_r , V_θ…) are modeled by cubic or quartic expansions; the VdW is modeled by an exponential type potential.
Class 3 FFs are used to model system with strong electronic polarization effects. In this case the parameters depend on the neighboring atoms.

In my thesis work, the choice was a Class 2 FF. In general, the inputs required for the calculation are:

- The atom types
- How the atoms are bonded and their hybridization
- A plausible initial geometry

The first two inputs determine the parameters to be used in the selected FF, thereafter the molecular structure can be optimized by minimizing V starting from the initial geometry (third input).
Modern programs, like Avogadro, a molecule editor and visualizer designed for computational chemistry and materials science (Hanwell 2012), allow one to define all required input (connectivity, bond types, initial geometry) by simply drawing the molecule on the screen, using a graphical interface.

In my thesis work the MM simulations have been carried out through the use of OBSolv, a program written in C++ (iso 2011) based on a Monte Carlo approach, implemented through OpenBabel. The latter is a very practical C++ library for MM that contains several FFs optimized for studying small molecules (O'Boyle 2011, The Open Babel Package 2011). The OBSolv code was developed to simulate the solvation of small molecules like drugs in water or organic solvents. In a series of tests carried out during the development of OBSolv, which were presented by M. Tommasini at the VOA6 conference (https://www.voa6.org), the program has shown a good correlation between the calculated energies of solvation and the measured solvation free energy of a series of small organic compounds (Nicholls 2008). The code reads the structures of the solute and of the solvent and proceeds by placing an increasing number of solvent molecules at random positions and orientations around the solute. For each step of solvent molecule addition, the program evaluates the most energetically favorable placement among a selected number of the successful trial positions (N_{hits}). The evaluation of the energy is carried out by suitable calls to the OpenBabel library that implements a selection of molecular mechanics force fields (O'Boyle 2011, The Open Babel Package 2011). The target number of successful trials (N_{hits}) sampled by the Monte Carlo algorithm is initially defined by the user. At each trial the algorithm evaluates the distance between the solvent [A] and the solute [B] by the formula $d(A,B) = min[d_{ij}]$, where d_{ij} is the distance between the atom $i \in A$ and the atom $j \in B$. Then the reduced distance defined as $\xi_{ij} = d_{ij}/(R_{VdW,i} + R_{VdW,j})$, is compared with respect to a given threshold to assess whether the solvent molecule is close enough to the solute, so that it is relevant to compute the energy (which is the time-consuming part of the algorithm). The condition ξ =1.0 corresponds to a distance d_{ij} equal to the sum of Van der Waals radii for the solute-solvent pair.

To proceed with the evaluation of the energy, the program considers acceptable a range of reduced distances ($t_{mid} < \xi_{ij} < t_{max}$). This condition allows the program to quickly discard many Monte Carlo trials that are associated to energy unfavorable conditions (i.e., a solvent molecule too close or too far from the solute). The program can also consider a more relaxed condition ($t_{min} < \xi_{ij} < t_{mid}$) which

is done to take into account hydrogen bonds that imply a closer approach between molecules.

In my work the OBSolv program has been used in a low coverage condition of the solute. In particular I consider just one solvent molecule. Therefore, the solvation energy computed by OBSolv corresponds to the interaction energy of the two molecules of interest. Also, the calculation is faster because it involves less atoms. Among the force fields offered by OpenBabel, I selected the MMFF94s force field, which is a slightly modified version of the Merck Molecular Force Field (MMFF94), optimized for the study of small drug molecules. The MMFF94 uses a novel "Buf-14-7" form for the Van der Waals interactions and also employs a buffered Coulomb term for electrostatic interactions. This approach not only confers generality but also enables MMFF94 to describe a broad range of chemical structures and interactions very well (Halgren 1995).

2.2. Raman Spectroscopy

Raman spectroscopy (RS) is a widely used technique for the characterization of materials that measures the quantum transitions between vibrational states. RS may provide information about the chemical and physical state of the analyte. The Raman spectra provide characteristic fingerprint by which the molecular structure of the samples can be identified. When light interacts with matter two kind of scattering processes are possible: the elastic scattering, which is the one most probable, and the inelastic scattering that is at the basis of RS. Due to inelastic scattering the photon energy is modified, and the energy difference between the incoming photon and the outgoing (scattered) photon is equal to the energy taken or given by the molecule during the transition between two vibrational states. To obtain an acceptable signal a powerful light source is mandatory, usually a laser (Kneipp 2001).

We can distinguish three possible scattering processes:

- Rayleigh scattering: an elastic collision process that conserves the photon energy; the frequency of the incoming photon (ν) is equal to the frequency of the outgoing photon (ν').
- Stokes scattering: an inelastic collision process. Part of the photon energy is given to the molecule which is excited to a higher vibrational level. The scattered photon has a frequency lower than the incoming photon ($\nu' < \nu$) – it is red shifted.
- Anti-Stokes scattering: an inelastic collision process. Part of the vibrational energy of the molecule is given to the photon. The scattered photon has a frequency higher than the incoming photon ($\nu' > \nu$) – it is blue shifted.

The relative intensity of the Stokes and anti-Stokes processes depends on the population of the initial vibrational state, which is given by the Boltzmann distribution.

Is possible to increase the Raman signal of some orders of magnitude by working at resonance condition, which leads to resonance Raman spectroscopy (RRS). Another possibility is to use Surface Enhanced Raman spectroscopy (SERS): thanks to plasmonic effects at the surface of metallic nanostructures this technique may reach an enhancing factor of several orders of magnitude.

2.2.1. Classical description of the Raman effect

The oscillating electric field associated to light induces a polarization of the electron charge distribution of the irradiated molecule. At first order, the associated dipole moment P can be expressed as:

$$P = \alpha E$$

where α is the polarizability tensor. The electric field can be considered to be just time dependent, as the wavelength of the light used in Raman experiments is much larger than the size of the sampled molecule. The time dependence of the normal coordinates (Q_j) can be used to describe (classically) the vibration of a molecule:

$$Q_j = q_0^j \cos(\omega_j t + \varphi_j)$$

As the molecule vibrates, the polarizability is modulated by the different normal modes. By expanding the polarizability as a Taylor expansion in the normal coordinates, and dropping any non-linear term, the molecular polarizability is written as:

$$\alpha = \alpha_0 + \sum_j \left(\frac{\partial \alpha}{\partial Q_j}\right)_0 Q_j$$

As the electric field oscillates, $E(t) = E_0 \cos(\omega_L t)$, the dipole moment can be finally written as:

$$P = \alpha E = \alpha_0 E_0 \cos \omega_L t + E_0 \left(\frac{\partial \alpha}{\partial Q}\right)_0 \cdot \frac{1}{2}\left[\cos\big((\omega_L - \omega)t - \varphi\big) + \cos\big((\omega_L + \omega)t + \varphi\big)\right]$$

This equation shows that light scattering may occur at three different frequencies. The first term is associated with Rayleigh scattering, and it takes place at the same frequency of the light source ω_L. The second term takes place at $\omega_L - \omega$ and it is associated with Stokes scattering; the third term is associated to anti-Stokes scattering and it takes place at $\omega_L + \omega$. It is important to underline that by changing the frequency of the light source ω_L the relative position (and the shape) of the Raman spectrum does not change, but only shifts as a whole.

2.2.2. Quantum description of the Raman effect

The potential energy of a diatomic molecule can be described by the Morse potential and can be further simplified by using the harmonic approximation. The system is then described by a harmonic quantum oscillator with energy levels that are labelled by the quantum number n ($n = 0,1,2 \ldots$).

$$E_n = \hbar\omega\left(n + \frac{1}{2}\right)$$

Off-resonance Raman scattering can be described by introducing non-stationary states. Such states are highly unstable because they do not correspond to any eigenstate of the system. If the energy of the incoming photon is not matching any electronic transition, the molecule will be excited in the virtual state and then rapidly decay in a lower level emitting a new photon. The relative position of the three states (initial, virtual and final) lead to the three cases of Rayleigh, Stokes and anti-Stokes Raman scattering, as one can see in Figure 2.

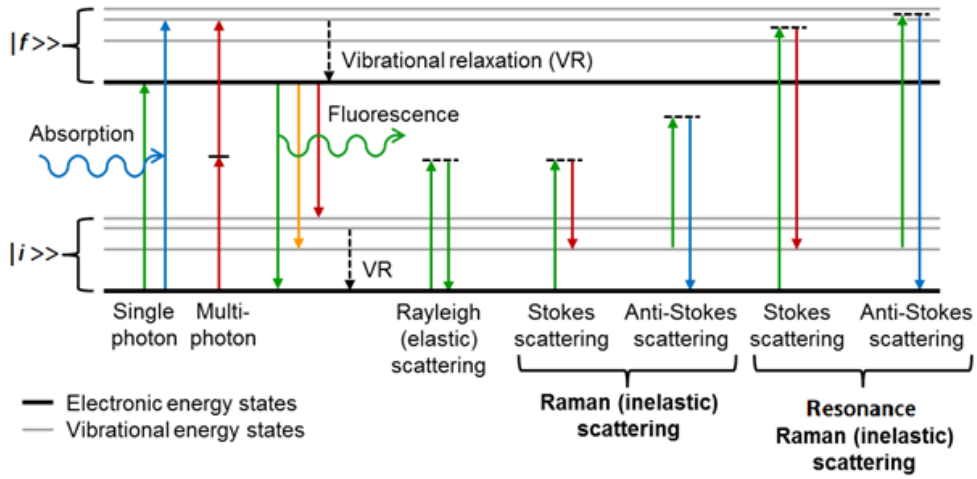

Figure 2: scheme describing absorption/fluorescence, normal Raman scattering and resonance Raman scattering (from left to right) in relation with the electronic and vibrational energy levels of the molecule.

Due to the Boltzmann population distribution, the anti-Stokes scattering will be always less probable than the other two scattering processes, because his initial state is higher in energy, so it is less populated.

2.3. Surface Enhanced Raman Scattering (SERS)

The Surface Enhanced Raman Scattering (SERS) effect was observed for the first time by M. Fleischmann in 1974. He measured an intensified Raman spectrum of pyridine on a rough silver electrode (Fleischmann 1974). Three years later, two independent research groups, gave two possible interpretations of the enhancement phenomena. The first group proposed that the enhancement was due to the intensification of the electromagnetic field in the surrounding of the silver surface (Jeanmaire 1977). The second group suggested that the interaction between the surface and the analyte led to the modification of the molecular electronics level, obtaining a chemical enhancement factor associated to the polarizability (α) (Albrecht 1977). Nowadays is well know that the SERS phenomena involves both effects, the electromagnetic enhancement (with a factor of 10^{10}-10^{11}) and the chemical enhancement (with a factor of the order of 10^{4}-10^{5}). Overall, in the best condition, the net enhancement may reach a factor of the order of 10^{14}-10^{15}.

2.3.1. Optical proprieties of a metal

To understand the SERS phenomena, we need to introduce the description of plasmonic effects. In a first approximation, the interaction between the electromagnetic radiation and the metal can be explained by the Drude model. The metal is assumed to consist of fixed positive ions surrounded by a gas of free electrons. In this kind of systems, the most relevant collision phenomena happen between electrons and ions. After the collision process (associated to a relaxation time) the velocity of the electrons depends only on the temperature. By applying a uniform electric field E, the electrons will accelerate, as described by the following equation of motion:

$$\frac{dp}{dt} = -\frac{p}{\tau} - eE$$

p is the momentum and τ is the characteristic relaxation time. One may consider a time dependent electric field $E = E_0 e^{-i\omega t}$ and rewrite the equation of motion as:

$$m\frac{d^2x}{dt^2} + \gamma m\frac{dx}{dt} = -eE$$

The solution is:

$$x(t) = \frac{e}{m(\omega^2 + i\gamma\omega)} E(t)$$

This result implies that the displacement of the electrons contributes to the macroscopic polarization of the metal (P) because $P = n\mu$, with n being the electron density, and μ being the electron dipole moment, $\mu = -ex$. Hence, P can be written as:

$$P = n\mu = -\frac{ne^2}{m(\omega^2 + i\gamma\omega)} E(t)$$

By remembering that the dielectric displacement is $D = \varepsilon_0 E + P$, which can be simplified for linear, isotropic nonmagnetic media as $D = \varepsilon_0 \varepsilon E$, one obtains:

$$\varepsilon(\omega) = 1 + \frac{P}{\varepsilon_0 E} = 1 - \frac{\omega_p}{(\omega^2 + i\gamma\omega)}$$

With the plasma frequency: $\omega_p = \sqrt{\frac{ne^2}{\varepsilon_0 m}}$. Now, by considering the real and the imaginary part of $\varepsilon(\omega)$ we can identify three regimes:

$$\varepsilon_1(\omega) = Re\{\varepsilon(\omega)\} = 1 - \frac{\omega_p^2}{\omega^2 + \gamma^2}$$

$$\varepsilon_2(\omega) = Im\{\varepsilon(\omega)\} = \frac{\omega_p^2 \gamma}{\omega(\omega^2 + \gamma^2)}$$

- $\omega > \omega_p$: the real part is positive, and the imaginary part is negligible. The medium is transparent to the radiation.
- $\omega < \omega_p$: the real part is negative. The radiation can propagate in the metal with an intensity that decay exponentially over the penetration depth, the material absorbs the wave and produces an evanescent field.
- $\omega = \omega_p$: the imaginary part is not negligible; in this resonant condition, the electrons collectively oscillate at the same frequency, the quanta of this oscillation are called plasmons.

Plasmon can be both bulk plasmons or surface plasmons (SP). By using Maxwell's equations and the correct boundary conditions, one finds that SP are bound modes at the interface between metal and another dielectric medium that happen with frequency ω_{sp}. Such bound modes are those relevant for explaining the SERS effect.

$$\omega_{sp} = \frac{\omega_p}{\sqrt{1 - \varepsilon_{dielectric}}}$$

2.3.2. Electromagnetic enhancement

SERS substrates are made of metal nanoparticles (NP). A fundamental collective excitation of electron for those systems are the so-called localized surface plasmon (LSP). LSP are non-propagating excitations resulting from the confinement of SP in a metal nanoparticle with a size lower than the excitation wavelength. By applying an external field (E_0) to such nanoparticles will generate the oscillation of the electron cloud in the NP that generates a new electric field E_{sp}. The total electric field is:

$$E_{tot} = E_0 + E_{sp}$$

E_{tot} can be considered an electrostatic field because the dimension of the NP is much lower than the excitation wavelength. By solving the Laplace equation, it is possible to obtain the internal and the external potential, from which we obtain the expression for the polarizability and for the induced electric field.

$$\alpha = 4\pi a^3 \frac{\varepsilon - \varepsilon_m}{\varepsilon + 2\varepsilon_m}$$

$$E_{in} = \frac{3\varepsilon_m}{\varepsilon + 2\varepsilon_m} E_0$$

The plasmon resonance condition becomes:

$$\varepsilon_1(\omega) = 1 - \frac{\omega_p^2}{\omega^2} = -2\varepsilon_m$$

And so, the localized surface plasmon frequency is:

$$\omega_{lsp} = \frac{\omega_p}{\sqrt{1 + 2\varepsilon_m}}$$

And the electric field associated to the LSP is:

$$E_{lsp} = \frac{\varepsilon - \varepsilon_m}{\varepsilon + 2\varepsilon_m} \frac{a^3}{(a+d)^3} E_0$$

Where a is the radius of the NP and d is the distance between the analyte and the surface of the NP. For simplicity, we consider the case $E_{lsp} \gg E_0$. In this condition we can define the enhancement factor as:

$$A(\omega) = \frac{E_{TOT}}{E_0} \approx \frac{E_{lsp}}{E_0} = \frac{\varepsilon - \varepsilon_m}{\varepsilon + 2\varepsilon_m} \frac{a^3}{(a+d)^3}$$

The intensity of the SERS signal (I_{SERS}) due to electromagnetic enhancement contains the excitation contribution and the scattered radiation contribution.

$$I_{SERS}(\omega) = |A(\omega_s)|^2 |A(\omega_L)|^2 \approx \left|\frac{\varepsilon(\omega_s) - \varepsilon_m}{\varepsilon(\omega_s) + 2\varepsilon_m}\right|^2 \left|\frac{\varepsilon(\omega_L) - \varepsilon_m}{\varepsilon(\omega_L) + 2\varepsilon_m}\right|^2 \frac{a^{12}}{(a+d)^{12}}$$

Another important contribution can be found by considering arrays of NPs which can create a hot spot. In those points the interaction between different plasmons further intensify the EM field.

The different contributions are summarized in Figure 3.

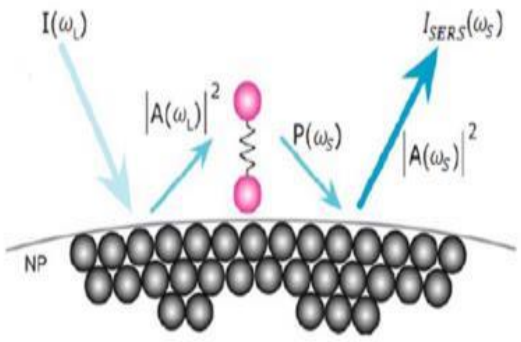

Figure 3: schematic representation of the electromagnetic enhancement (VILLA, 2018).

2.3.3. Chemical enhancement

The chemical enhancement is the intensification of the SERS signal due to the chemical interaction between the analyte and the nanostructured metallic substrate. The interaction mechanism is the so-called charge transfer effect that modifies the electronic structure of the analyte molecule by creating new electronic states that undergo resonance conditions with the incoming radiation. The chemical enhancement is strongly dependent on the nature of the molecule, its concentration and the interaction with the metal surface. Also, the geometry and morphology of the substrate and the orientation of the analyte molecule play a relevant role, as they affect the degree of superposition of the orbitals and lead to changes in the polarizability. The enhancement factor depends on the number of adsorbed molecules N_{ads} and on the Raman cross-section σ_{ads} of the adsorbate:

$$I(\omega_s) \propto N_{ads}\sigma_{ads}$$

2.3.4. Global enhancement

The global enhancement is the product of the two contributions, electromagnetic and chemical. The intensity of the SERS signal can be thus written as:

$$I_{SERS}(\omega_s) = n_L N_{ads}\sigma_{ads}|A(\omega_s)|^2|A(\omega_L)|^2$$

Where n_L is the number of incident photons per unit surface and time, and the subscripts S and L stand for Stokes and Laser.

3. A molecular mechanics survey of Antiepileptic drugs and Azahelicenes

3.1. Antiepileptic drugs (AEDs)

Nowadays, many different AEDs are available on the market. Such AEDs act in different ways and are therefore suitable for specific epilepsies and conditions. In theory, taking into account the medical history and the other characteristics of the patients is possible to select the best AEDs to use for each specific case.

In the first step of my thesis work, I have considered a number of AEDs that are currently used for treating epilepsy, and that are interesting for TDM. I have investigated their structure to assess those which might be the best candidates for showing a good molecular interaction with Azahelicenes. I have initially considered 26 AEDs. I have designed their molecular structures in Avogadro, and I have obtained their geometry-optimized structures. In this way we can assess the bond distances, bond angles and also other geometrical parameters of the selected AEDs. From the initial set of 26 AEDs, I have selected eight AEDs on the basis of some structural characteristics:

- π-conjugation (the more, the better)
- rigidity (i.e., minimal conformational flexibility)
- presence of aromatics rings

Such requirements were selected to choose candidate molecules that are likely to exhibit a good SERS signal, thanks to π-conjugation, and at the same time do not presents a large number of possible conformers because of their rigidity. Conformational flexibility may give rise to a manifold of interaction geometries with the SERS substrate, which could significantly broaden the SERS signals, making them less robust for sensing purposes. Finally, the presence of aromatic rings was required, so to possibly achieve a good π-stacking interaction mechanism with Azahelicenes.

The selected eight candidates are represented in Figure 4.

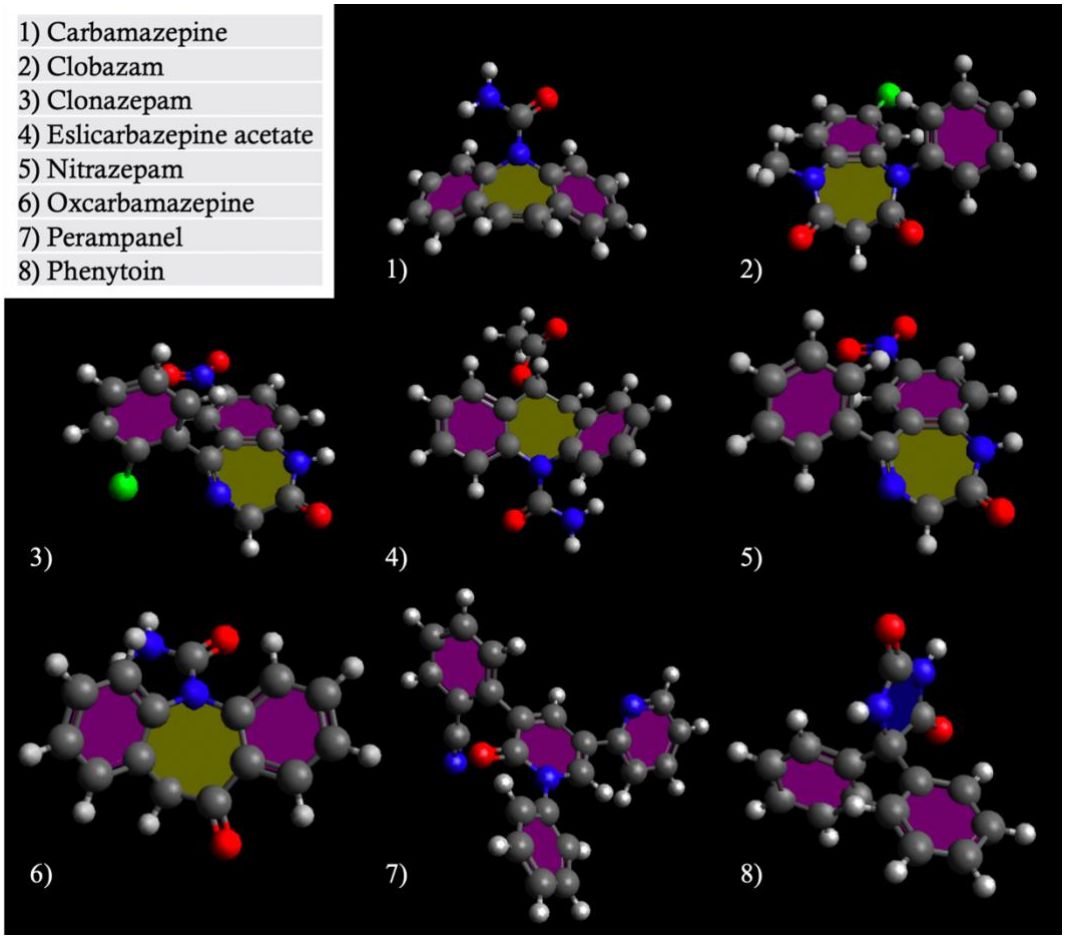

Figure 4: ball & stick representation of the equilibrium structures of the eight AEDs selected for the assessment of the interaction with Azahelicenes.

By using Avogadro I could explore the energetically stable conformers of the selected eight AEDs. It is important to underline that to correctly investigate the conformers of such drugs, I had to select a force field optimized for simulating the molecular mechanics of small molecules. Among the choices available in Avogadro, I adopted the MMFF94s force field.

Here below, I report the conformers and associated energies of the eight selected AEDs.

Carbamazepine displays rather a rigid structure, with just two isoenergetic stable conformers (named A and B – see Table 1, Figure 5).

Table 1: energy of the two conformers of Carbamazepine

Carbamazepine	E_1^0 *(kcal/mol)*
A & B	39.41

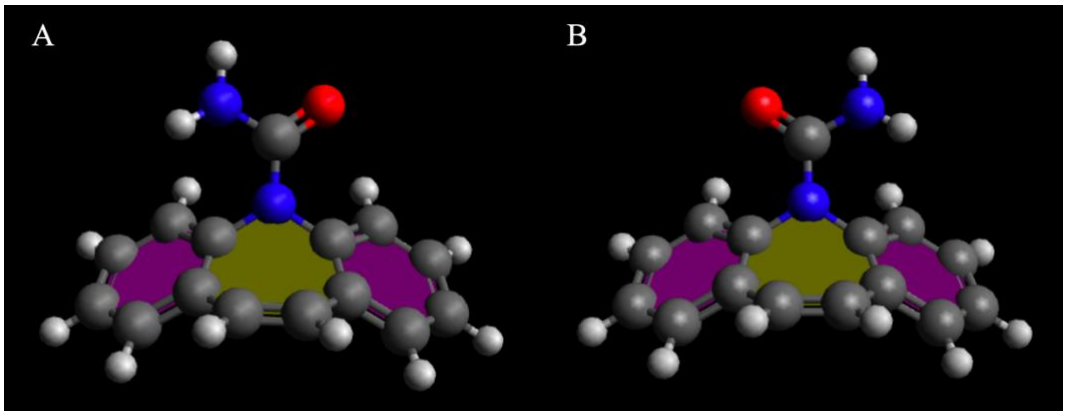

Figure 5: conformers of Carbamazepine

The situation for Clobazam (Table 2, Figure 6), Nitrazepam (Table 3, Figure 7) and Phenytoin (Table 4, Figure 8) is similar to that of Carbamazepine, with just two stable conformers.

Table 2: energy of the two conformers of Clobazam

Clobazam	E_1^0 *(kcal/mol)*
A & B	58.82

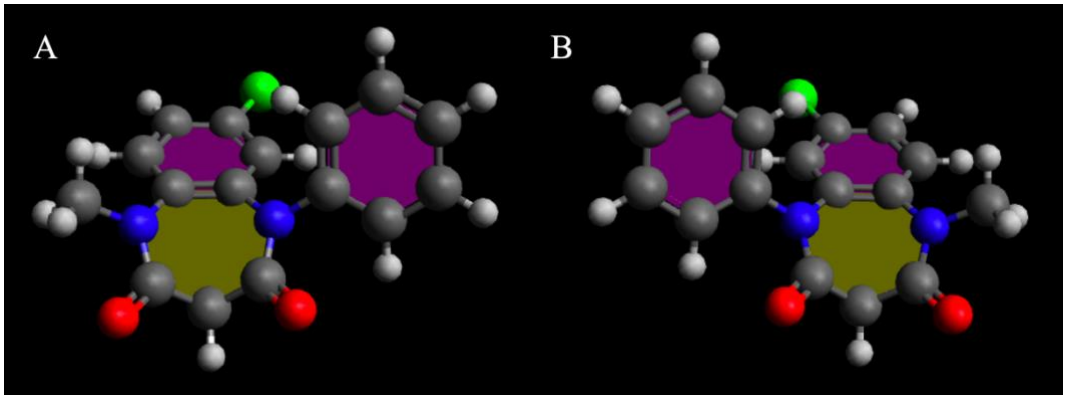

Figure 6: conformers of Clobazam. The transition between the two specular conformations occurs by inversion of the heptagon, represented in yellow

Table 3: energy of the two conformers of Nitrazepam

Nitrazepam	E_1^0 *(kcal/mol)*
A & B	131.60

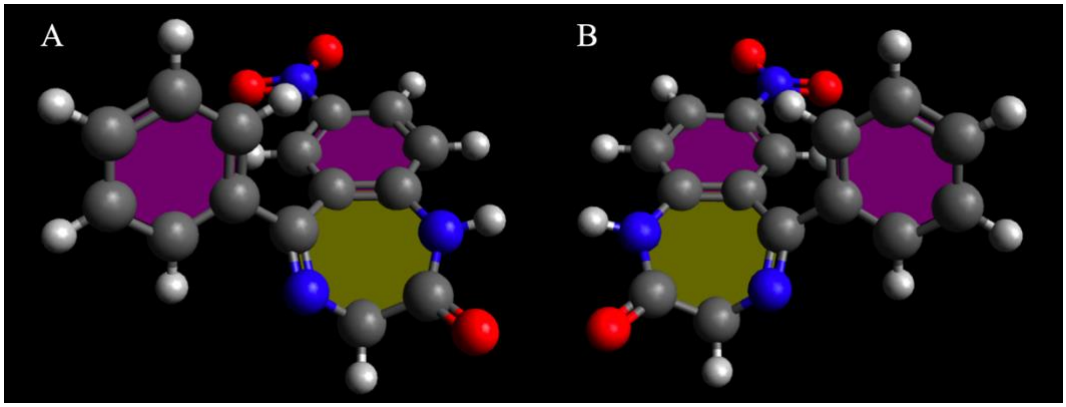

Figure 7: conformers of Nitrazepam. The transition between the two specular conformations occurs by inversion of the heptagon, represented in yellow

Table 4: energy of the two conformers of Phenytoin

Phenytoin	E_1^0 *(kcal/mol)*
A & B	-1.20

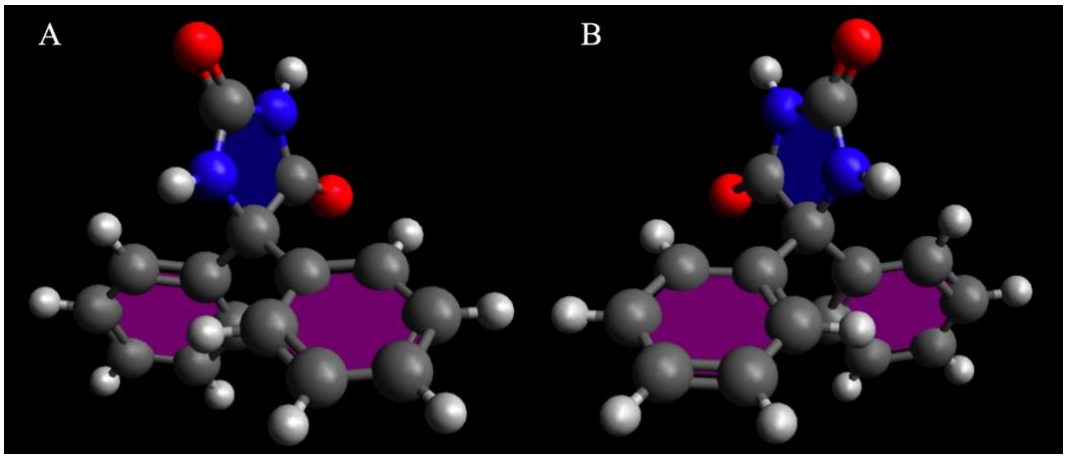

Figure 8: conformers of Phenytoin

For Clonazepam (Table 5, Figure 9) and Oxcarbamazepine (Table 6, Figure 10) I have obtained four possible conformers, as detailed below.

Table 5: energy of the four conformers of Clonazepam

Clonazepam	E_1^0 *(kcal/mol)*	ΔE *(kcal/mol)*
B & D	132.12	0.00
A & C	134.54	2.42

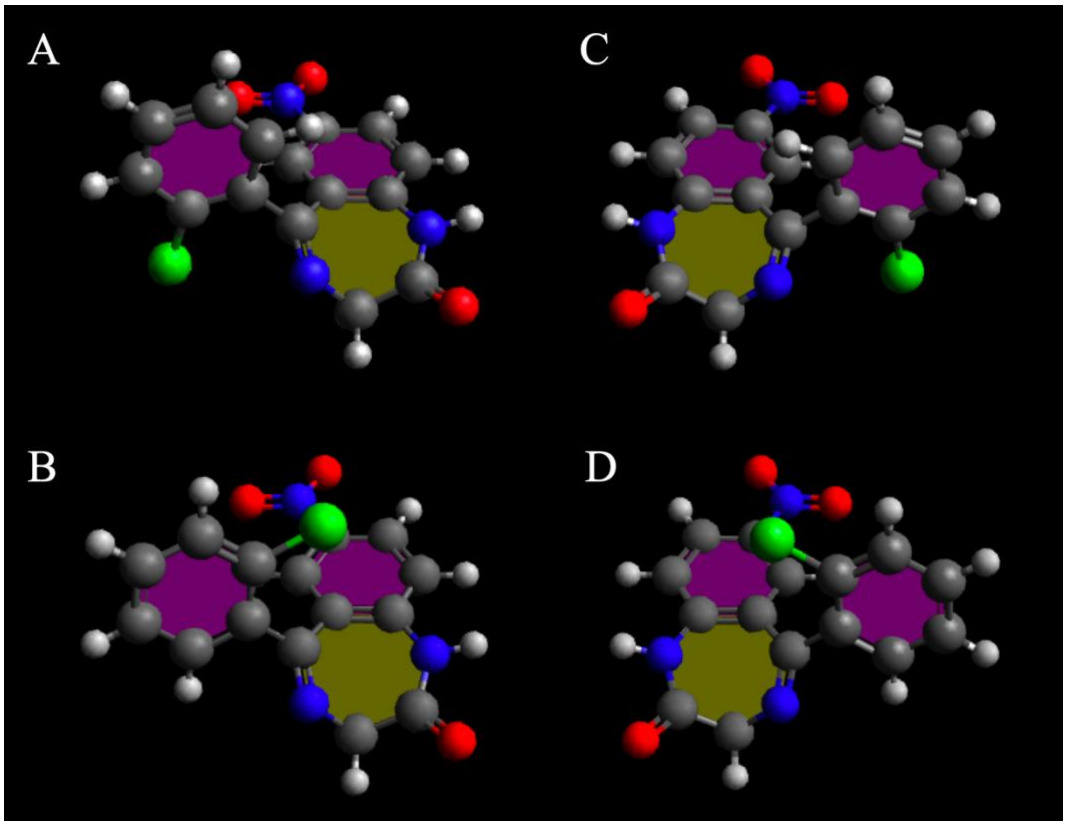

Figure 9: conformers of Clonazepam. The transition between two specular conformations (A/C & B/D) occurs by inversion of the heptagon, represented in yellow

Table 6: energy of the four conformers of Oxcarbamazepine

Oxcarbamazepine	E_1^0 *(kcal/mol)*	ΔE *(kcal/mol)*
A & C	46.10	0.00
B & D	48.02	1.92

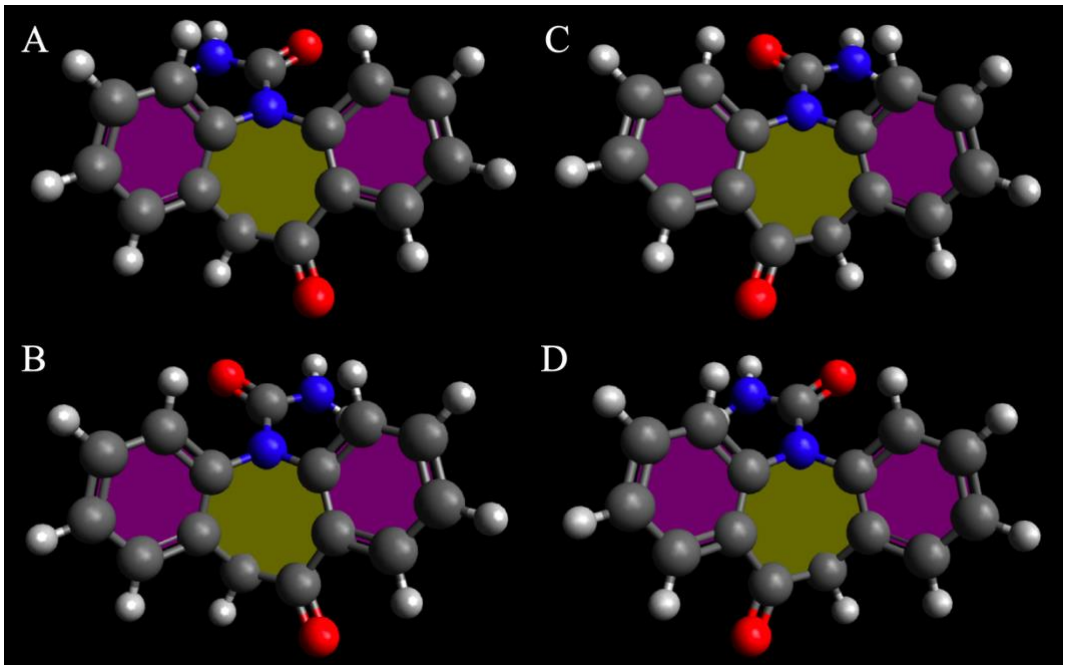

Figure 10: conformers of Oxarbamazepine. The transition between two specular conformations (A/C & B/D) occurs by inversion of the heptagon, represented in yellow

For Perampanel, the number of the conformers that I have found rises up to eight (Table 7, Figure 11).

Table 7: energy of the eight conformers of Perampanel

Perampanel	E_1^0 *(kcal/mol)*	ΔE *(kcal/mol)*
B & G	113.31	0.00
A & H	113.67	0.36
D & E	113.96	0.65
C & F	114.32	1.01

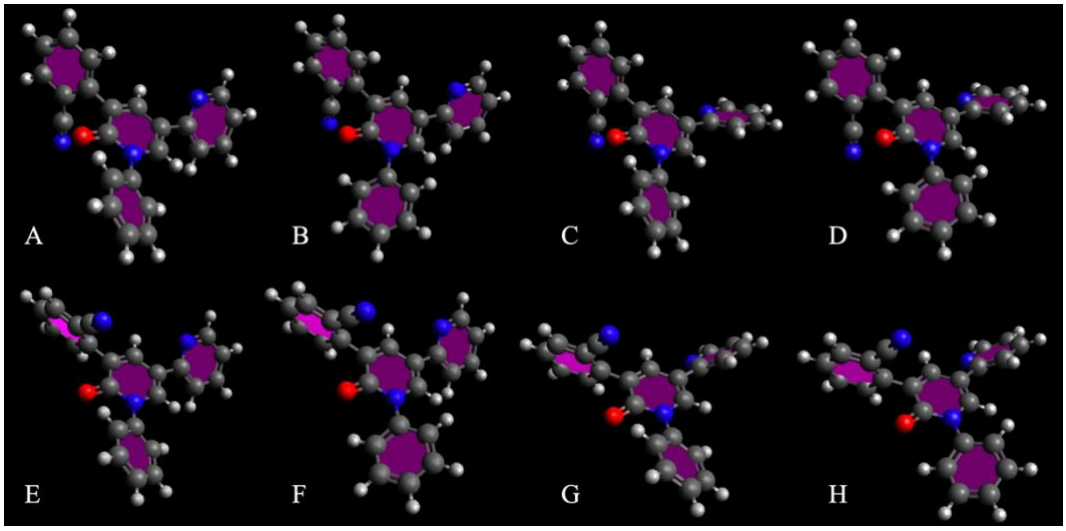

Figure 11: conformers of Perampanel

The Eslicarbazepine acetate was initially chosen for his similarity with Carbamazepine and the Oxcarbamazepine but during the early stage of my analysis it was immediately clear that this AED was outside the scope of my work. This was due to the fact that the Esilcarbamazepine acetate, unlike the other two molecules, has a flexible chain in its structure and so the number of possible conformers is too high. For this reason, it was discarded.

3.2. Aza[5]helicene & Aza[6]helicene

Besides the AEDs described in the previous section, in my thesis work I also consider the two Azahelicenes named Aza[5]helicene and Aza[6]helicene. By studying the interaction of the selected AEDs with those Azahelicenes, I want to understand if there is an appreciable difference of interaction for the two cases. The initial hypothesis is that the two Azahelicenes may act differently, because of the different extension of the helical coil – which is longer in Aza[6]helicene than in Aza[5]helicene. Such structure difference between the two Azahelicenes may lead to specific interaction differences for a given interacting AED. This could lead to the possibility of optimizing our SERS substrate based on the specific antiepileptic drug that we are studying.

Brief chemical background. Azahelicenes can be synthetized in many different ways. One of the simplest strategies that can be used starts from the photochemical ring-closure of 1,2-diarylethylenes. The key step of this process is the photolysis of differently nitrogen-substituted 1,2-diarylethylenes (Bazzini 2005). The photochemical cyclization is represented in Figure 12.

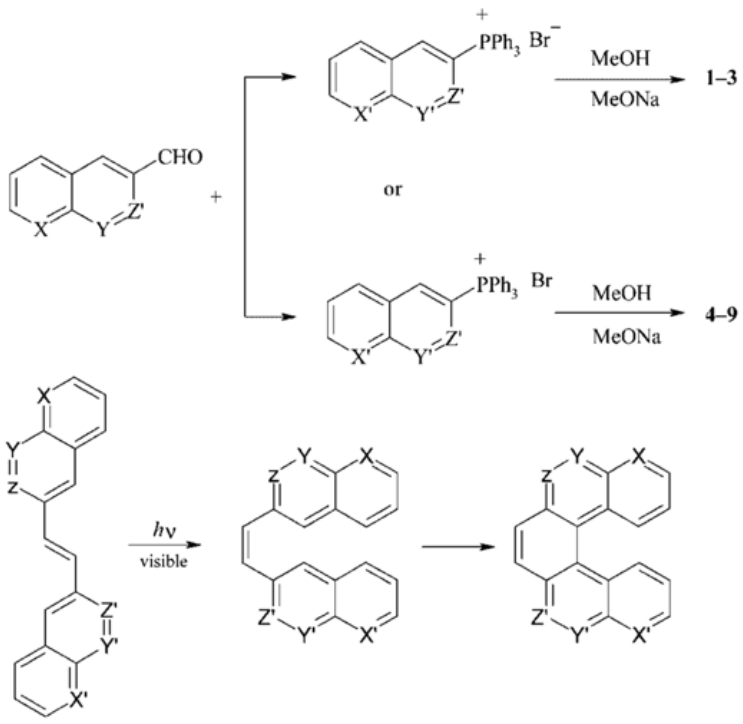

Figure 12: scheme of photochemical cyclization and precursor for the syntheses of Aza[5]helicene (Bazzini 2005)

Such Azahelicenes can be endowed with a tail containing a terminal double bond by heterocyclic aromatic substitution (Minisci 1993). The double bond is then used to introduce a thioester functionality, obtaining an alkyl chain that can react whit the nanostructured gold surface. This approach avoids risk of oxidation by air of the -S-H group (Zanchi 2018). The process is reported in Figure 13.

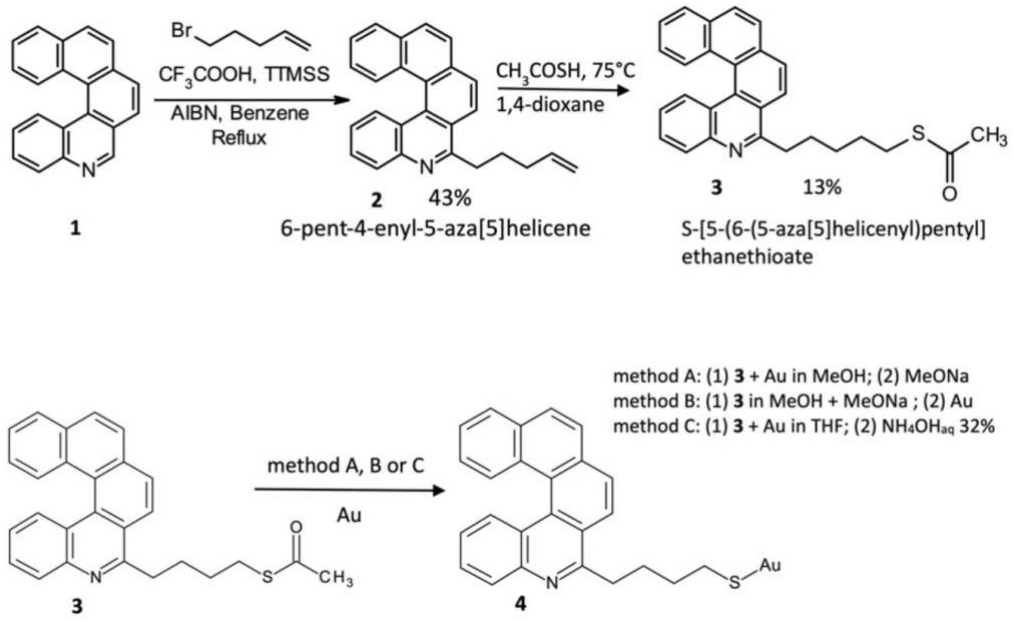

Figure 13: shematic process of chain addition to Aza[5]helicene and functionalization of nanostructured Au substrates by in situ thiol generation and S-Au bond formation (Zanchi 2018).

The models. To focus my study on the interaction between the selected AEDs and the functional core of the Azahelicenes (*i.e.*, the helical coil), I need to simplify out the alkyl chain that was introduced to anchor the Azahelicene to the gold substrates. This is necessary because in our simulation the two molecules are located in free space and the gold substrate is not present, so if the chain is not removed, it would bend on itself and change the interaction mechanism of our molecules. This simplification avoids the interactions between the chain and the AEDs. Such interactions are not likely, because in the real case the chain is already bonded to the substrate and is less reachable by the AED than the Azahelicene moiety. Moreover, this simplification allows us to focus our study on the most spectroscopically interesting part of the molecule, that is the π-conjugated part. Therefore, instead of the alkyl chain I decided to consider only a methyl group as can be seen in Figure 14. The energies of the two simplified geometries are shown in Table 8.

Table 8: energy of Aza[5]helicene and Aza[6]helicene

	E_2^0 *(kcal/mol)*
Aza[5]helicene	117.03
Aza[6]helicene	142.26

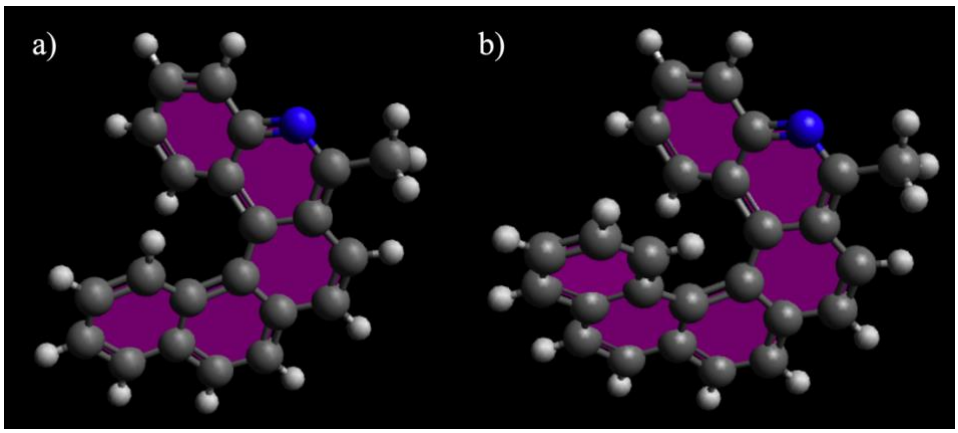

Figure 14: geometrical strucute of Azahelicenes. a) on the left the Aza[5]helicene. b) on the right the Aza[6]helicene

3.3. Early investigation of molecular interaction

After having determined the conformers of the selected AEDs, and the molecular structures of Aza[5]helicene and Aza[6]helicene, I can setup an early investigation on the molecular interaction between the Azahelicenes and the AEDs.

The first step is to create suitable Avogadro files that contain the two molecules that define a given dimer. Then, by using the *Auto Optimization Tool* of Avogadro, I can simulate the interaction process using the MMFF94s force field. In this way the program optimizes the geometry of the dimer and from the final optimized structure I can obtain the total energy (E_i).

From the energy data it is easy to recover the interaction energy (W_i) by simply subtracting the energy of the two isolated molecules ($E_0 = E_1^0 + E_2^0$) from the total energy (E_i) of the dimer, as reported below:

$$W_i = E_i - E_0 = E_i - E_1^0 - E_2^0$$

With this method is also possible exploring other interaction geometries for our dimers by changing the initial relative position of the molecules. This process must be done manually, of course. For this reason, this approach is not able to generate an acceptable number of simulations in a reasonable amount of time. So, this kind of "manual" simulation was only done in the early stage of this work in order to get more familiarity with these results and for a better understanding of the interaction energy (W_i) range (see Table 9) and of the possible kind of interaction geometries that we may expect. As we will see later on, to actually obtain thousands of simulations with a Monte Carlo approach, I had to use the OBSolv program.

Table 9: interaction energy (W_i) of the dimers. AED + Aza[X]helicene.

		Aza[5]helicene		*Aza[6]helicene*	
	E_1^0 *(kcal/mol)*	E_i *(kcal/mol)*	W_i *(kcal/mol)*	E_i *(kcal/mol)*	W_i *(kcal/mol)*
Carbamazepine	39.41	145.55	-10.89	169.82	-11.85
Clobazam	58.82	167.7	-8.15	190.81	-10.27
Clonazepam	134.54	221.65	-29.92	248.86	-27.94
Nitrazepam	131.60	234.49	-14.14	261.64	-12.22
Oxcarbamazepine	46.10	151.97	-11.16	176.11	-12.25
Perampanel	113.67	221.11	-9.59	245.44	-10.49
Phenytoin	-1.20	102.98	-12.85	127.82	-13.24

4. Using OBSolv to investigate interactions between CBZ and the two Azahelicenes

4.1. Carbamazepine (CBZ)

Carbamazepine (CBZ, chemical formula $C_{15}H_{12}N_2O$) is a widely used antiepileptic drug on the market and it is a first-generation AED. Nowadays it is no longer protected by a patent and thanks to its versatility and low cost it is widely used also in developing countries (Jallon 1997). Synthesized for the first time in 1965 has been extensively studied and his mechanism of action is relatively well understood. CBZ is slowly but well absorbed by oral administration. The plasma half-life is about 30 hours for patients who need it sporadically, if it is given repeatedly the plasma half-life shortens to about 15 hours due to the induced strong hepatic enzymes response (Pistaffa 2017). His molecular structure is tricyclic and fairly rigid. Upon rotation of the -$CONH_2$ group, CBZ displays two isomeric forms that are – from a geometric standpoint – enantiomers. Due to the low barrier to such internal rotation. ($8.2\ kcal/mol$, computed at B3LYP/6-31G(d,p) level)[1], CBZ is not a chiral compound. However, in our simulation, we can assume that the two isomers are different from the point of view of chiral recognition phenomena. This is a good approximation because OBSolv does not take in account thermal fluctuations and so the energetic barrier between the two isomers is never exceeded. The SERS spectrum of CBZ obtained by using Au substrates with controlled nanostructure shows a perfect correspondence of the peak position with respect to the normal Raman spectrum of CBZ, whose signals are also coherent with the position of the normal modes computed by DFT (Zanchi 2016). This is due to the weak interaction between CBZ and the Au substrate, which occurs only by physisorption rather than chemisorption. Hence the structure of CBZ is not modified by the absorption on the Au surface. The consequence is the loss of signal enhancement related to the chemisorption factor, in the order of 10^3 (Zanchi 2016). In principle, by functionalizing the nanostructured gold substrate with Azahelicenes one may increase the interaction between the plasmonic substrate and CBZ, and

[1] this energy barrier is similar to the values found in molecules that are known to undergo inversion at room temperature, such as Corannulene (Biedermann 1999).

consequently increase the SERS signal. By considering the different interaction of the two enantiomers of CBZ with the two enantiomers of the Azahelicenes, one may also obtain a theoretical proof of concept for the possible chiral recognition behavior of the SERS substrate. For those reason CBZ was selected for our purposes, and OBSolv was used to generate an appreciable number of Monte Carlo simulations concerning all the possible interactions between the two enantiomers of CBZ (A, B) and the two chiral forms of the Azahelicenes (A, B) – for both Aza[5]helicene and Aza[6]helicene. This led me to consider a total number of 8 possible interaction pairs, reported in tables Table 10 and Table 11.

Table 10: the name of the 4 interaction pairs for the case of Aza[5]helicene.

Aza[5]helicene	*CBZ*	
A	*A*	*AA[5]*
A	*B*	*AB[5]*
B	*A*	*BA[5]*
B	*B*	*BB[5]*

Table 11: the name of the 4 interaction pairs for the case of Aza[6]helicene.

Aza[6]helicene	*CBZ*	
A	*A*	*AA[6]*
A	*B*	*AB[6]*
B	*A*	*BA[6]*
B	*B*	*BB[6]*

4.2. Simulated interaction between CBZ and Azahelicenes

I used OBSolv to simulate the interaction of CBZ with Aza[5]helicene and Aza[6]helicene. I first assessed the interaction energy (W_i) and the geometry of the simulated dimers by running OBSolv locally (on my personal computer) and by considering a small number of simulations. This initial step was done to check the consistency of the data produced by OBsolv with respect to the results previously obtained by Avogadro with a manual approach. As discussed above CBZ has two isoenergetic conformers (A, B) that can be considered as enantiomers in our simulation. The Aza[5]helicene and Aza[6]helicene have also two enantiomers each (A, B). Hence, we have to consider the total number of combinations, that is 4 for each couple of molecules, namely AA, AB, BA, BB. I expect the cases AA and BB to be indistinguishable, as well as the cases AB and BA. This is due to the fact that AA is the mirror image of BB, and AB is the mirror image of BA. One of the aims of my thesis work is to assess chiral recognition with Azahelicenes. To check whether this is possible, we should find an appreciable difference in the interaction mechanism of the two classes of interactions, *i.e.*, AA/BB *vs*. AB/BA. In the next two tables (Table 12 & Table 13) I report the results from the 120 OBSolv simulations that have been obtained by running the OBSolv code locally, with no automatization. It is pleasant to verify that the results are in the correct expected range, in line with the values I have previously determined by using Avogadro on manually selected interacting pairs of Azahelicenes and CBZ.

CBZ-Aza[5]helicene

Table 12: interaction energy (W_i) calculated by OBSolv for the four cases (AA, BB, AB, BA) of CBZ-Aza[5]helicene.

Run	W_{AA} *(kcal/mol)*	W_{BB} *(kcal/mol)*	W_{AB} *(kcal/mol)*	W_{BA} *(kcal/mol)*
1	-12.79	-12.70	-12.03	-10.85
2	-10.56	-12.85	-10.43	-8.95
3	-9.41	-12.80	-12.08	-10.80
4	-11.63	-12.73	-11.31	-10.05
5	-10.07	-12.30	-12.03	-11.44
6	-9.74	-12.77	-11.79	-10.29
7	-10.47	-10.50	-12.05	-12.21
8	-10.77	-10.93	-12.11	-11.69
9	-10.48	-11.51	-11.99	-12.24
10	-12.79	-8.86	-11.74	-12.18
11	-9.65	-9.82	-10.51	-12.20
12	-8.89	-9.21	-12.19	-11.39
13	-11.04	-11.42	-11.77	-9.83
14	-11.13	-9.61	-11.84	-9.36
15	-9.94	-11.15	-10.48	-12.15
MED	-10.62	-11.28	-11.62	-11.04
DEV	1.13	1.42	0.63	1.12

CBZ-Aza[6]helicene

Table 13: interaction energy (W_i) calculated by OBSolv for the four cases (AA, BB, AB, BA) of CBZ-Aza[6]helicene.

Run	W_{AA} *(kcal/mol)*	W_{BB} *(kcal/mol)*	W_{AB} *(kcal/mol)*	W_{BA} *(kcal/mol)*
1	-9.79	-10.29	-11.94	-13.80
2	-9.77	-9.68	-10.43	-12.24
3	-9.96	-11.55	-10.34	-12.67
4	-10.41	-10.57	-10.23	-10.97
5	-10.35	-10.03	-12.29	-11.16
6	-10.75	-9.34	-12.30	-13.52
7	-9.45	-9.45	-11.61	-10.92
8	-10.39	-9.53	-9.80	-11.69
9	-11.38	-9.65	-11.88	-8.94
10	-10.48	-9.81	-12.02	-11.91
11	-11.62	-11.06	-11.80	-9.61
12	-10.13	-10.42	-11.36	-12.34
13	-10.78	-10.35	-11.67	-13.43
14	-11.77	-10.11	-12.56	-11.27
15	-11.64	-8.44	-10.13	-10.26
MED	-10.58	-10.02	-11.36	-11.65
DEV	0.74	0.75	0.92	1.42

As we can see from this small batch of data, we cannot find any appreciable difference between the two classes of enantiomeric interactions (*e.g.*, AA *vs*. AB). As it will be clearer later on, this is just due to the lack of data due to poor statistical sampling. Therefore, to obtain an appreciable number of OBSolv simulations it was necessary to run the code on the HPC servers available to the FunMat research group at the *Dipartimento di Chimica, Materiali ed Ingegneria Chimica (Politecnico di Milano)*. This allowed to drastically reduce the time required for each simulation, and also allowed to run in parallel a significant number of independent OBSolv instances. We decide to run 2000 simulations for each enantiomeric case, for a grand total of 16000 simulations. Of course, such amount of data cannot be manually analyzed as I have done so far. For this reason, it was necessary to develop a Matlab code for the data analysis.

5. Data analysis with MatLab

To process the sizeable amount of data provided by OBSolv it was necessary to write a suitable Matlab code. For this purpose, first of all, I had to determine the relevant data available from OBSolv simulations and examine their organization. The data are exposed in matrix form, as a multivariate dataset, where each row contains all the information of a single Monte Carlo run for a given dimer (*e.g.*, CBZ(A)-Aza[5]helicene(B)). Each rows of the dataset matrix provide the results of independent runs of the OBSolv code – for the same kind of dimer. For a better representation of the situation, I report the details of the format of the dataset matrix in Table 14. As we can see, the first column of the matrix contains the unique ID code of each simulated dimer – this allows to quickly identify a given specific output of OBSolv for possible subsequent analysis. In the second and third columns one finds the energy of the two isolated molecules (Azahelicene, E_2; CBZ, E_1). These two energy values are computed by OBSolv by taking the given structure of dimer and considering the structure of one molecule (in the dimer) in the absence of the other molecule – this calculation was implemented in OBSolve for internal checking and debugging purposes and provides anyway a useful output. The program does not fully optimize the geometry of such isolated species, so we do not obtain E_1^0 and E_2^0. Nevertheless, for rigid molecules, the difference is small enough, and we may assume as a first approximation that $E_1 \approx E_1^0$ and $E_2 \approx E_2^0$.

The 4th column contains the interaction energy of the dimer (W_i). Starting from the 5th column onwards, we find the Cartesian coordinates of the structure of the dimer. The position in space of each atom is identified by three consecutive columns in which we find the three X, Y, and Z coordinates (this is the so-called 3N Cartesian format, widely used in theoretical vibrational spectroscopy).

From this matrix the Matlab code extracts the interaction energy (W_i) and the Cartesian coordinates required to compute the geometric parameters that we will define later on to describe the interaction geometry of each dimer. The aim of this analysis is to inspect the interaction energy (W_i) as a function of selected geometric parameters that are related to the relevant interaction mechanisms that can be found in our specific case.

Of course, I do not expect to be able to compute the total interaction energy just based on the most relevant contributions. This is not possible, even for small molecules (Mark 1994). However, I expect to use those parameters to identify the most relevant geometries of the interacting molecules, so to group in the output of OBSolv a few different families of interaction structures. In this way, I can also establish if chiral recognition is possible and if there is an advantage in using Aza[5]helicene instead of Aza[6]helicene (or vice versa).

Table 14: exemple of the structure of the multivariable dataset generated by OBSolv

ID	E_2	E_1	W_i	*Cartesian coordinates …*
10000	117.03	39.41	-10.89	…

5.1. Molecular interactions

As we know, it is not possible to reliably model the total interaction mechanism between two relatively complex shaped molecules based only on the most relevant interaction contributions (Mark 1994). However, in my case the total interaction is already calculated by OBSolv through molecular mechanics (MMFF94s force field). Hence, the aim of my work is not to recover (or to model) the total interaction energy. Instead, I want to define a selection of suitable geometrical parameters to highlight the most relevant kind of interactions for our application. From the inspection of the results from the early calculations carried out manually with Avogadro, it is evident that the most relevant interactions in the CBZ-Azahelicene dimers are Hydrogen bonds and π-stacking. Therefore, I have defined three parameters related to those interactions, namely, the N…H distance, the N…HN angle and the distance between the centers of selected aromatic rings (Centroid distance). This approach gives the possibility to describe the geometry of the possible dimers by using simple geometrical parameters that are related to the nature of the interaction itself. I introduce below the details of these kind of interactions and the definition of the associated geometric descriptors.

5.1.1. Hydrogen bonds

Hydrogen bonds are one of the most relevant interactions in drug molecules and are formally represented as XH···Y. The electronegative atom covalently bonded to hydrogen (X) is named proton donor; the atom non-covalently bonded to hydrogen (Y) is named proton acceptor; the hydrogen itself (H) is named protic center.

The hydrogen bond can be described as an electrostatic dipole-dipole interaction with some covalent bonding features like the strong interaction, the directionality and the interatomic distances shorter than the sum of the Van der Waals radii in the range 1.6 – 2 Å. The covalent features of the hydrogen bond are proportional to the electronegativity of the proton donor (X). In our case, the NH_2 group of CBZ (X) and N atom of Azahelicene (Y) strongly interact but not always a "real" hydrogen bond is formed because other kind of interactions prevent the two interested atoms (X, Y) to be close enough. In some geometries of interaction this bond will be more similar to a dipole-dipole interaction without the covalent contribution.

Hydrogen bonds can vary in strength from 0.2-0.5 kcal/mol to 38 kcal/mol as in the case of the HF_2^- ion (Larson 1984, Emsley 1980). In our case the interaction is represented as (CBZ side) N−H···:N (Azahelicene side). The order of magnitude of the interaction energy can be inferred from, *e.g.*, the Ammonia-Ammonia interaction, whose reported hydrogen bond energy is 3.1 kcal/mol (Morrison 2003).

5.1.2. π-stacking

π-π interaction is a direct attractive non-covalent interaction between aromatic rings that is very relevant in drug intercalation (Kryger 1999, Wakelin 1986). Aromatic rings can interact in several different conformations: Face-to-Face, offset staked, T-shaped conformation (Figure 15). The strength of the interaction is proportional to the overlapping of p orbitals in π-conjugated systems, so it becomes stronger as the number of π-electrons increases (C. Hunter 2001).

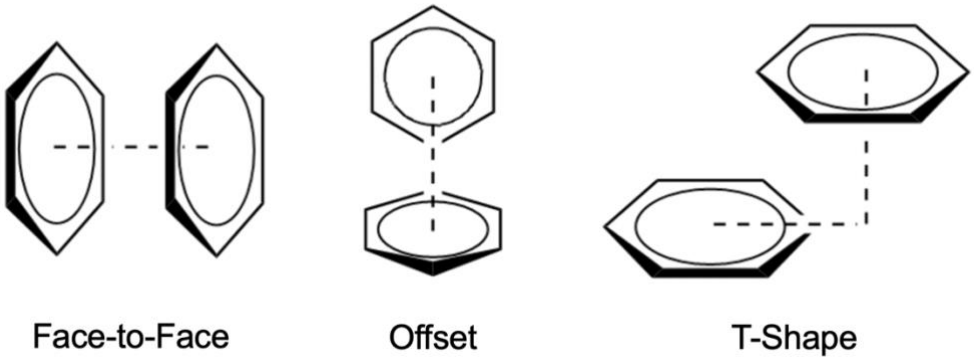

Figure 15: schematic representation of the three possible π-stacking configurations

As we have seen, the interaction can take place in three different conformation each of which is favored by different combination of forces. In order to study the real nature of π-π interaction, Hunter and Sanders proposed an electrostatic interaction model (Hunter and Sanders 1990). This model is based on two key points:

- it considers the π-electrons and the σ-bonding systems separately.
- it considers the net favorable π-π accretive interactions as the result of π-σ attractions that overcome the π-π repulsions.

So, the π-systems can be imagined as a sandwich of the positively charged σ-framework between two negatively charged π-electrons clouds where the attraction contribution is higher than the repulsion. The electrostatic interactions between such systems act as a function of the orientation. As result, the Face-to-Face and offset stacked conformation will always meet repulsion and attraction, respectively, while T-shape conformations will afford repulsion and attraction effects at the same time according to different orientations (Zhao 2015).

In general, Face-to-Face conformation is unfavorable while T-shaped conformation is favored. In the case of CBZ-Azahelicene interactions the T-shaped conformation is almost never achieved, for different reason. First of all, in the CBZ-Azahelicene systems there are other kind of interactions to take in account (i.e., hydrogen bonds). Secondly, the high number of aromatic rings and the rigidity of both CBZ and Azahelicenes does not allow T-shaped π-stacking to be overall energetically favorable. The position of an aromatic ring can be naturally represented in terms of the ring centroid, that is the center of the six-membered ring (Thomas 2002).

5.2. Definition of the geometrical parameters to describe CBZ-Azahelicene interactions

After having identified the main possible interaction mechanisms, I have then defined the associated geometric parameters. For the hydrogen bond interaction, I selected the N…H-distance (NH) and the N…HN-angle (NHN). For the π-stacking I defined the Centroid distance (C). To obtain the data associated with such parameters it was necessary to design three different Matlab scripts, one for each parameter. In general, those scripts calculate the related parameter, from the multivariate dataset in matrix form that stores the results of the several OBSolv runs. Thereafter, the Matlab scripts plot the interaction energy (W_i) as a function of the computed parameters. I also designed a script that plots the interaction energy in a 3D color-coded graph, as a function of the three geometrical parameters (associated to the three Cartesian axes); the interaction energy is encoded as a color value in a gradient palette. The Matlab code that I wrote for the data analysis is specific to the case of CBZ-Azahelicene. However, it is important to underline that the code is modular, and it was designed with a bottom-up approach. Hence, it is very easy to define new parameters, for other kinds of molecular interactions that are specific of other AEDs.

5.2.1. N…H distance

This parameter characterizes the hydrogen bond interaction. In our case it is defined by the distance between the nitrogen atom in the second ring of Azahelicene, and one of the hydrogens of the NH_2 group of CBZ (Figure 16). In the NH_2 group of CBZ there are two hydrogen atoms that can lead to hydrogen bond interaction. Hence, to properly define the N…H distance parameter, the Matlab code needs to calculate both distances ($N{-}H_1{\cdots}{:}N$, $N{-}H_2{\cdots}{:}N$) and then chooses the smallest value.

5.2.2. N…HN angle

The second parameter describing the hydrogen bond is the bond angle formed by the $N{-}H{\cdots}{:}N$ atoms (Figure 16). This is computed using the hydrogen atom that is closer to the N atom of Azahelicene (such hydrogen atom was selected in the calculation of the N…H distance – see above). The N…HN angle describes the directionality of the bond, that is maximal for 180°.

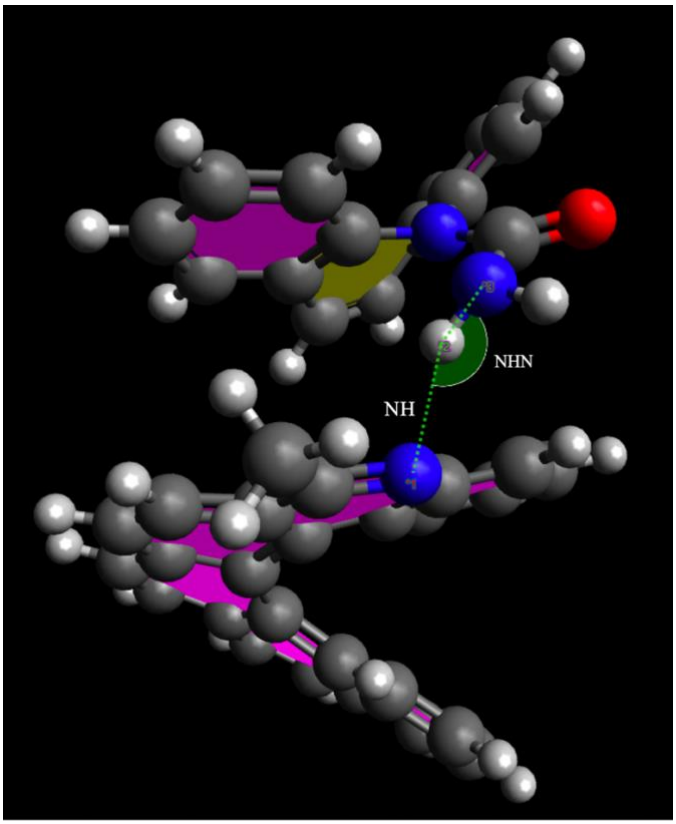

Figure 16: representation of the two parameters used to describe the hydrogen bond interaction: the N…H distance and the N…HN angle.

5.2.3. Centroid distance

The Centroid distance allows to characterize the π-stacking interaction. First of all, I define the centroid as the center of selected six-membered rings. Such rings (Figure 17) are the two rings of CBZ and all the rings of the interacting Azahelicene. When π-stacking interaction is effective, the rings of the two molecules are proximal, and the distance is minimal. Therefore, the Matlab code that I designed computes, for each of the two rings of CBZ, the distance between its centroid and the centroids of all the rings of the Azahelicene, and finally takes the minimum value. The minimum distance values computed for the two rings of CBZ are then summed together to define the Centroid distance.

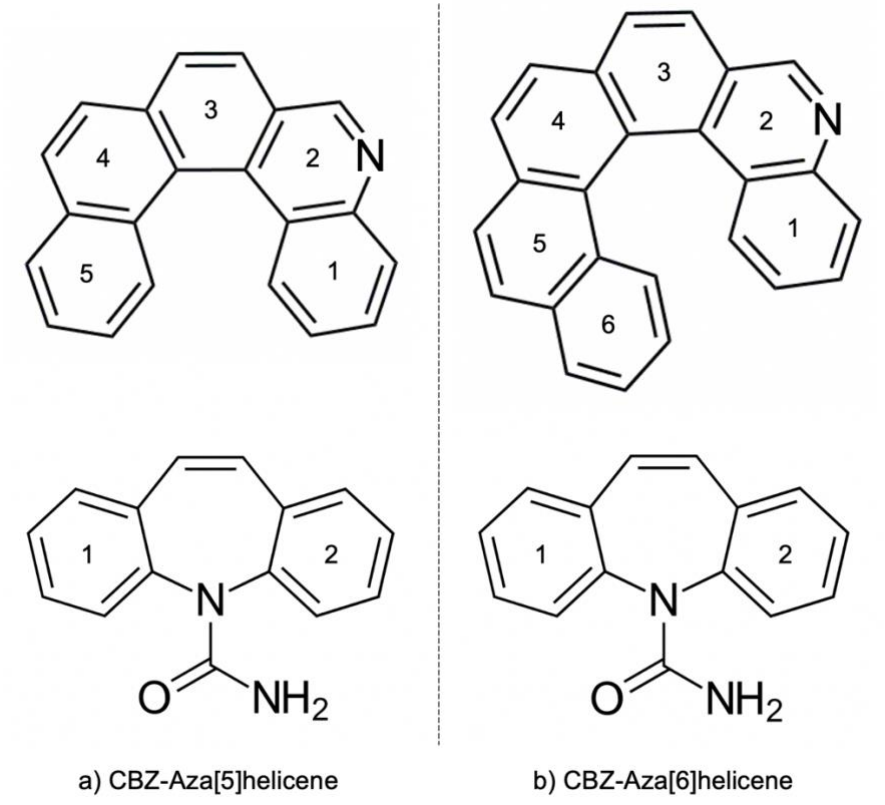

Figure 17: representation of the centroids of the aromatic rings of CBZ and Azahelecenes.

5.3. The interaction between CBZ and Aza[5]helicene

I describe below all the steps that I made to analyze the data produced by OBSolv for the investigation of the CBZ-Aza[5]helicene interaction. First of all, after the definition of the three geometrical parameters described above, I wrote a Matlab code for each parameter, using which I could plot the interaction energy (W_i) as a function of the selected parameter. From such plots I could identify absolute minima of the interaction energy and other local minima of interest. Then I cross-checked those points for all the three parameters, so to establish if such minima were common in all plots. At this point, I realized that some of those minima were effectively found in all the three plots, whereas some other points of interest were only defined in one or two plots (*i.e.*, they were recognized as minima just as a function of one or two variables). Therefore, I used the ID code to identify the simulation associated with those minima with one or more missing variables and then, by using the Matab code selectively only on those simulations, I calculated the missing parameters. In this way I could find the three geometrical parameters and the interaction energy (W_i) for all the points of interest. I then sorted them by increasing interaction energy, labelled them with numbers and plot them in suitable graphs. The last step was to use the ID code to extract the simulated file.xyz of the dimer and use Avogadro to visualize the associated geometry. This workflow has been repeated for all four possible combinations: AA, BB, AB, BA. In the next sections I discuss in detail the results of this analysis for the four interaction pairs.

5.3.1. AA & BB pairs of the CBZ-Aza[5]helicene interaction

Figure 18 and Figure 19 plot the interaction energy of the AA & BB pairs, respectively, as a function of the three geometric parameters selected for describing the intermolecular interactions, namely the N…H-distance (NH), the N..HN-angle (NHN) and the Centroid distance (C). Each point in this three-dimensional (3D) space is associated with a specific simulated dimer identified by its ID code. The color of each point in the plot is mapped as a function of the interaction energy (W_i). In Figure 18 and Figure 19 one observes seven regions of clustered points (identified by numerical labels) in which are located the points of interest for the description of the geometry of the intermolecular interactions. As one can see, the AA and BB case are almost identical, and the cluster regions are located within the same coordinate and energy range. This fact perfectly follows the expectation that the AA and BB cases must be analogous, as the two dimers are mirror images of each other. This point will be described with the related geometries of the dimers of interest at the end of this section.

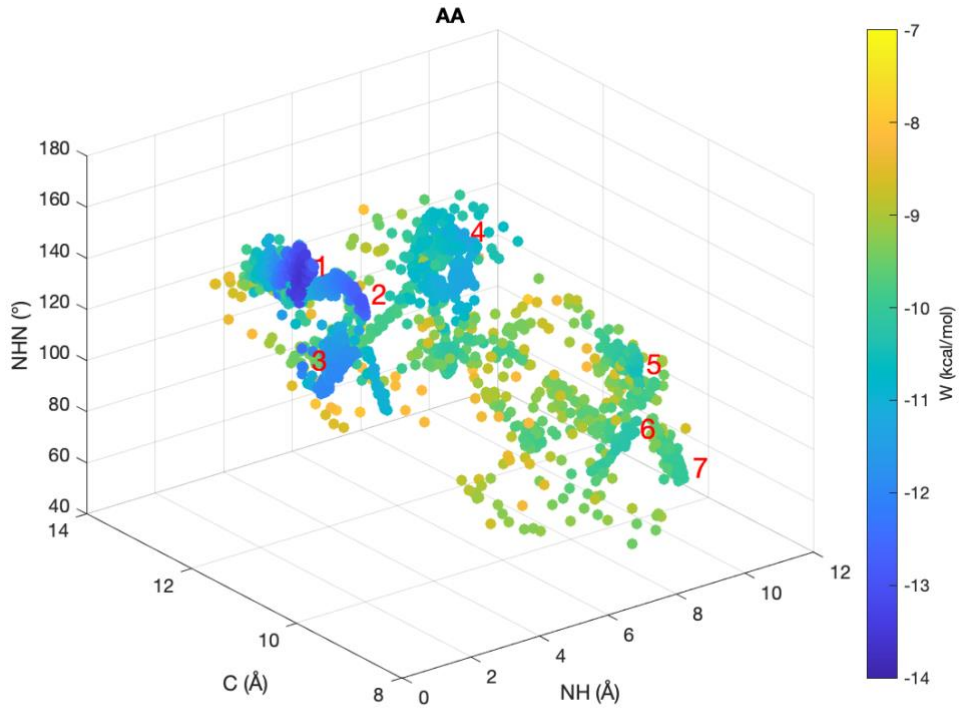

Figure 18: 3D graph of geometrical parameters and interaction energy (W_i) for the AA case of CBZ-Aza[5]helicene.

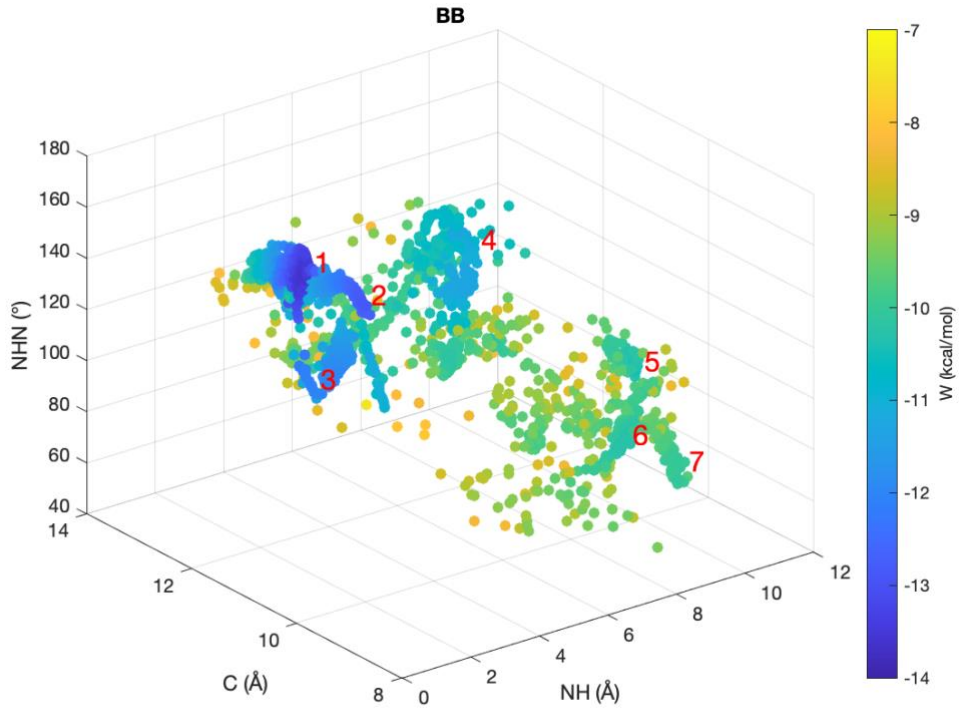

Figure 19: 3D graph of geometrical parameters and interaction energy (W_i) for the BB case of CBZ-Aza[5]helicene.

To better characterize the location of such clusters of points and identify the different interaction energy minima, one can examine in detail a series of 2D graphs, where the interaction energy (W_i) is plotted as a function of each geometrical parameter.

The **N…H-distance (NH)** – Figure 20 and Figure 21 – was introduced for characterizing the hydrogen bond interaction. As we know, the typical interaction length of an hydrogen bond is between 1.6 Å and 2 Å; therefore the only point where such interaction is present corresponds to the dimer conformation labelled 1. However, the hydrogen bond interaction is clearly still relevant also for the conformations of dimers 2 and 3. On the contrary, for conformations 5, 6 and 7 the hydrogen bond contribution is not relevant and other kinds of interactions should be considered for justifying the local minimum in the interaction energy. The conformation 4 is peculiar. As will be discussed later on by inspecting the structure of the different dimers, in dimer 4 the NH_2 group of CBZ creates an interaction different from an hydrogen bond, and such kind of interaction is not present in any other interaction geometry.

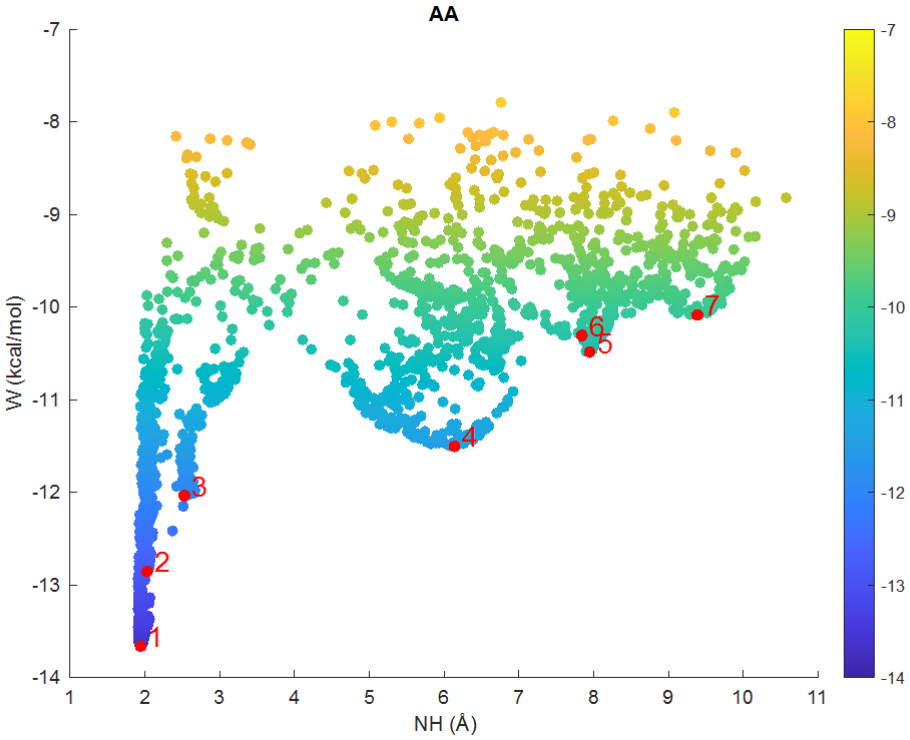

Figure 20: graph of the interaction energy (W_i) as a function of the N…H distance for the AA case of CBZ-Aza[5]helicene.

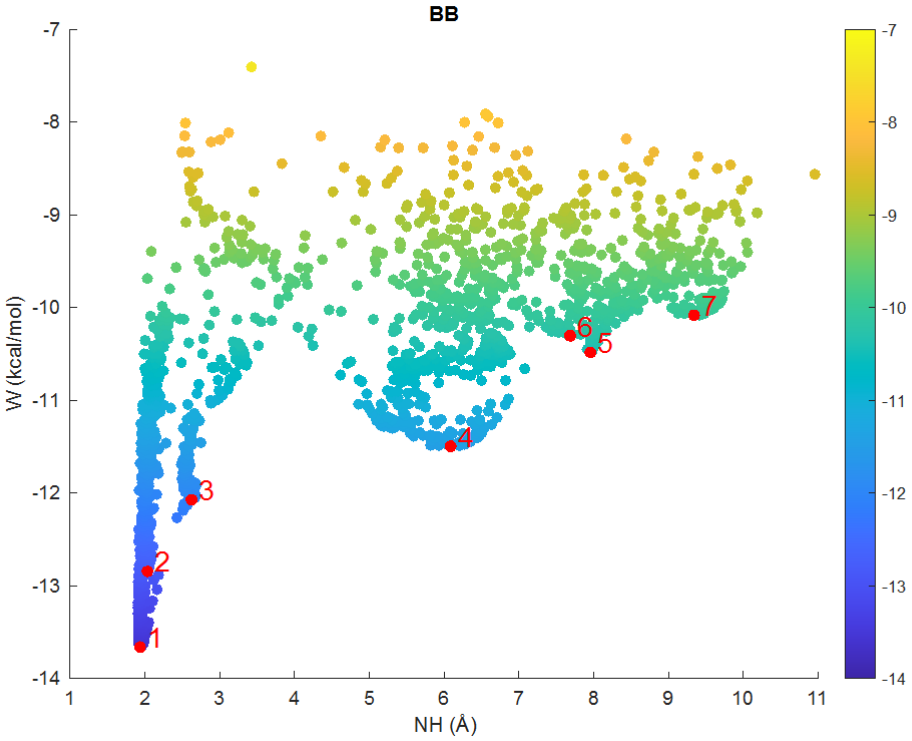

Figure 21: graph of the interaction energy (W_i) as a function of the N…H distance for the BB case of CBZ-Aza[5]helicene.

The second parameter defined for characterizing the hydrogen bond is the **N…HN-angle (NHN)** – Figure 22 and Figure 23. This parameter does not take into account geometries where the N…H distance is too high. In such cases, the contribution of the hydrogen bond is not relevant. Therefore, the conformations of dimers 4, 5, 6 and 7 are not considered. The conformations of interest for the discussion of the N…HN angle are those of the dimers labeled 1, 2 and 3. The first two display a similar angle, between 155° and 160°, whereas the third shows a much smaller angle, between 110° and 115°. So, we can say that:

- Dimer 1 has a good directionality and the interatomic distances is shorter than the sum of the Van der Waals radii. Hence, the hydrogen bond is present.
- Dimer 2 has just a good directionality. The interaction is present (weaker), and we can define a sort of hydrogen bond.
- Dimer 3 has the lowest directionality and the hydrogen bond contribution is significantly less effective than the previous two cases.

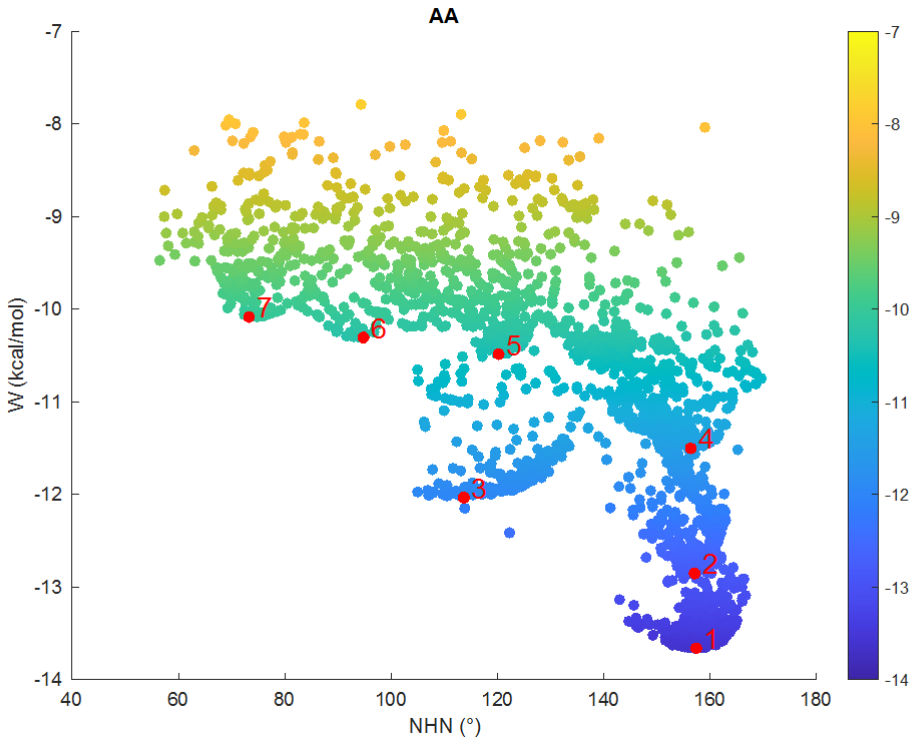

Figure 22: graph of the interaction energy (W_i) as a function of the N…HN angle for the AA case of CBZ-Aza[5]helicene.

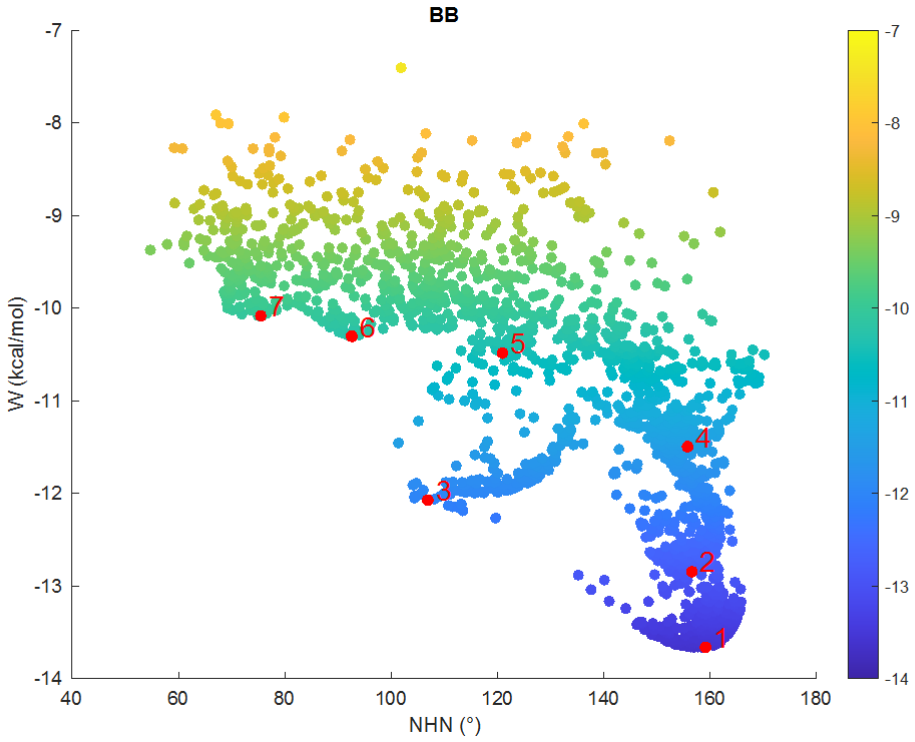

Figure 23: graph of the interaction energy (W_i) as a function of the N…HN angle for the BB case of CBZ-Aza[5]helicene.

To identify π-stacking interactions I have introduced the **Centroid distance (C)**. Here we can see in Figure 24 and Figure 25 that the conformations of dimers 1, 2, 3 and 4, that are the more energetically stable, are found at C values between 10 Å and 12 Å. Instead, in dimers 5, 6 and 7 the C values are below 9 Å. We have already established that for such three conformations the hydrogen bond is not relevant. Hence, we can expect that the dimers 5, 6 and 7 are associated to geometries that maximize π-stacking but minimize the NH_2 group contribution, thus leading to higher interaction energy. The dimers 1, 2, 3 and 4 are associated to geometries where we can find good compromises between different interactions, and this process leads to more stable configurations and lower interaction energy.

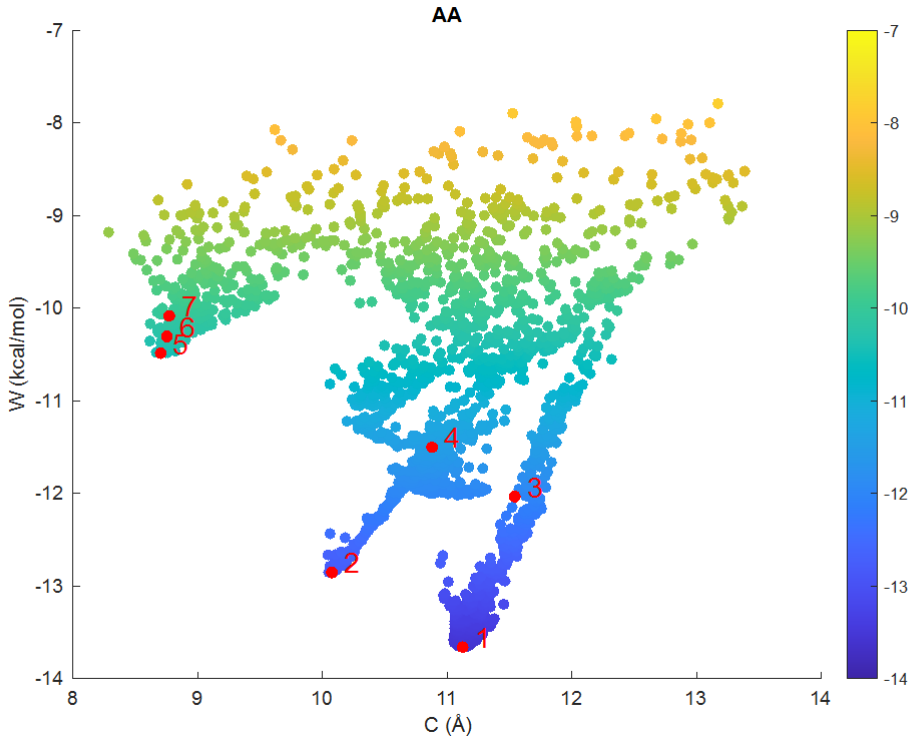

Figure 24: graph of the interaction energy (W_i) as a function of the Centroid distance (C) for the AA case of CBZ-Aza[5]helicene.

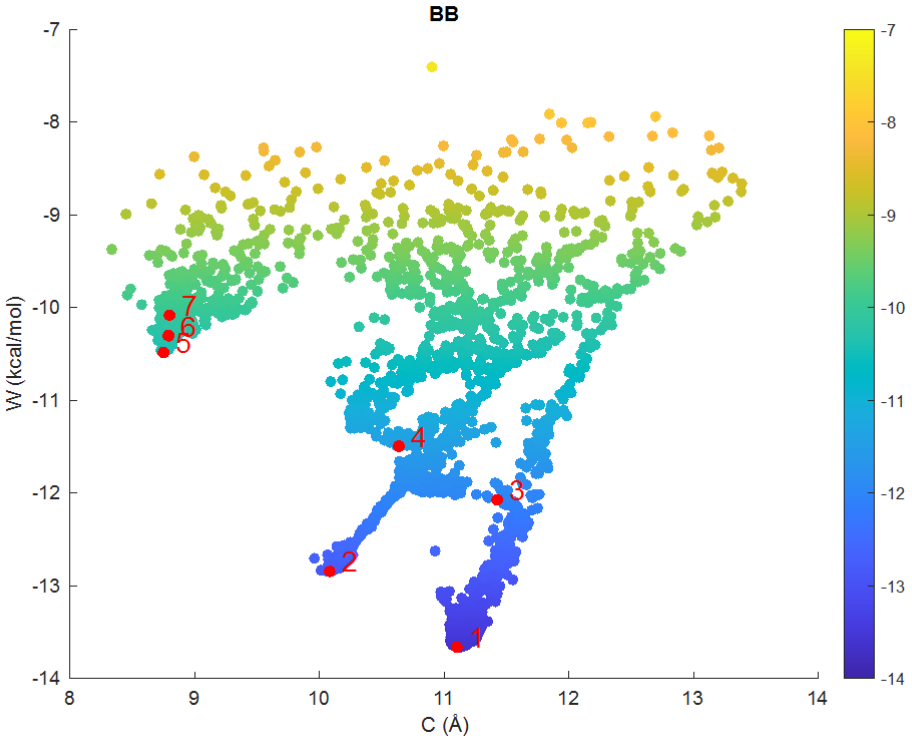

Figure 25: graph of the interaction energy (W_i) as a function of the Centroid distance (C) for the BB case of CBZ-Aza[5]helicene.

To conclude the discussion of the plots of the interaction energy, I report in Table 15 and Table 16, for each AA and BB dimer, the ID code, the interaction energy (W_i), and the three geometrical parameters that describe the intermolecular interactions.

Table 15: geometrical parameters, interaction energy (W_i) and ID for each dimer configuaration in the case BB CBZ-Aza[5]helicene. The ID is reported for archive purposes.

AA[5]	*ID*	*W_{AA} (kcal/mol)*	*N…H (Å)*	*N…HN (°)*	*C (Å)*
1	141695	-13.66	1.943	157	11.12
2	170195	-12.85	2.029	157	10.07
3	126936	-12.03	2.525	114	11.54
4	101874	-11.50	6.137	156	10.88
5	59473	-10.49	7.944	120	8.70
6	101343	-10.30	7.840	95	8.75
7	159806	-10.08	9.384	73	8.77

Table 16: geometrical parameters, interaction energy (W_i) and ID for each dimer configuaration in the case BB CBZ-Aza[5]helicene. The ID is reported for archive purposes.

BB[5]	*ID*	*W_{BB} (kcal/mol)*	*N…H (Å)*	*N…HN (°)*	*C (Å)*
1	178795	-13.66	1.935	159	11.10
2	75171	-12.84	2.034	157	10.08
3	98725	-12.07	2.622	107	11.42
4	65925	-11.49	6.086	156	10.64
5	62795	-10.48	7.953	121	8.75
6	84513	-10.30	7.684	93	8.79
7	102475	-10.08	9.339	75	8.80

Here below, I examine the geometries of the dimers that have been identified as local minima of the interaction energy.

Dimer 1: the NH_2 group of CBZ is forming a hydrogen bond with the N atom in the Aza[5]helicene (Figure 26). Also π-stacking is contributing to some extent to the total interaction. As expected, the AA and BB configurations are mirror images, and this is confirmed for all the dimers below.

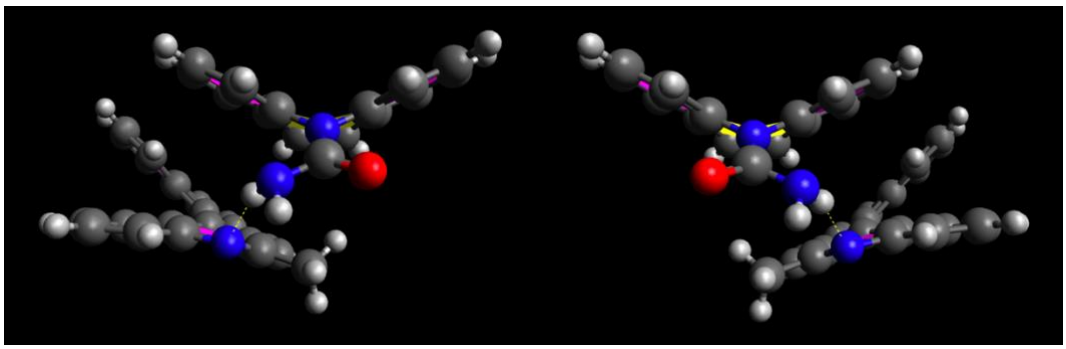

Figure 26: visual representation of the interacting Dimer 1. On the left AA and on the right BB. For the CBZ-Aza[5]helicene case.

Dimer 2: as one can see in Figure 27, in this geometry the hydrogen bond is present but weaker, and π-stacking is also present to some extent.

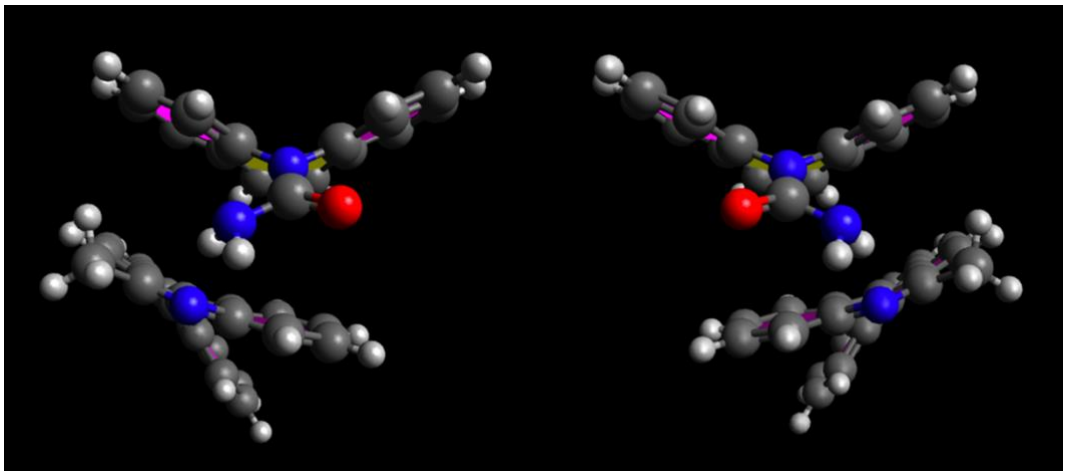

Figure 27: visual representation of the interacting Dimer 2. On the left AA and on the right BB. For the CBZ-Aza[5]helicene case.

Dimer 3: as shown in Figure 28, the hydrogen bond is less directional here than in the previous cases, but his contribution is still relevant, together with the contribution from π-stacking.

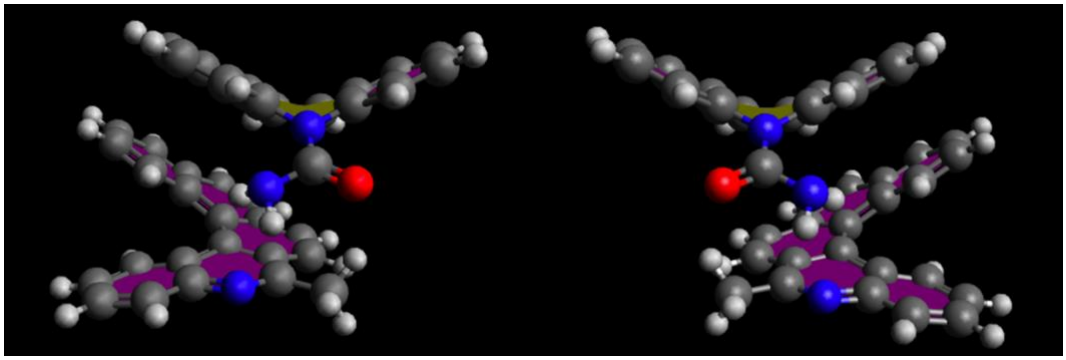

Figure 28: visual representation of the interacting Dimer 3. On the left AA and on the right BB. For the CBZ-Aza[5]helicene case.

Dimer 4: as one can see in Figure 29, the NH_2 group of CBZ is located away from the N atom of Aza[5]helicene, hence hydrogen bond is not possible. The NH_2 group is placed above one of the aromatic rings. This location of the NH_2 group leads to some sort of NH_2-π interaction that contributes to the intermolecular energy.

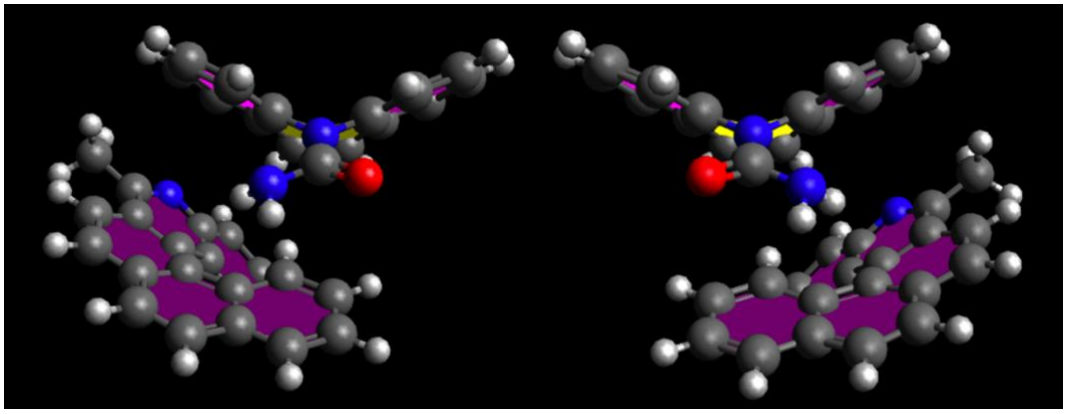

Figure 29: visual representation of the interacting Dimer 4. On the left AA and on the right BB. For the CBZ-Aza[5]helicene case.

Dimers 5, **6** and **7** can be seen in Figure 30, Figure 31 and Figure 32 respectively. In these three dimers the hydrogen bond contribution is irrelevant, because CBZ is rotated by 180° with respect to the previous four cases. Hence the NH_2 group is located too far away from the nitrogen of Aza[5]helicene to actually lead to any sort of hydrogen bonding contribution. The only possible contribution to the interaction energy of such dimers is from π-stacking. As reported above (Table 15 and Table 16), the Centroid distance (C) is below 9 Å for all the three configurations. The interaction energy (W_i) is much higher than the previous geometries because π-stacking is a weaker interaction than hydrogen bonding (for aromatic molecules of this size).

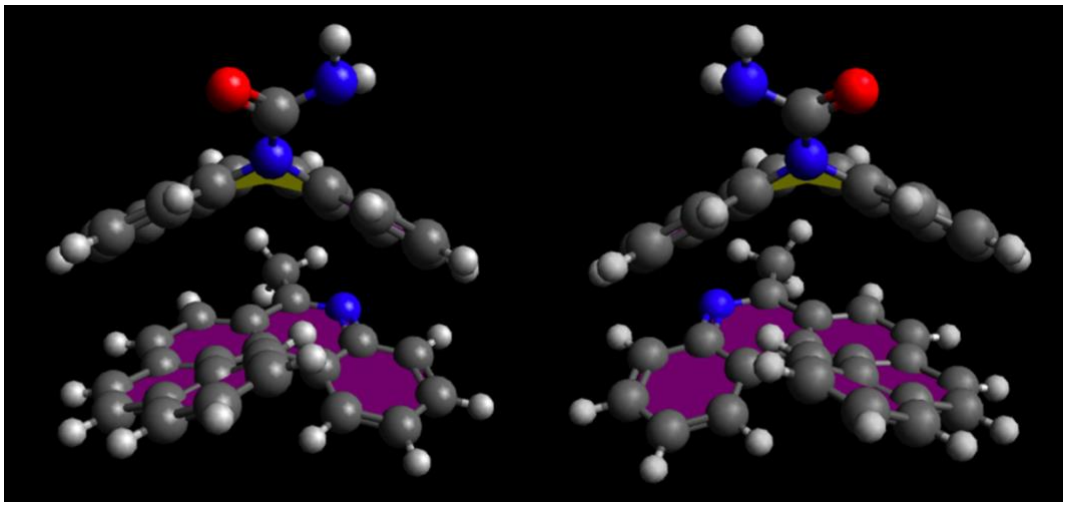

Figure 30: visual representation of the interacting Dimer 5. On the left AA and on the right BB. For the CBZ-Aza[5]helicene case.

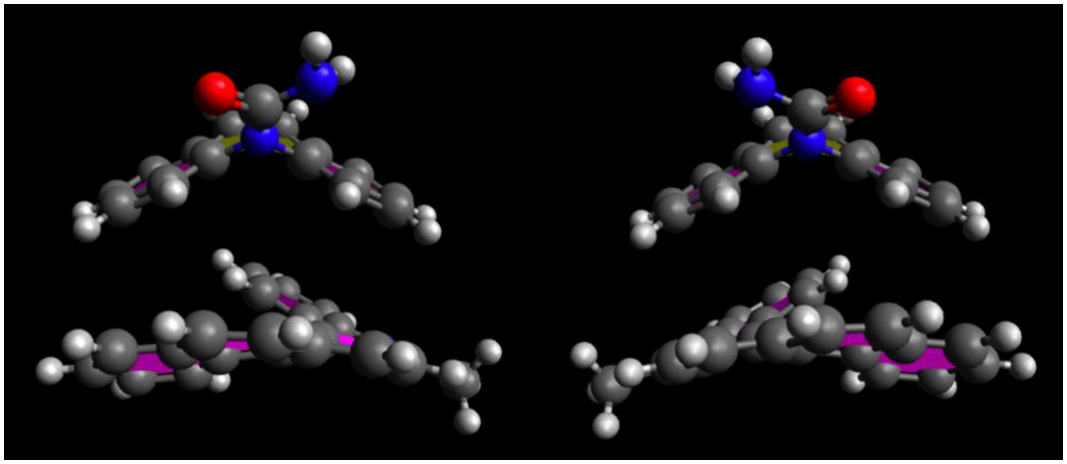

Figure 31: visual representation of the interacting Dimer 6. On the left AA and on the right BB. For the CBZ-Aza[5]helicene case.

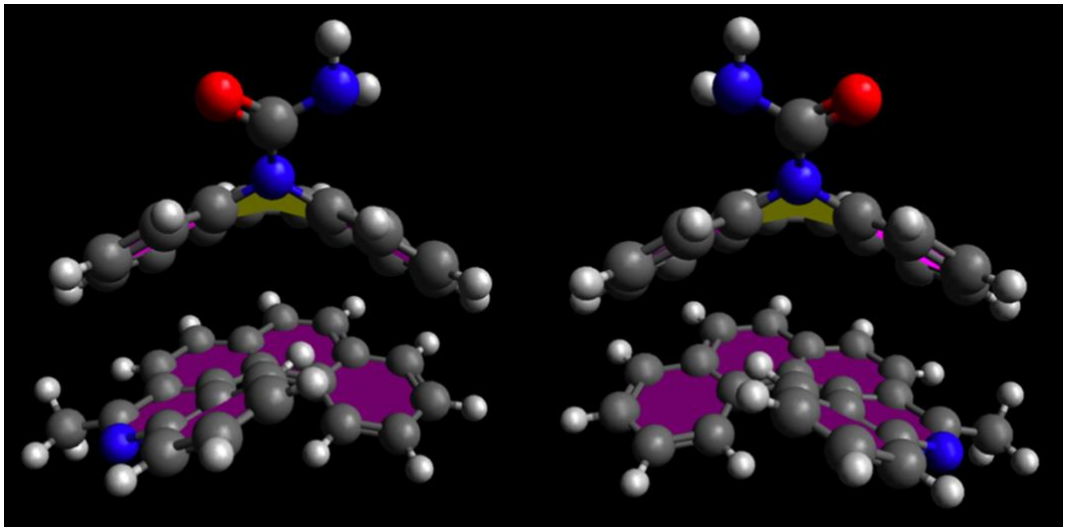

Figure 32: visual representation of the interacting Dimer 7. On the left AA and on the right BB. For the CBZ-Aza[5]helicene case.

By analyzing both the plots of the interaction energy and the three-dimensional representations of the dimers, one can confirm that AA and BB cases are perfectly analogous because they are mirror images of each other. This was expected from the theoretical point of view (based on stereochemistry), but it was necessary to check the correctness of the OBSolv code by carrying out the analysis anyway.

5.3.2. AB & BA pairs of the CBZ-Aza[5]helicene interaction

Figure 33 and Figure 34 show the 3D graphs of the two cases of interest. As expected, based on stereochemistry, the two plots are almost identical, and the local minima of the interaction energy cluster in the same coordinate and energy range. I have identified seven clusters, and there is no relation between such clusters and those identified previously for the AA & BB cases. This is due to the fact that the interacting enantiomers are different than before, and so the molecules can achieve new kinds of geometrical arrangement in forming the dimers.

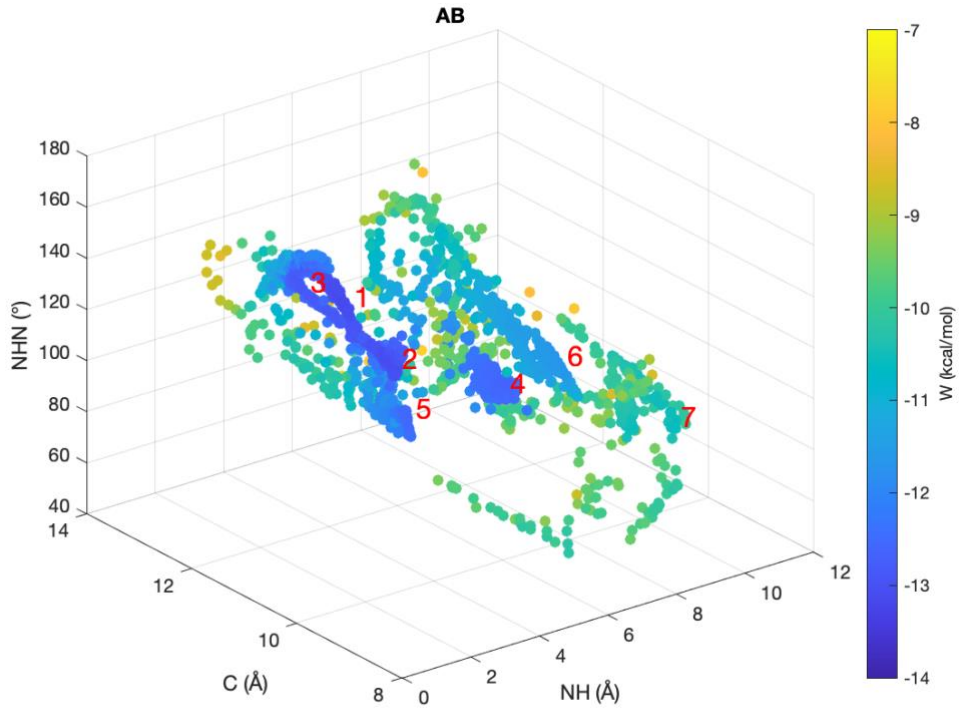

Figure 33: 3D graph of geometrical parameters and interaction energy (W_i) for the AB case of CBZ-Aza[5]helicene.

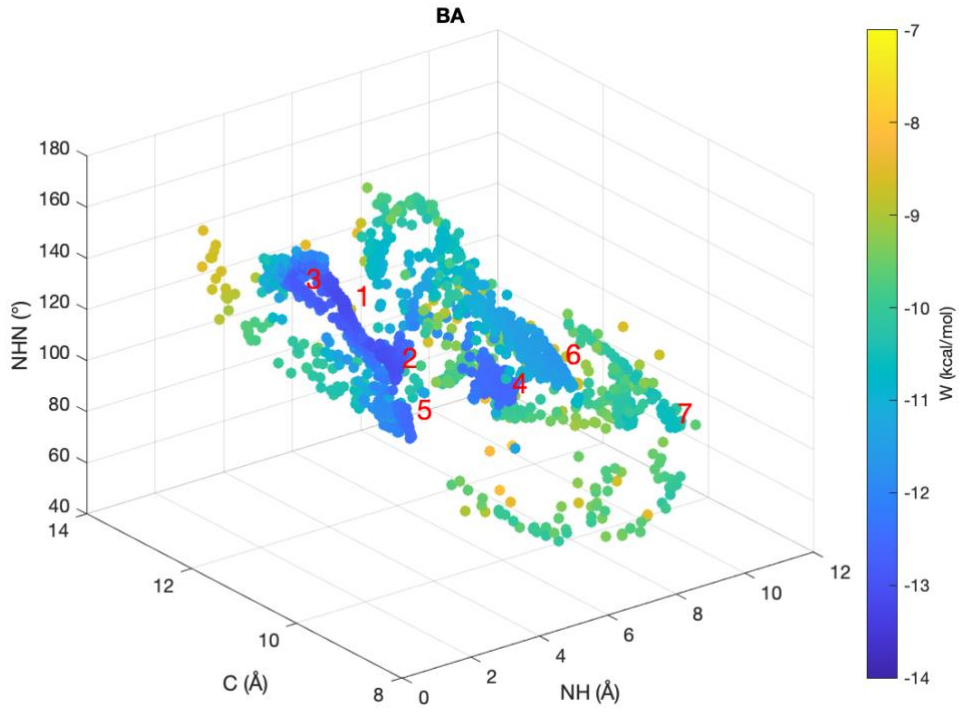

Figure 34: 3D graph of geometrical parameters and interaction energy (W_i) for the BA case of CBZ-Aza[5]helicene.

I discuss here below the plots of the interaction energy as a function of the different geometric parameters introduced before.

N…H-distance (NH) – Figure 35 and Figure 36. The dimer conformations 1 and 3 display N…H distances below 2 Å, which highlights the presence of an hydrogen bond; also for the dimers 2 and 5 this interaction is relevant, but it does not lead to the formation of a real hydrogen bond because the distances between the interested atoms are too high. The conformations of dimers 4 and 6 have a peculiar behavior that can be explained with another kind of interaction mechanisms. Clearly, dimer 7 has no hydrogen bond interaction, because the N…H distance is too large.

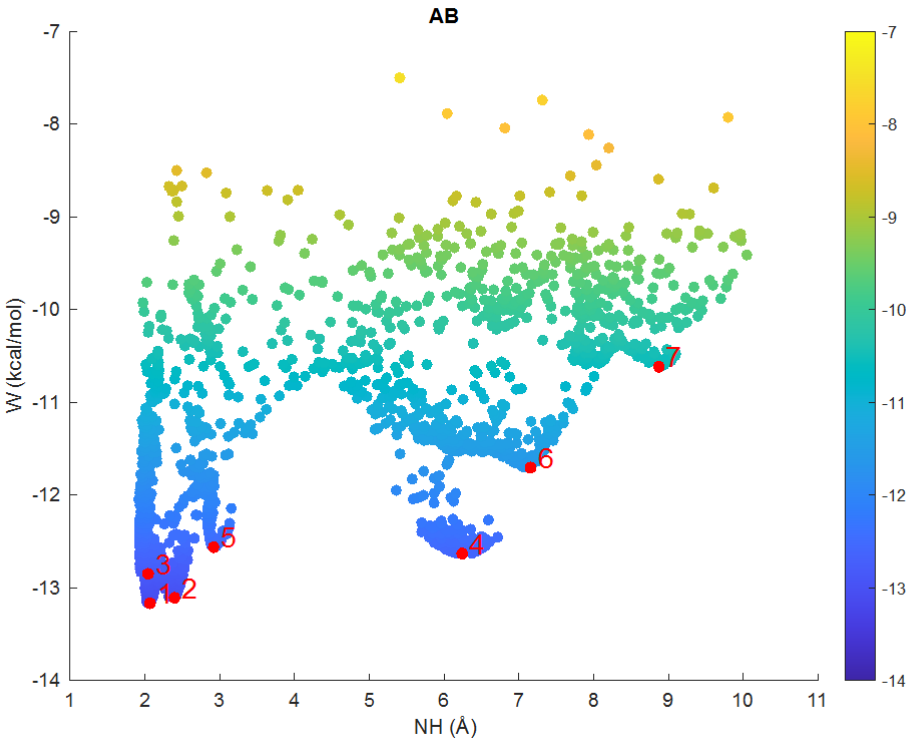

Figure 35 graph of the interaction energy (W_i) as a function of the N…H distance for the AB case of CBZ-Aza[5]helicene.

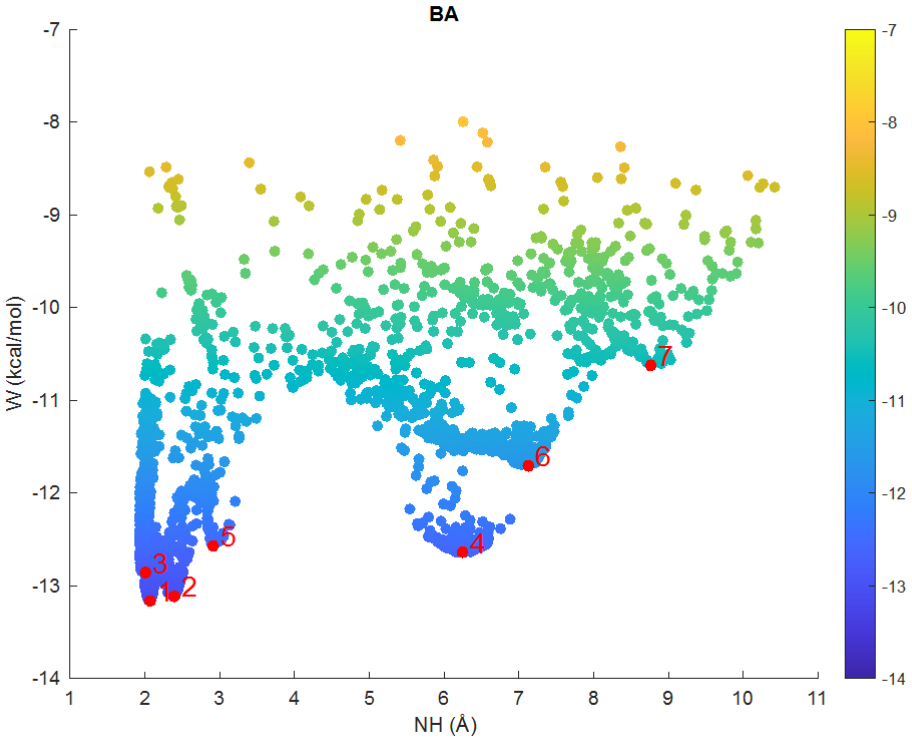

Figure 36 graph of the interaction energy (W_i) as a function of the N…H distance for the BA case of CBZ-Aza[5]helicene.

N…HN-angle (NHN) – Figure 37 and Figure 38. This parameter is not relevant for the geometries where the contribution of hydrogen bond is negligible because the N…H distances are too high, hence the dimers 4, 6 and 7 are not considered here. The following are the dimers of interest:

- Dimers 1 and 3 display values of the N…HN angle between 145° and 155°. They show good directionality and interatomic distance shorter than the sum of the Van de Walls radii. For those reason the hydrogen bond is relevant for both dimers.
- Dimer 2: the angle NHN ≈ 135.5° displays an intermediate directionality. The hydrogen bonding contribution to the interaction energy is still present, but less efficient.
- Dimer 5: the angle NHN ≈ 111.5° displays an even lower directionality than before. The hydrogen bond contribution is weaker than in Dimers 1, 3 and 2, but still contributes to the interaction energy.

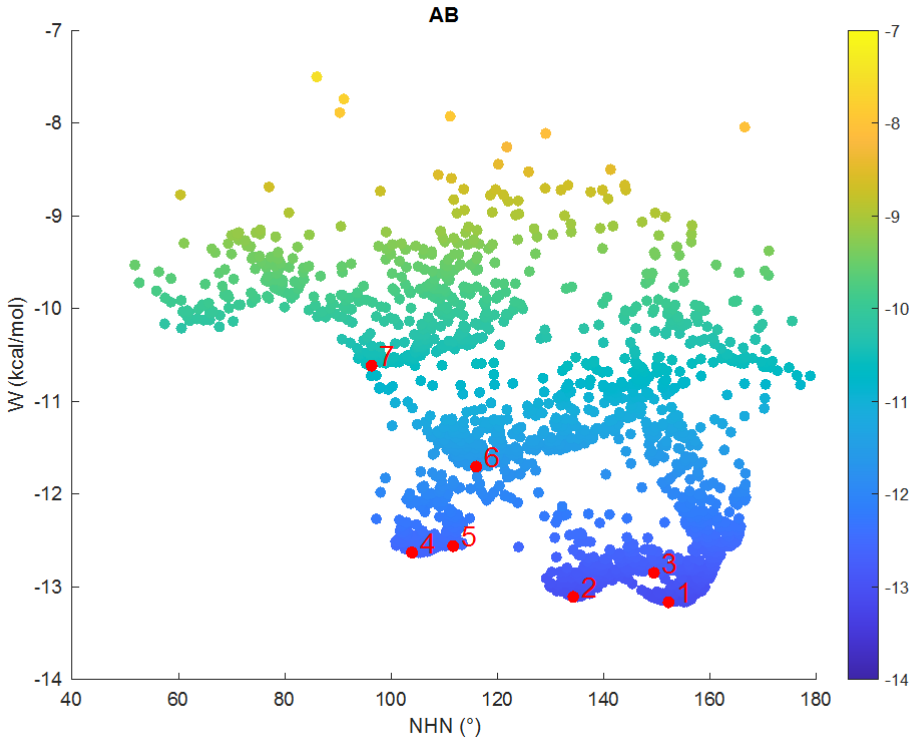

Figure 37: graph of the interaction energy (W_i) as a function of the N…HN angle for the AB case of CBZ-Aza[5]helicene.

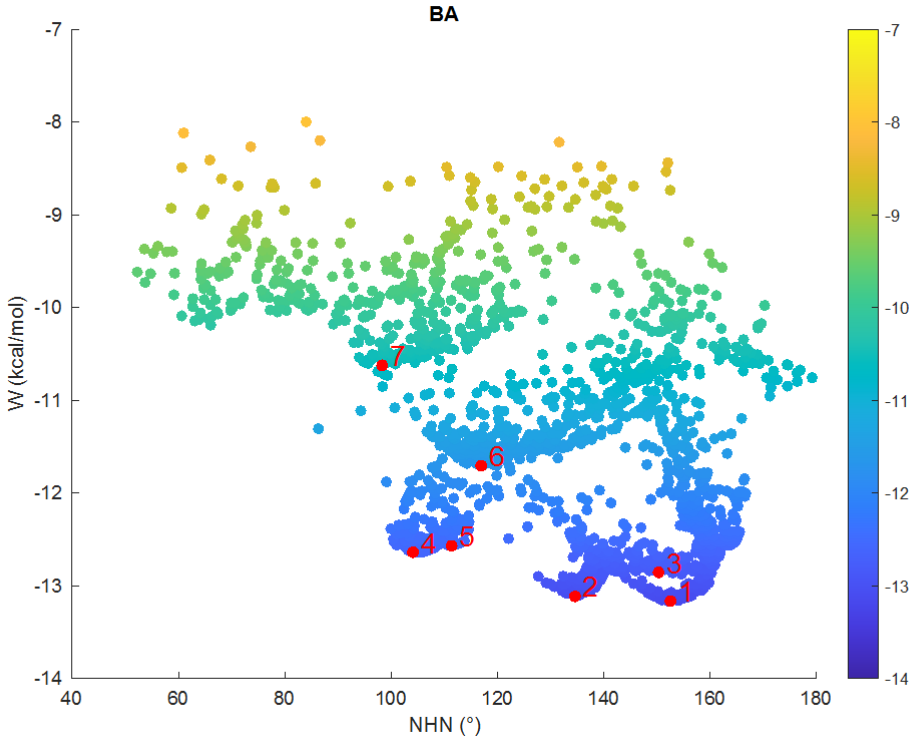

Figure 38: graph of the interaction energy (W_i) as a function of the N…HN angle for the BA case of CBZ-Aza[5]helicene.

Centroid distance (C) – Figure 39 and Figure 40. The dimer's conformations 7 is the one with the lower value of C, below 9 Å, and the least stable interaction energy. As it will be clearer by visual inspection of the three-dimensional representation of the dimers (see below), π-stacking is the only relevant interaction mechanism for dimer 7. This means that the NH_2 group of CBZ does not create any relevant contribution. So, the geometry of dimer 7 maximizes just the π-π interaction, which leads to an overall higher interaction energy (W_i). In all the other dimers (1, 2, 3, 4, 5 and 6), different kinds of interaction mechanisms compete with each other, leading to higher values of C, but stronger interaction and lower interaction energy (W_i).

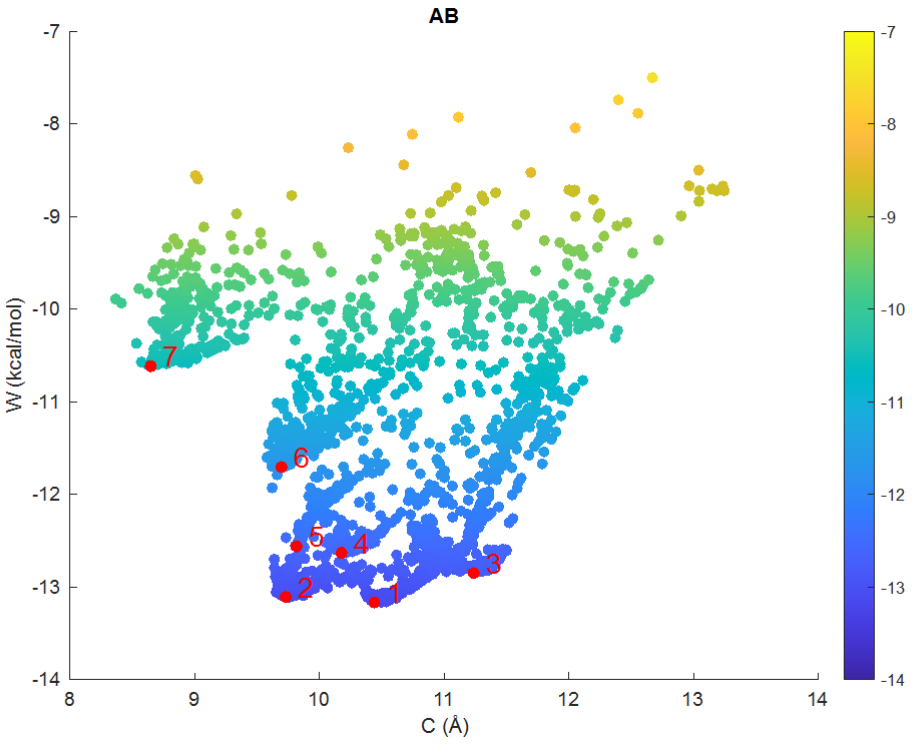

Figure 39: graph of the interaction energy (W_i) as a function of the Centroid distance (C) for the AB case of CBZ-Aza[5]helicene.

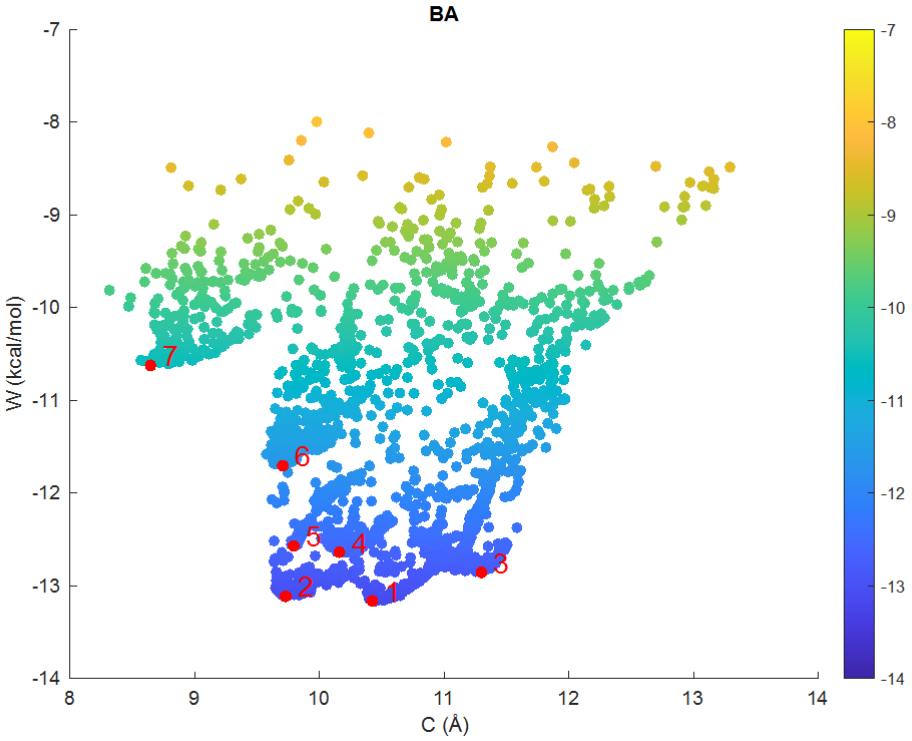

Figure 40: graph of the interaction energy (W_i) as a function of the Centroid distance (C) for the BA case of CBZ-Aza[5]helicene.

To conclude the discussion of the interaction energy plots, I report in Table 17 and Table 18, for each AB and BA dimers, the ID codes of the dimers, the interaction energy (W_i) and the values of the three geometrical parameters.

Table 17: geometrical parameters, interaction energy (W_i) and ID for each dimer configuaration in the case AB CBZ-Aza[5]helicene. The ID is reported for archive purposes.

AB[5]	*ID*	*W_{AB} (kcal/mol)*	*N…H (Å)*	*N…HN (°)*	*C (Å)*
1	61953	-13.17	2.066	152	10.44
2	132441	-13.11	2.396	134	9.73
3	98949	-12.85	2.042	149	11.24
4	75226	-12.63	6.239	104	10.18
5	176270	-12.56	2.919	112	9.82
6	154227	-11.70	7.154	116	9.69
7	177786	-10.61	8.867	96	8.65

Table 18: geometrical parameters, interaction energy (W_i) and ID for each dimer configuaration in the case BA CBZ-Aza[5]helicene. The ID is reported for archive purposes.

BA[5]	*ID*	*W_{BA} (kcal/mol)*	*N…H (Å)*	*N…HN (°)*	*C (Å)*
1	44200	-13.16	2.069	153	10.42
2	51292	-13.11	2.391	135	9.73
3	38226	-12.86	2.006	150	11.30
4	52083	-12.64	6.245	104	10.16
5	57346	-12.57	2.912	111	9.79
6	47107	-11.70	7.126	117	9.71
7	37966	-10.62	8.759	98	8.65

Here below, I discuss the three-dimensional representation of the seven dimers that have been identified above.

The first "family" of interacting geometries includes **dimers 1**, **2**, **3** and **5**. As one can see in Figure 41, Figure 42, Figure 43 and Figure 44 in such four cases, the NH_2 group is located nearby the N atom of Aza[5]helicene. Therefore, such geometries are strongly guided by the hydrogen bond interaction. Here, the π-stacking interaction can be considered a second order contribution.

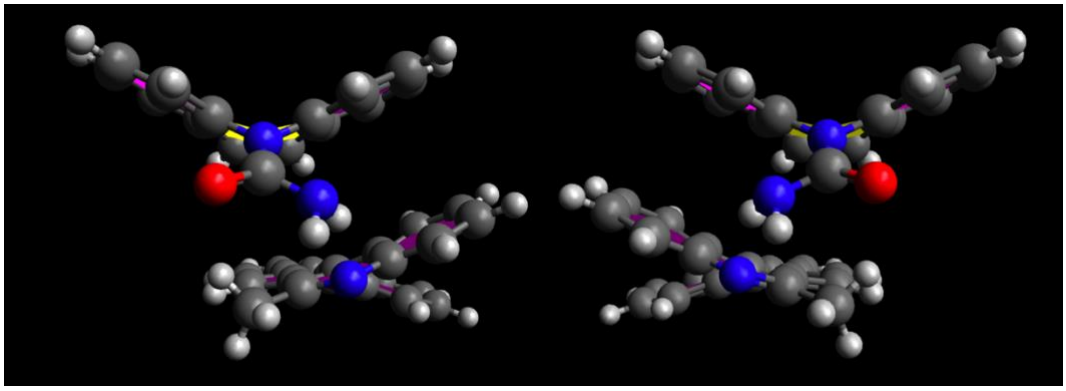

Figure 41: visual representation of the interacting Dimer 1. On the left AB and on the right BA. For the CBZ-Aza[5]helicene case.

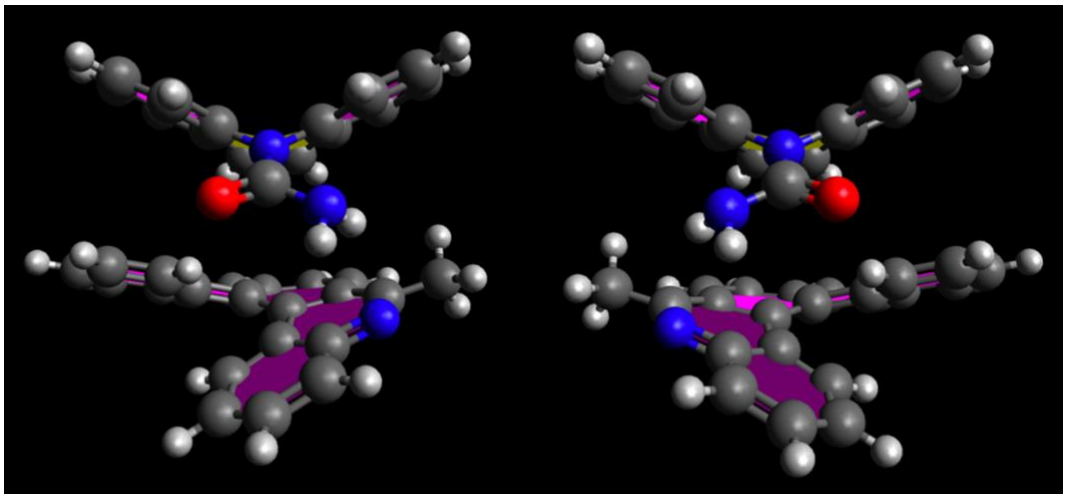

Figure 42: visual representation of the interacting Dimer 2. On the left AB and on the right BA. For the CBZ-Aza[5]helicene case.

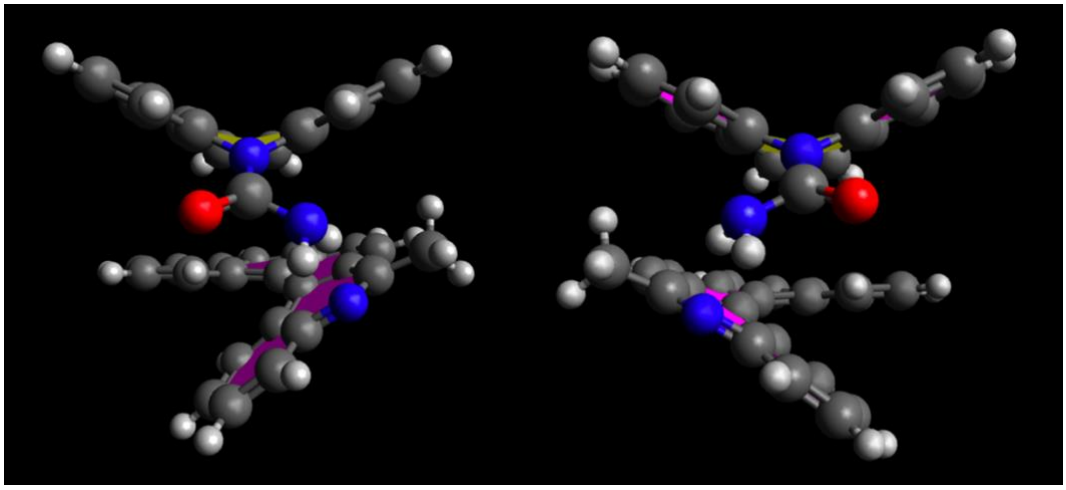

Figure 43: visual representation of the interacting Dimer 3. On the left AB and on the right BA. For the CBZ-Aza[5]helicene case.

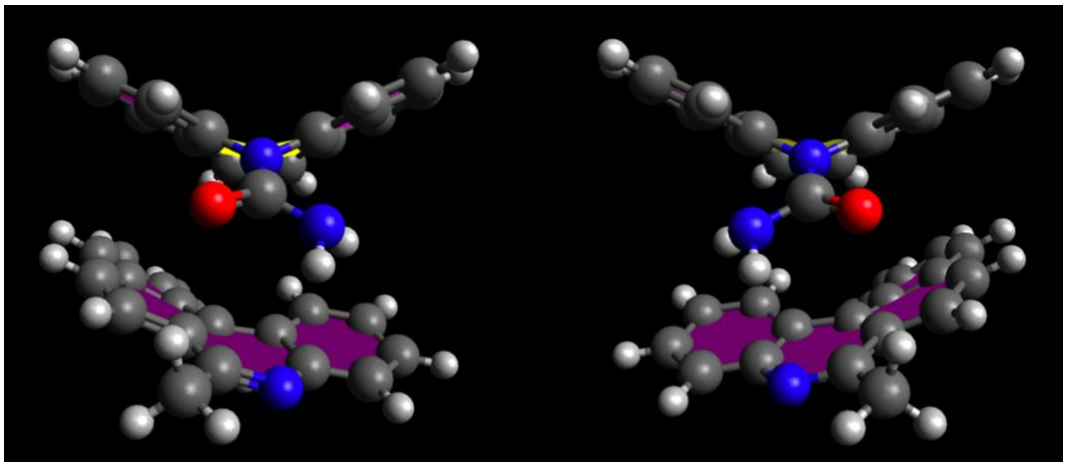

Figure 44: visual representation of the interacting Dimer 5. On the left AB and on the right BA. For the CBZ-Aza[5]helicene case.

A second “family” of interaction geometries contains **dimers 4** and **6** (Figure 45 and Figure 46). In these two dimers the NH_2 group is facing towards one of the aromatic rings of Aza[5]helicene, thus leading to a different kind of contribution to the interaction energy, that can be described by the NH_2-π interaction.

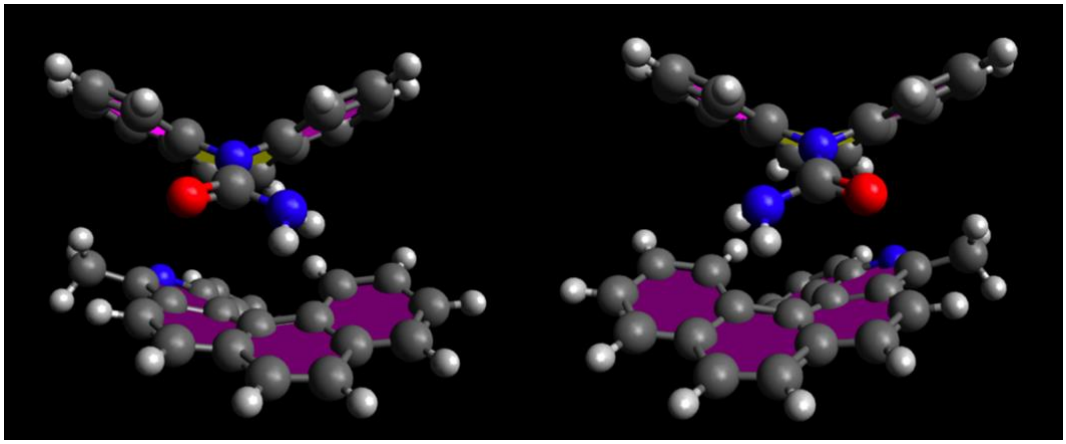

Figure 45: visual representation of the interacting Dimer 4. On the left AB and on the right BA. For the CBZ-Aza[5]helicene case.

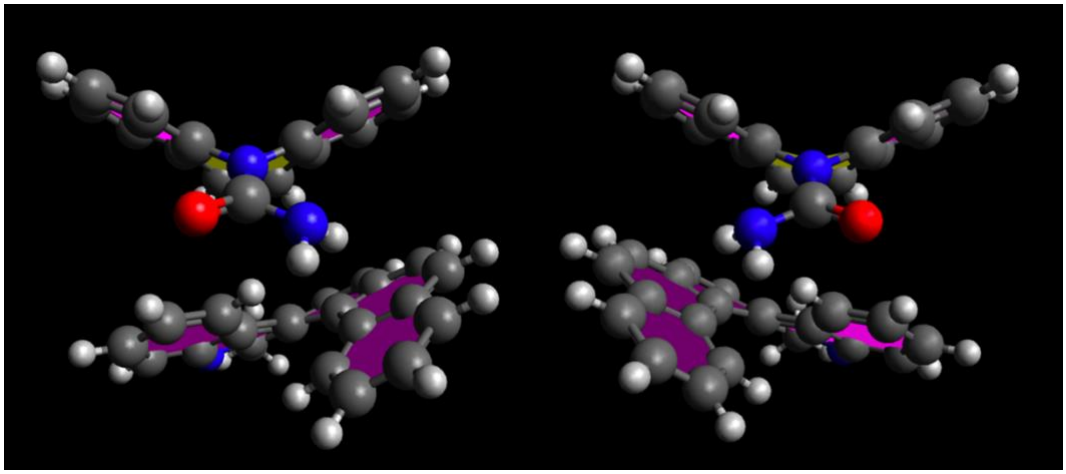

Figure 46: visual representation of the interacting Dimer 6. On the left AB and on the right BA. For the CBZ-Aza[5]helicene case.

Finally, the position of the NH_2 group in the interaction geometry of **dimer 7** (Figure 47) is in the opposite direction with respect to the nitrogen atom of Aza[5]helicene. In this kind of situation, hydrogen bonding is not possible, and the only relevant contribution is π-stacking.

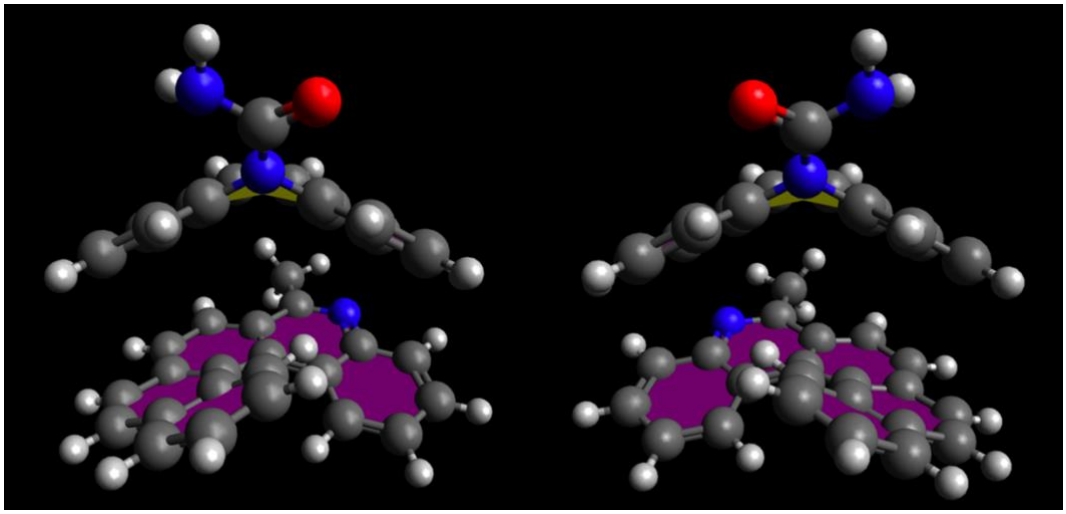

Figure 47: visual representation of the interacting Dimer 7. On the left AB and on the right BA. For the CBZ-Aza[5]helicene case.

5.3.3. NH_2-π

As discussed above, some dimers (AA/BB: dimer 4; AB/BA: dimes 4 and 6) interact in a peculiar way and display a sort of NH_2-π interaction. This type of interaction can be more generally defined as a polar-π interaction that involves molecules with permanent dipoles interacting with the quadruple moment of a π-system (Figure 48). So, in general any molecule with a hydrogen bond donor will electrostatically interact with the π-system of a conjugated molecule. In ours cases the NH_2 group of CBZ may interact with one aromatic ring of Aza[5]helicene. This is similar to the interaction between aniline and benzene that is worth 1.6 kcal/mol (Anslyn 2004). For this reason, I expect that this kind of interaction will be present in our systems as well.

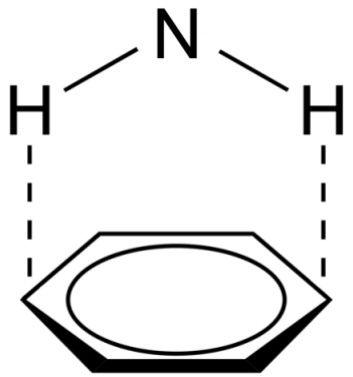

Figure 48: schematic representation of the polar-π interaction.

5.4. Boltzmann distribution of the interacting pairs

In the previous sections I have investigated the structures of the interacting dimers. Such structures were found in points of interest of the plots of the interaction energy, and it was possible to define three families of geometries of interaction based on the most relevant kind of interactions that were operating in each case.
Here I want to characterize better which one among such dimers are statistically relevant at a given temperature. To this aim, I apply the Boltzmann distribution to find the conformations that are more stable at ambient temperature ($T = 298K$). In general, the Boltzmann distribution is given by the expression:

$$f_i = \frac{e^{-E_i/k_bT}}{Z(T)}$$

$Z(T)$ is the partition function, which is given by $Z(T) = \sum_i e^{-E_i/k_bT}$

In our case the energy E_i is the total energy of the interacting dimer, and it is equal to the sum of the interaction energy (W_i) and the energy of the two non-interacting molecules, at the conformation of the dimer ($E_0 = E_1 + E_2$):

$$E_i = E_0 + W_i = E_1 + E_2 + W_i$$

As we discussed before (Chapter 5), we can approximate E_1 and E_2 to two constant values (respectively E_1^0 and E_2^0) that are the energies of the two isolated molecules in their equilibrium (this is acceptable when the molecules are relatively rigid – as the present case). For this reason, E_0 can be considered as a constant:

$$E_0 = E_1 + E_2 \approx E_1^0 + E_2^0$$

By using the previous approximation, the Boltzmann distribution can be rewritten as:

$$f_i = \frac{e^{-E_0/k_bT} e^{-W_i/k_bT}}{\sum_i e^{-E_0/k_bT} e^{-W_i/k_bT}} = \frac{e^{-E_0/k_bT} e^{-W_i/k_bT}}{e^{-E_0/k_bT} \sum_i e^{-W_i/k_bT}}$$

That is:

$$f_i = \frac{e^{-W_i/k_bT}}{\sum_i e^{-W_i/k_bT}}$$

Therefore, one can apply Boltzmann statistics directly to the interaction energies (instead of the total energies). This expedient allows to relate more straightforwardly the outcomes of Boltzmann analysis to the previous discussion of the interaction energies and identify the most relevant configurations of those dimers.

The results of the Boltzmann analysis for the case of CBZ-Aza[5]helicene are reported in the tables below (Table 19, Table 20, Table 21 and Table 22).

Table 19: Boltzman distribution (f_i), interaction energy (W_i) and ID for each dimer configuaration in the case AA CBZ-Aza[5]helicene. The ID is reported for archive purposes.

AA[5]	*ID*	W_{AA} *(kcal/mol)*	ΔW_{AA} *(kcal/mol)*	Z_i	f_i
1	141695	-13.66	0.00	1.00	73.7%
2	170195	-12.85	0.81	0.26	18.8%
3	126936	-12.03	1.63	0.06	4.7%
4	101874	-11.50	2.16	0.03	1.9%
5	59473	-10.49	3.17	0.00	0.3%
6	101343	-10.30	3.36	0.00	0.3%
7	159806	-10.08	3.58	0.00	0.2%

Table 20: Boltzman distribution (f_i), interaction energy (W_i) and ID for each dimer configuaration in the case BB CBZ-Aza[5]helicene. The ID is reported for archive purposes.

BB[5]	*ID*	W_{BB} *(kcal/mol)*	ΔW_{BB} *(kcal/mol)*	Z_i	f_i
1	178795	-13.66	0.00	1.00	73.8%
2	75171	-12.84	0.82	0.25	18.5%
3	98725	-12.07	1.59	0.07	5.0%
4	65925	-11.49	2.17	0.03	1.9%
5	62795	-10.48	3.18	0.00	0.3%
6	84513	-10.30	3.36	0.00	0.3%
7	102475	-10.08	3.58	0.00	0.2%

Table 21: Boltzman distribution (f_i), interaction energy (W_i) and ID for each dimer configuaration in the case AB CBZ-Aza[5]helicene. The ID is reported for archive purposes.

AB[5]	*ID*	W_{AB} *(kcal/mol)*	ΔW_{AB} *(kcal/mol)*	Z_i	f_i
1	61953	-13.17	0.00	1.00	29.9%
2	132441	-13.11	0.06	0.90	27.0%
3	98949	-12.85	0.32	0.58	17.4%
4	75226	-12.63	0.54	0.40	12.1%
5	176270	-12.56	0.61	0.36	10.7%
6	154227	-11.70	1.46	0.08	2.5%
7	177786	-10.61	2.55	0.01	0.4%

Table 22: Boltzman distribution (f_i), interaction energy (W_i) and ID for each dimer configuaration in the case BA CBZ-Aza[5]helicene. The ID is reported for archive purposes.

BA[5]	*ID*	W_{BA} *(kcal/mol)*	ΔW_{BA} *(kcal/mol)*	Z_i	f_i
1	44200	-13.16	0.00	1.00	29.4%
2	51292	-13.11	0.05	0.92	27.1%
3	38226	-12.86	0.31	0.60	17.6%
4	52083	-12.64	0.52	0.41	12.2%
5	57346	-12.57	0.59	0.37	10.8%
6	47107	-11.70	1.46	0.09	2.5%
7	37966	-10.62	2.54	0.01	0.4%

5.5. Aza[5]helicene: comparing the AA/BB vs. the AB/BA pairs

In the previous sections I have examined the four possible interactions cases, and I have grouped them in the two classes AA/BB and AB/BA, based on enantiomeric equivalence (*i.e.*, AA (AB) is the mirror image of BB (BA)). For this reason, from now on, I will consider only the AA and AB cases, as the mirror cases BB and BA are just equivalent.

It was also shown above that in the two interaction families one can achieve different kinds of geometrical arrangement of the dimers. This fact is at the basis of the possibility to provide chiral recognition of AEDs using a SERS substrate functionalized with a specific enantiomeric form of Azahelicenes.

For instance, by considering the isomeric A form of Aza[5]helicene, it will interact differently with the two geometric enantiomers of CBZ (A & B) Figure 49, potentially leading to different SERS signals in the two cases (AA *vs*. AB). As a perspective, if the differences between the two SERS spectra are appreciable, one may hope to establish, by spectroscopy, which enantiomer of a chiral drug is interacting with the chiral functionalized SERS substrate – and this would lead to chiral recognition behavior.

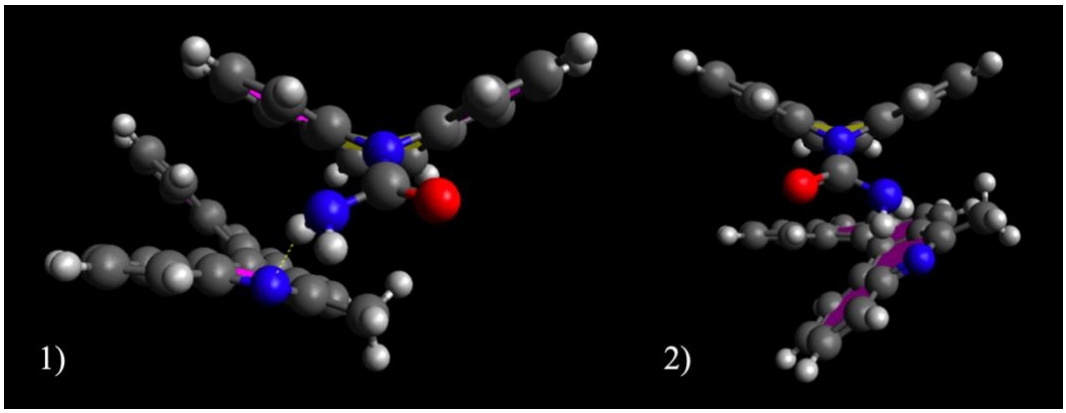

Figure 49: visual representation and comparison of the interaction dimers for the Aza[5]helicene. 1) case AA. 2) case AB.

Therefore, I start by considering here below the distribution of interaction energies for the two cases (AA, AB), so as to establish if the chiral recognition is actually possible by using Aza[5]helicene. As one can see from the histograms reported in Figure 50 and Figure 51 in AA case there is a significant shift of the interaction energies (W_i) towards more negative values. This is already a good indication of chiral-sensitive intermolecular interactions. However, it is important to underline that those histograms do not represent a probability. They simply show the energy spread of the total number of conformations probed by OBSolv in a series of 2000 repeated runs. For this reason, the different energy distributions are not sufficient to fully confront the AA *vs*. the AB case.

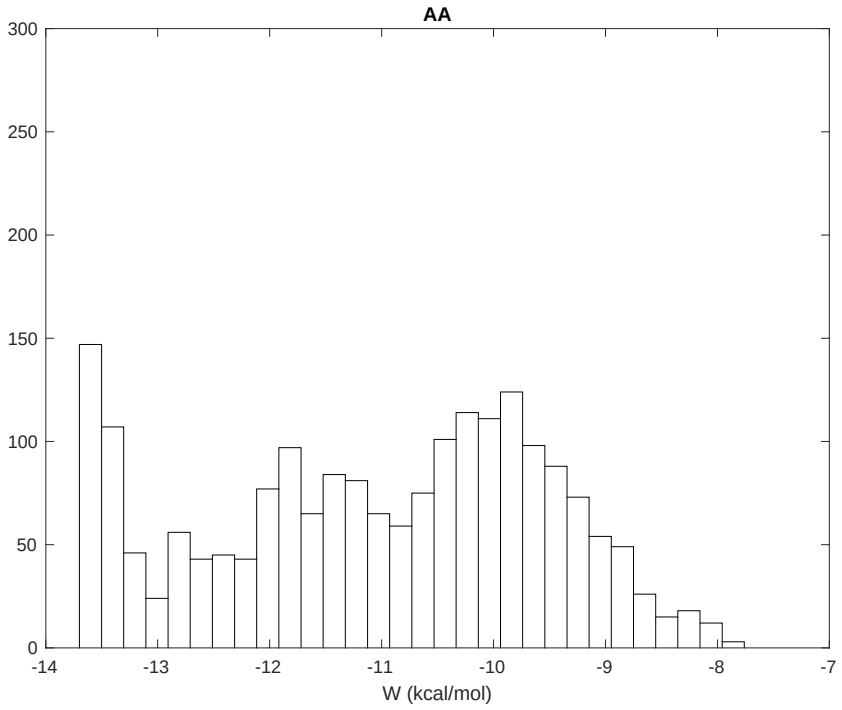

Figure 50: histogram of the OBSolv simulated dimers as a function of he interaction energy (W) for the case AA CBZ-Aza[5]Helicene. The histogram was obtained on a data set of 2000 independent simulations.

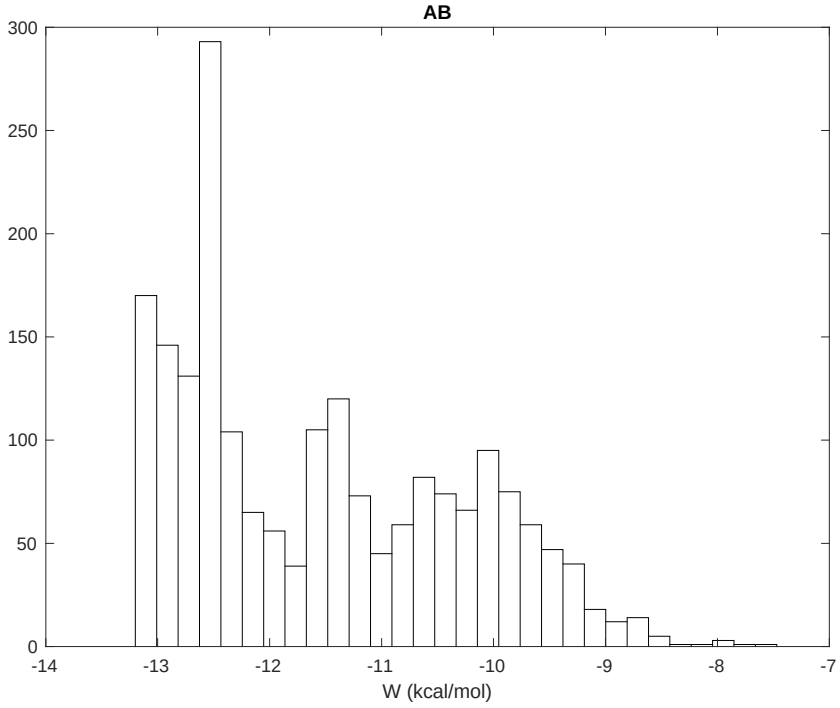

Figure 51: histogram of the OBSolv simulated dimers as a function of he interaction energy (W) for the case AB CBZ-Aza[5]Helicene. The histogram was obtained on a data set of 2000 independent simulations.

In other terms, OBSolv does not consider thermal fluctuations and the thermodynamic behavior of the simulated molecular system. To actually assess the probability of observing a chiral-dependent different interaction, I recur to the Boltzmann distribution. By taking the data previously calculated in the section 5.4, I can plot in Figure 52 the probability (f_i) vs. the interaction energy (W_i). As expected, the plot has an exponential behavior. The AA and BB cases are almost overlapping in Figure 52.

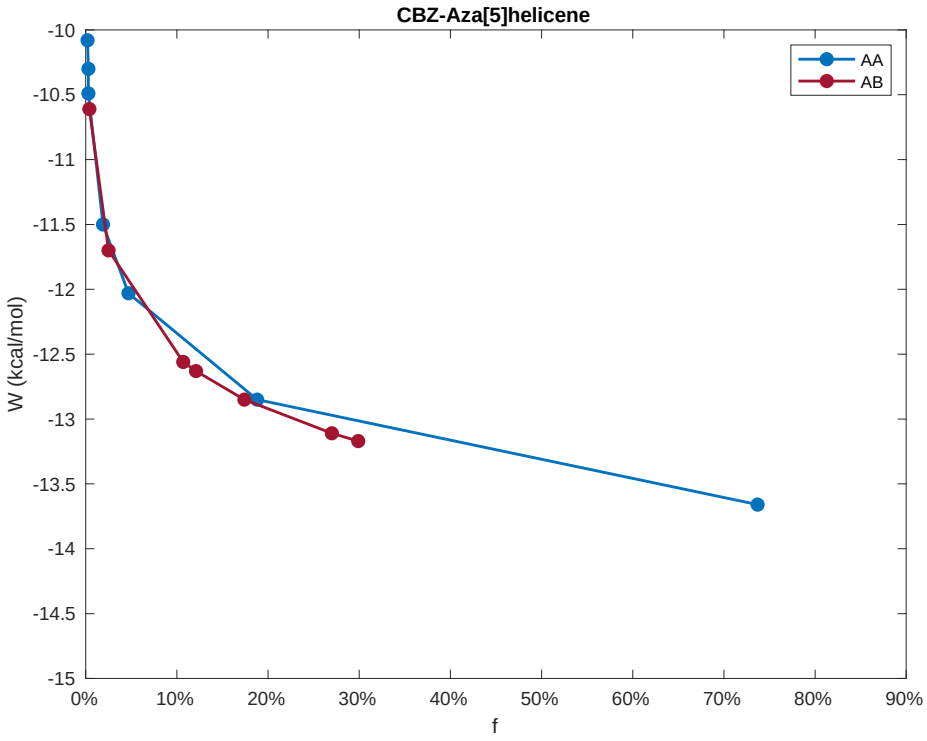

Figure 52: graph of the Boltzmann distribution (f) as a function of the interaction energy (W) for AA and AB in the case of CBZ-Aza[5]helicene.

The energy difference between the two most stable conformers of the two cases is just ~ 0.5 kcal/mol. At room temperature, this energy difference could be not enough to obtain an efficient chiral recognition. Moreover, for the AB case, the probability for the first and second more stable conformer is comparable. This would lead to a competition between the two at ambient temperature and may originate a more complex SERS spectrum that would be a convolution of the signal of the two geometries, thus hampering chiral recognition by spectroscopic analysis. Hence, for Aza[5]helicene, chiral recognition is hard to prove, at least based on just these data. The simulation of Raman spectra may provide useful additional information and will be taken into consideration later on.

5.6. The Case of Aza[6]helicene

In the following sections I focus on the case of Aza[6]helicene interacting with CBZ. As I have previously established for Aza[5]helicene, AA interactions are equivalent to BB interactions (similarly, AB interactions are equivalent to BA interactions). Hence, for the sake of conciseness, I just focus on the AA and AB cases.

5.6.1. CBZ-Aza[6]helicene: the AA pairs

I report in Figure 53 the 3D plot of the interaction energy of the AA pairs as a function of the usual selected geometric parameters (N…H distance, N…HN angle, Centroid distance C). The plot evidences the presence of four clustered regions corresponding to local interaction energy minima (that correspond to dimers of interest). I will examine and comment such minima by using a series of 2D plots, where the interaction energy is reported as a function of one geometric variable at the time.

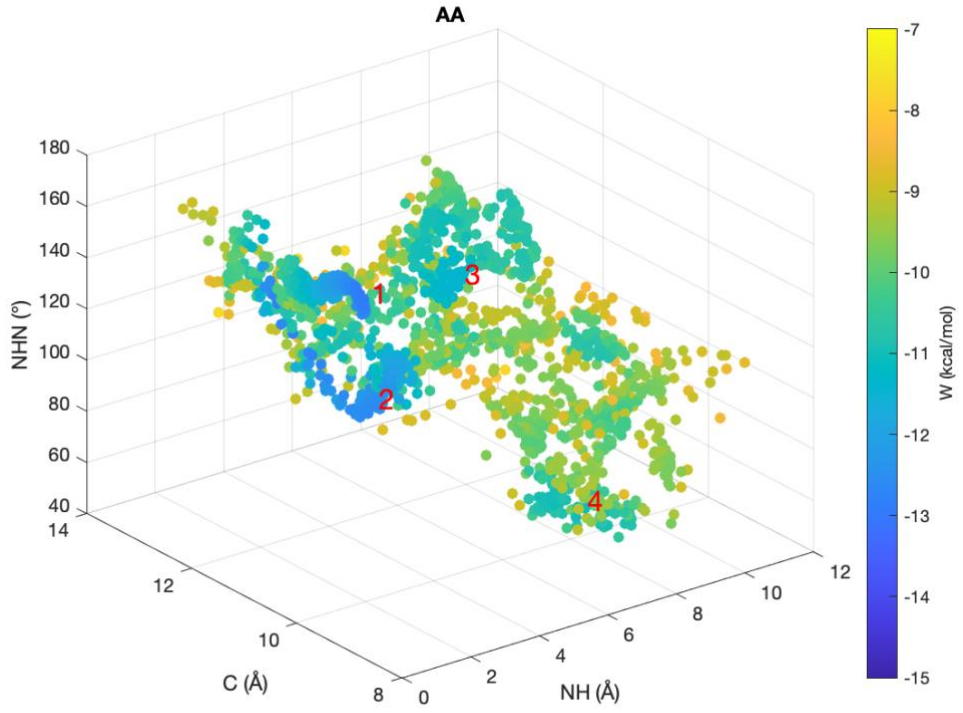

Figure 53: 3D graph of geometrical parameters and interaction energy (W_i) for the AA case of CBZ-Aza[6]helicene.

By starting from the **N…H-distance (NH)** – Figure 54 – I can state that for dimers 1 and 2 the hydrogen bond contribution to the interaction energy is relevant, as the associated N…H distance is relatively short. However, for dimers 3 and 4 the hydrogen bonding is negligible because the NH_2 group of CBZ is too far away from the N atom of the Aza[6]helicene (distance > 4 Å).

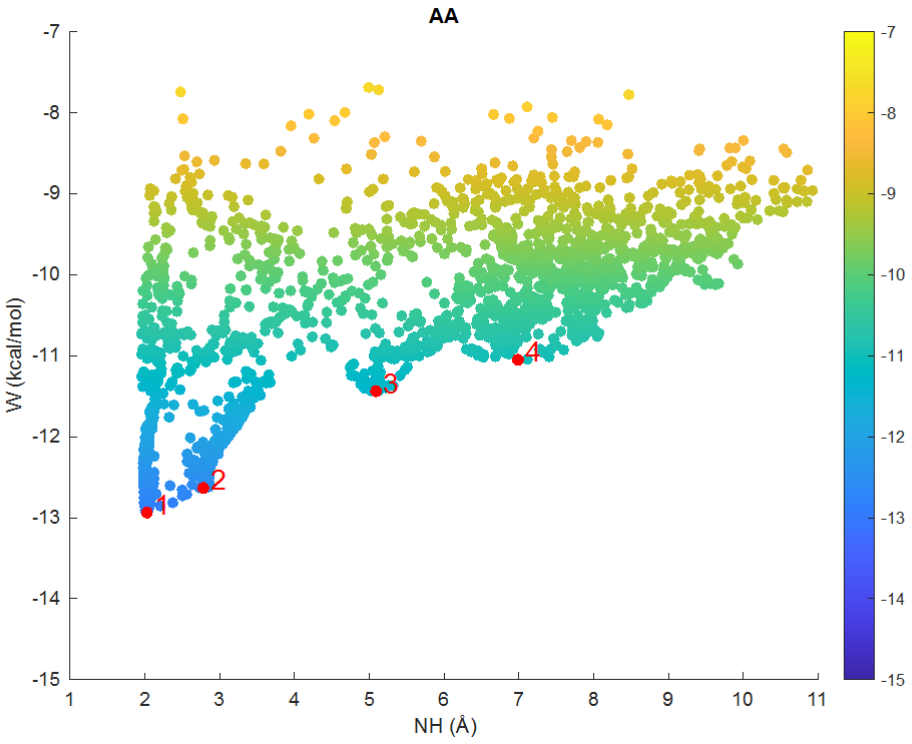

Figure 54: graph of the interaction energy (W_i) as a function of the N…H distance for the AA case of CBZ-Aza[6]helicene.

The **N…HN-angle (NHN)** – Figure 55 – is relevant just for those dimers where the interaction mechanism is strongly related to the hydrogen bond contribution. Hence, I consider only dimer 1 and 2. For dimer 1 it is NHN = 157°, which highlights a quite directional interaction, with considerable hydrogen bond contribution. For dimer 2: it is NHN = 109°, and the interaction has lower directionality. The hydrogen bond contribution is less effective than in dimer 1, but still active.

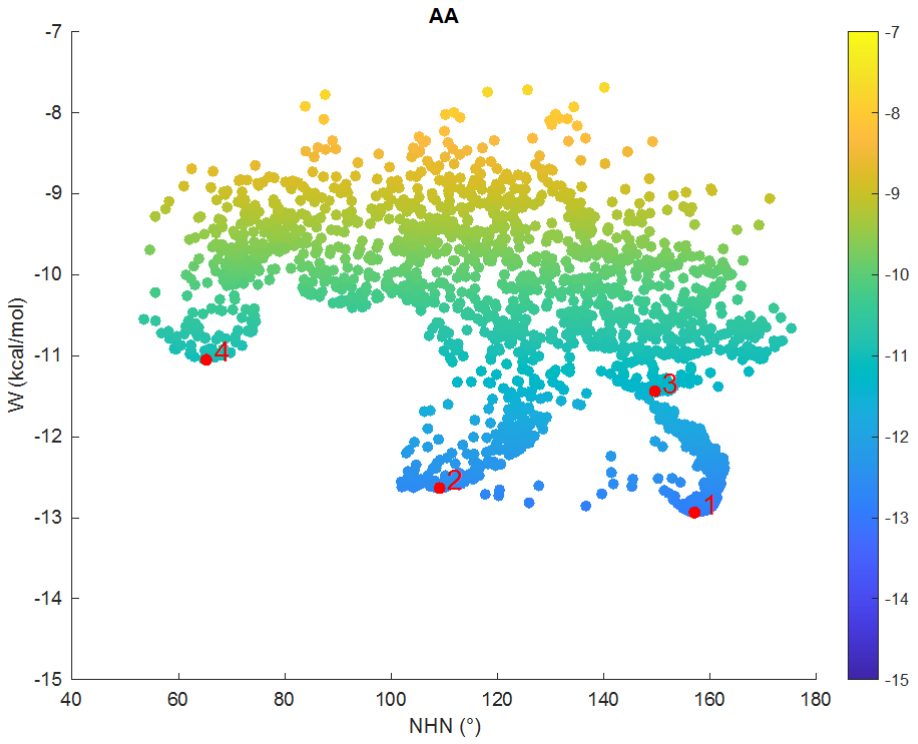

Figure 55: graph of the interaction energy (W_i) as a function of the N…HN angle for the AA case of CBZ-Aza[6]helicene.

The **Centroid distance (C)** plot of Figure 56 shows that the conformation of dimer 4 is the one with the lowest C value, and the highest interaction energy (W_i). This can be justified by a dimer geometry that maximizes π-stacking at the expense of other more effective interactions, as hydrogen bonding. Notably, dimers 1, 2 and 3 achieve a lower interaction energy by exploiting other interaction mechanisms that compete with each other leading to higher values of C but lower interaction energy (W_i).

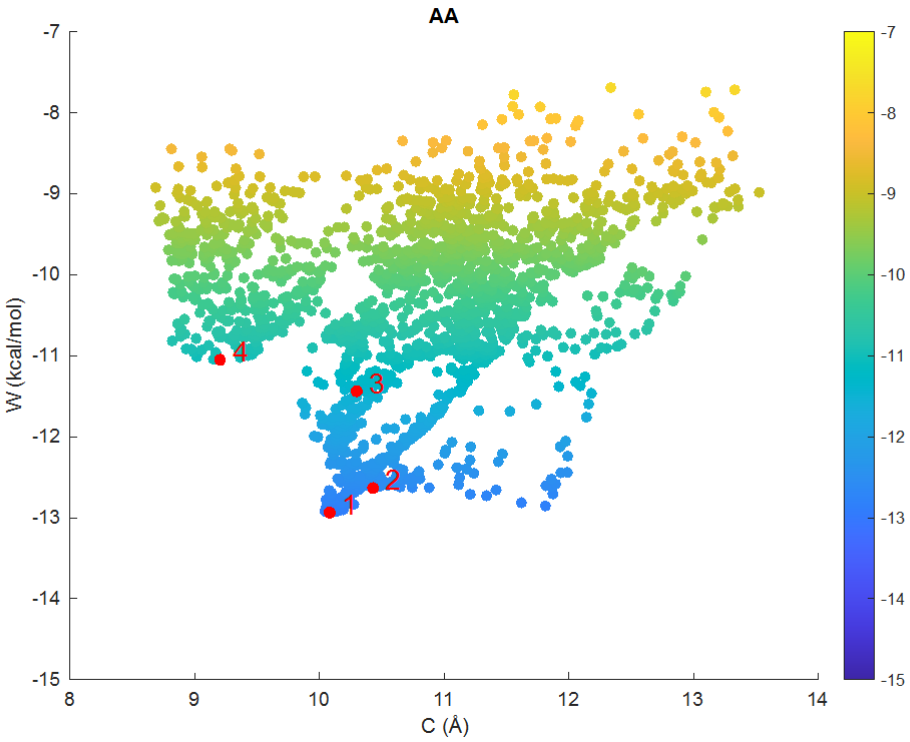

Figure 56: graph of the interaction energy (W_i) as a function of the Centroid distance (C) for the AA case of CBZ-Aza[6]helicene.

To conclude this section, I report in Table 23, for each dimer, the ID code, the interaction energy (W_i) and the three geometrical parameters for the case AA. The Boltzmann distribution is also reported in the following Table 24.

Table 23: geometrical parameters, interaction energy (W_i) and ID for each dimer configuaration in the case AA CBZ-Aza[6]helicene. The ID is reported for archive purposes.

AA[6]	*ID*	W_{AA} *(kcal/mol)*	*N…H (Å)*	*N…HN (°)*	*C (Å)*
1	118867	-12.94	2.026	157	10.08
2	113446	-12.63	2.781	109	10.43
3	141607	-11.44	5.088	150	10.30
4	143064	-11.05	6.985	65	9.20

Table 24: Boltzman distribution (f_i), interaction energy (W_i) and ID for each dimer configuaration in the case AA CBZ-Aza[6]helicene. The ID is reported for archive purposes.

AA[6]	*ID*	W_{AA} *(kcal/mol)*	ΔW_{AA} *(kcal/mol)*	Z_i	f_i
1	118867	-12.94	0.00	1.00	58.1%
2	113446	-12.63	0.30	0.60	34.8%
3	141607	-11.44	1.50	0.08	4.6%
4	143064	-11.05	1.89	0.04	2.4%

Here below, I report the three-dimensional representation of the dimers of interest that have been selected during the previous analysis. As one can see in Figure 57 and Figure 58, for **dimers 1** and **2**, the NH_2 group is located nearby the N atom of Aza[6]helicene, hence the two structures are strongly guided by the formation of the hydrogen bond, and π-stacking is a second order contribution.

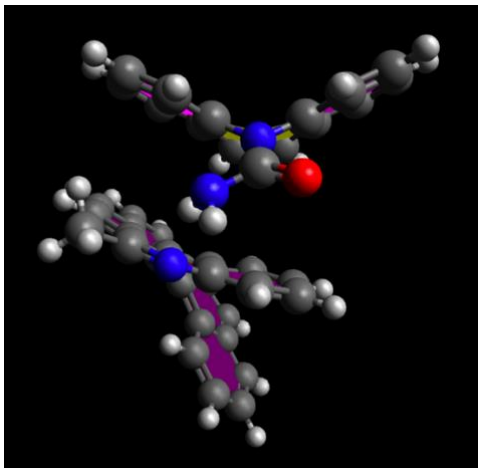

Figure 57: visual representation of the interacting Dimer 1. For the AA CBZ-Aza[6]helicene case.

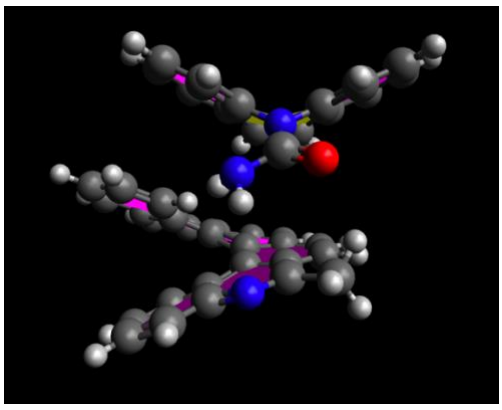

Figure 58: visual representation of the interacting Dimer 2. For the AA CBZ-Aza[6]helicene case.

For **dimer 3** (Figure 59), the NH_2 group is facing one of the aromatic rings of Aza[6]helicene, so the NH_2-π contribution is relevant instead of the hydrogen bond. Similar to dimers 1 and 2, also in this case π-stacking is a second order contribution.

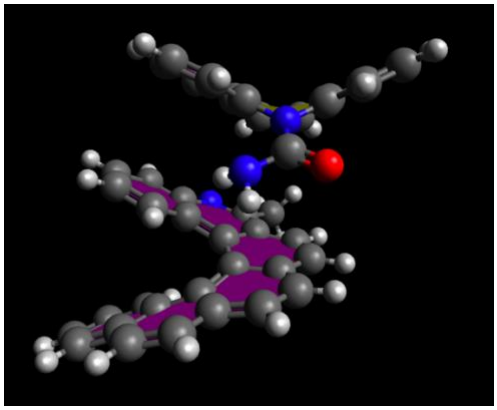

Figure 59: visual representation of the interacting Dimer 3. For the AA CBZ-Aza[6]helicene case.

Finally, in **dimer 4** (Figure 60) CBZ is facing in the opposite direction than the previous cases, hence the NH_2 group is too far away from Aza[6]helicene to be able to contribute to the interaction mechanism. In this case the only relevant contribution to molecular interaction is π-stacking.

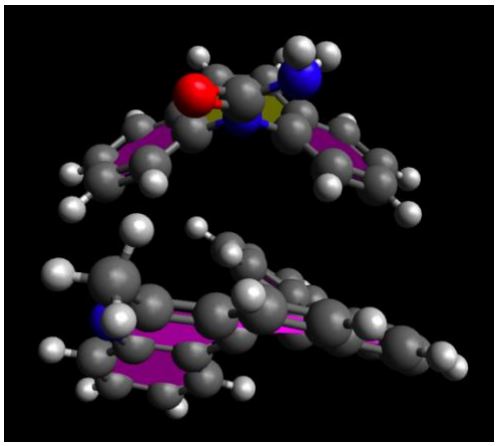

Figure 60: visual representation of the interacting Dimer 4. For the AA CBZ-Aza[6]helicene case.

5.6.2. CBZ-Aza[6]helicene: the AB pairs

I report in Figure 61 the 3D graph of the interaction energy of the AB pairs as a function of the three selected geometric parameters. I can highlight in the plot the presence of six clusters of dimers that are better shown and discussed in the following 2D graphs.

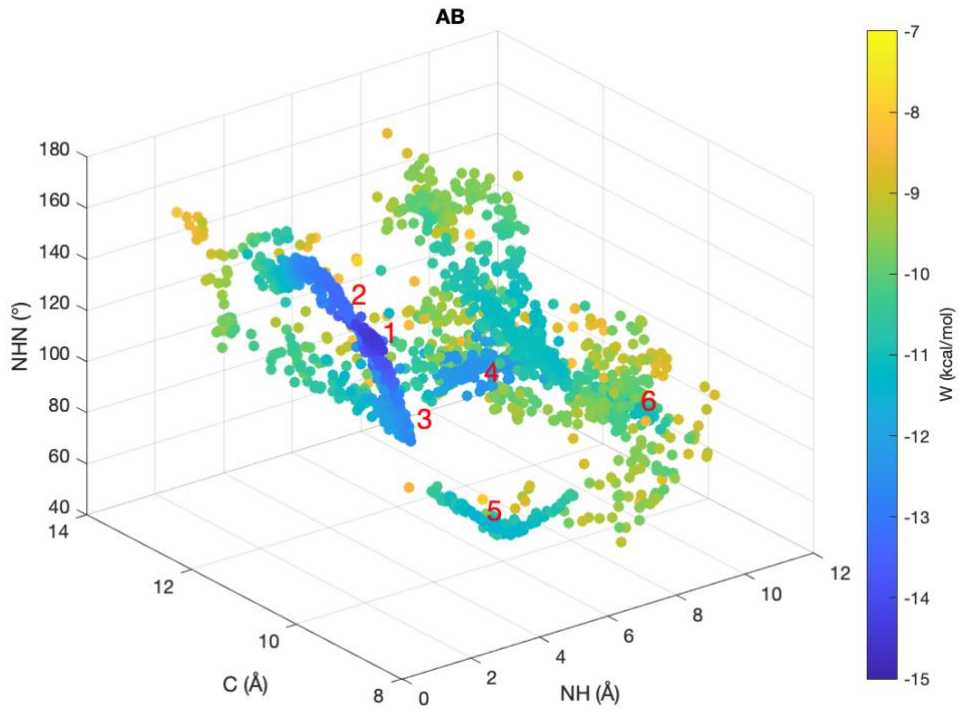

Figure 61: 3D graph of geometrical parameters and interaction energy (W_i) for the AB case of CBZ-Aza[6]helicene.

The **N…H-distance (NH)** – Figure 62 – in dimers 1, 2 and 3 is short enough, so that the hydrogen bonding contribution to the interaction energy is relevant in the three cases. However, in dimers 4, 5 and 6 the N…H distance is large (> 4 Å), and other kind of interactions contribute to the interaction energy.

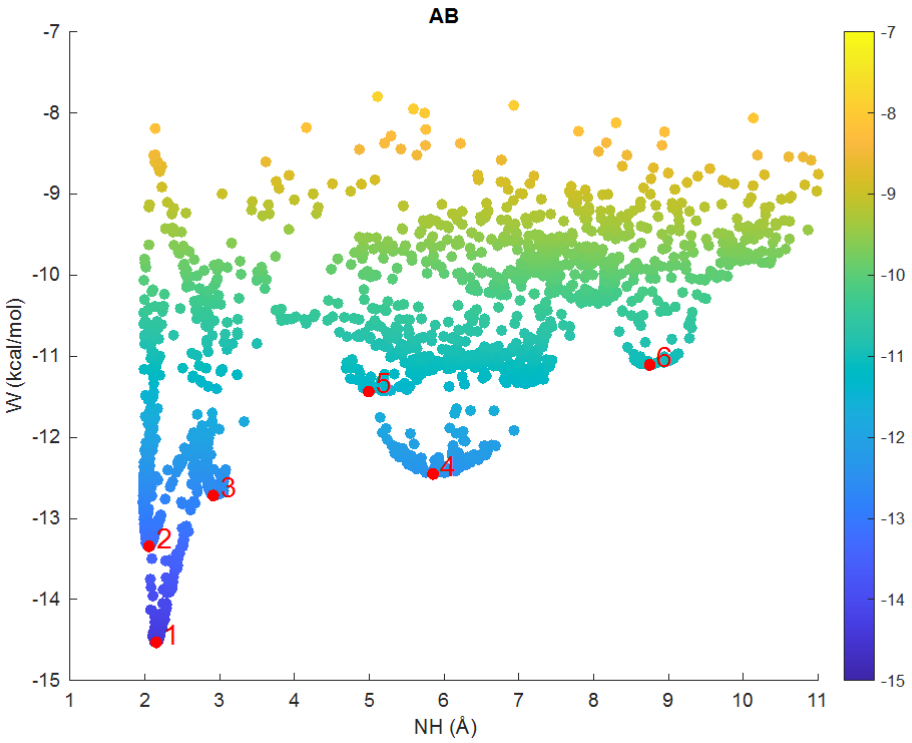

Figure 62: graph of the interaction energy (W_i) as a function of the N…H distance for the AB case of CBZ-Aza[6]helicene.

The **N…HN-angle (NHN)** – Figure 63 – is relevant only for those geometries where the interaction mechanism is strongly related to the formation of an hydrogen bond. Hence, I consider only dimers 1, 2 and 3. Dimer 1 is peculiar, it is characterized by a very low interaction energy and a low N…H distance, however the N…HN angle is smaller (NHN = 143°) than those found previously (section 5.6.1). Hence, in dimer 1 the hydrogen bond interaction has an intermediate directionality. The hydrogen bond contribution is relevant here, but this point deserves further analysis. Dimer 2 is characterized by a strongly directional interaction, and the hydrogen bond contribution is relevant. In dimer 3 the interaction has low directionality: the hydrogen bond contribution is less relevant but still present.

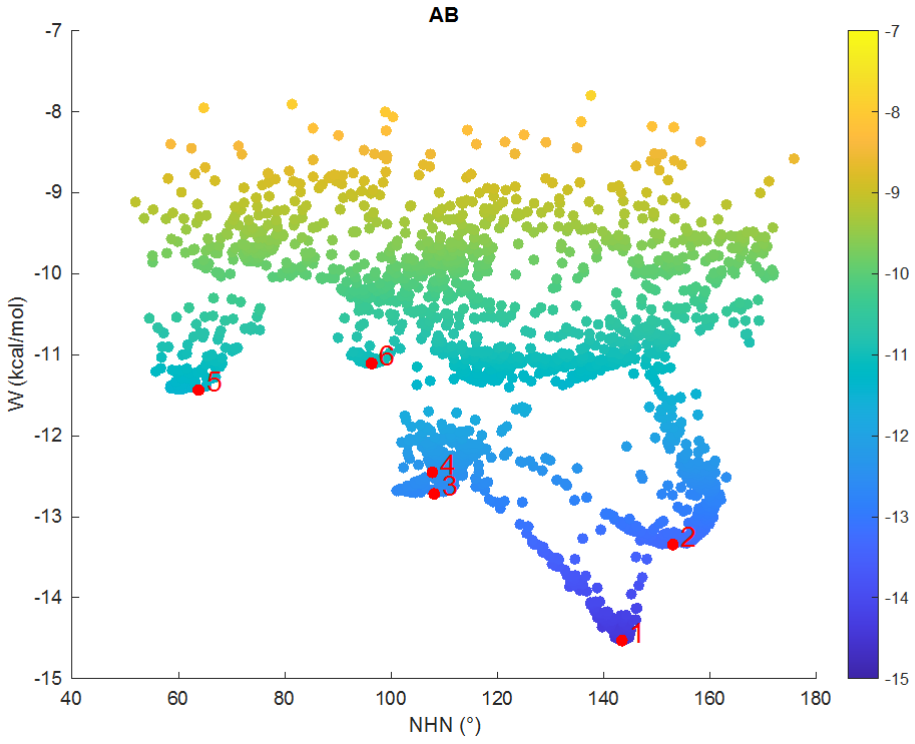

Figure 63: graph of the interaction energy (W_i) as a function of the N…HN angle for the AB case of CBZ-Aza[6]helicene.

The plot of the **Centroid distance (C)** – Figure 64 – shows that the dimers 1, 3 and 5 are located at similar C values but have very different W_i values. Such points are located between C = 9.8 Å and C = 9.9 Å and are not present in the AA case, where we can see a gap in that region of C values. For this reason, we can anticipate a strong chiral recognition for the Aza[6]helicene case. Dimer 6 is the one with the lowest C values and, as it will be confirmed later on, the interaction mechanism is only due to π-stacking.

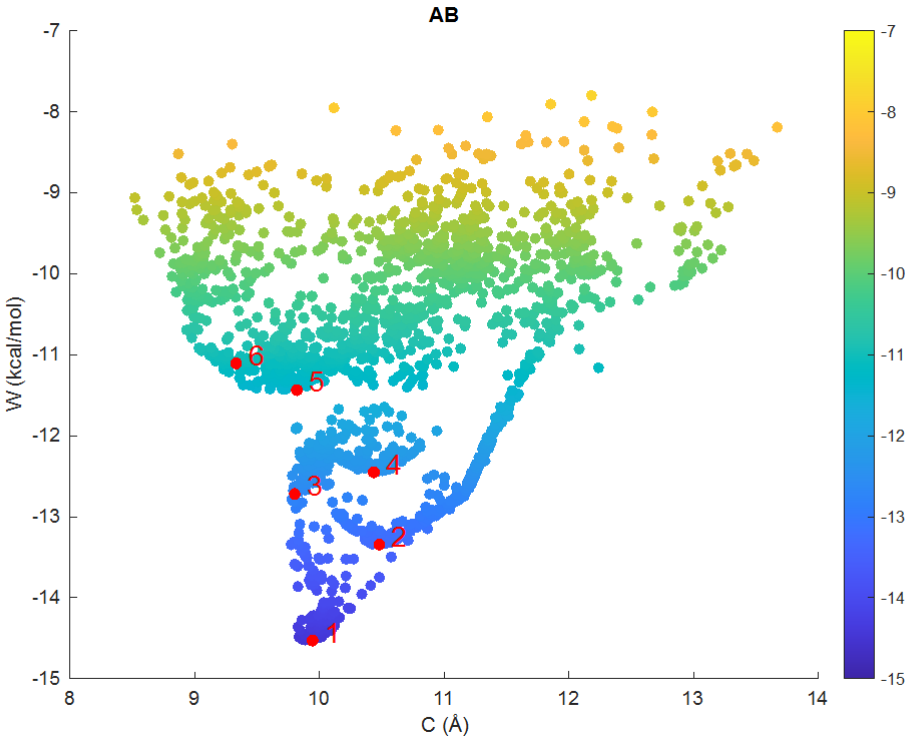

Figure 64: graph of the interaction energy (W_i) as a function of the Centroid distance (C) for the AB case of CBZ-Aza[6]helicene.

To conclude this section about the AB interaction case, I report in Table 25, for each dimer, the ID code the interaction energy (W_i) and the three geometrical parameters. The Table 26 that follows reports the corresponding Boltzmann distribution.

Table 25: geometrical parameters, interaction energy (W_i) and ID for each dimer configuaration in the case AB CBZ-Aza[6]helicene. The ID is reported for archive purposes.

AB[6]	*ID*	*W_{AB} (kcal/mol)*	*N…H (Å)*	*N…HN (°)*	*C (Å)*
1	13591	-14.53	2.151	143	9.94
2	27398	-13.34	2.053	153	10.48
3	44792	-12.72	2.913	108	9.80
4	24147	-12.45	5.849	108	10.44
5	24673	-11.44	4.989	64	9.82
6	38004	-11.11	8.746	96	9.33

Table 26: Boltzman distribution (f_i), interaction energy (W_i) and ID for each dimer configuaration in the case AB CBZ-Aza[6]helicene. The ID is reported for archive purposes.

AB[6]	*ID*	*W_{AB} (kcal/mol)*	*ΔW_{AB} (kcal/mol)*	*Z_i*	*f_i*
1	13591	-14.53	0.00	1.00	81.9%
2	27398	-13.34	1.18	0.14	11.1%
3	44792	-12.72	1.81	0.05	3.8%
4	24147	-12.45	2.08	0.03	2.5%
5	24673	-11.44	3.09	0.01	0.4%
6	38004	-11.11	3.42	0.00	0.3%

For **dimers 1**, **2** and **3**, as we can see in the next Figures (Figure 65, Figure 66 and Figure 67) the NH_2 group is located nearby the N atom of the Aza[6]helicene. Hence, such geometries are strongly guided by the formation of the hydrogen bond interaction, with minor contribution from π-stacking.

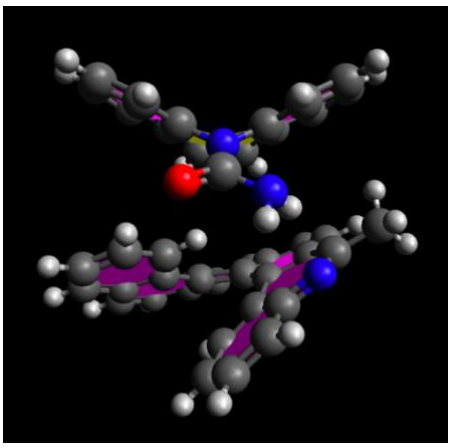

Figure 65: visual representation of the interacting Dimer 1. For the AB CBZ-Aza[6]helicene case.

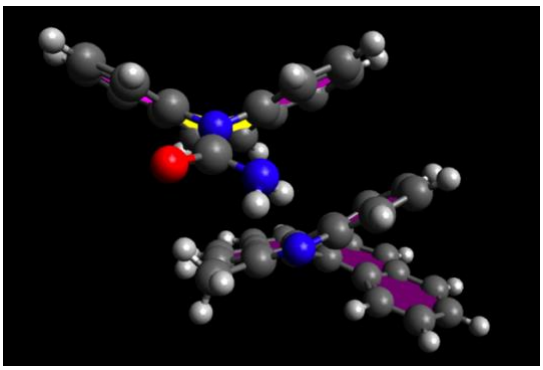

Figure 66: visual representation of the interacting Dimer 2. For the AB CBZ-Aza[6]helicene case.

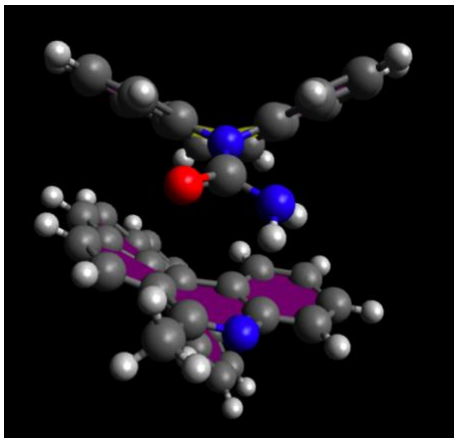

Figure 67 visual representation of the interacting Dimer 3. For the AB CBZ-Aza[6]helicene case.

In **dimer 4** (Figure 68) the NH_2 group is facing one of the aromatic rings of the Aza[6]helicene. Hence, I can deduce the presence of NH_2-π interaction as the relevant one, instead of the hydrogen bond. π-stacking should also contribute to a minor extent.

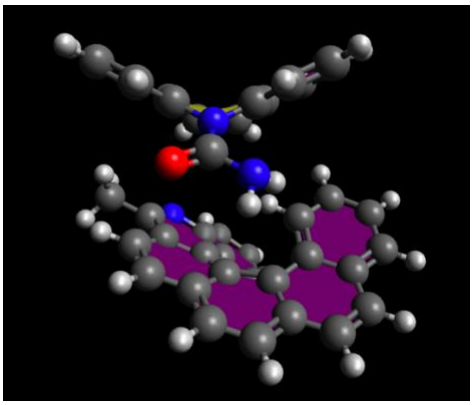

Figure 68: visual representation of the interacting Dimer 4. For the AB CBZ-Aza[6]helicene case.

For **dimers 5** and **6** (Figure 69 and Figure 70) CBZ is facing in the opposite direction with respect to all the previous geometries. However, dimer 5 is peculiar because the nitrogen of the NH_2 group and the oxygen of CBZ interact with the CH bonds of the methyl group of Aza[6]helicene and this process is tilting the whole molecule. This is not the case of dimer 6 where the only relevant contribution is π-stacking.

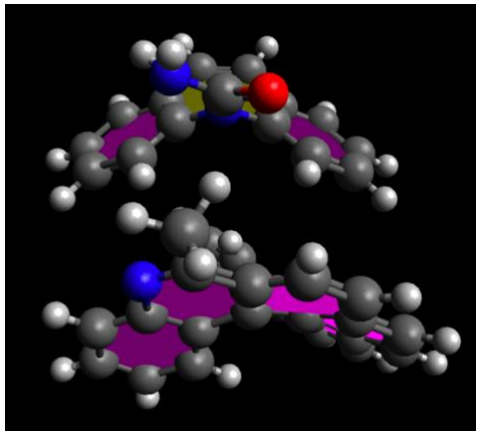

Figure 69: visual representation of the interacting Dimer 5. For the AB CBZ-Aza[6]helicene case.

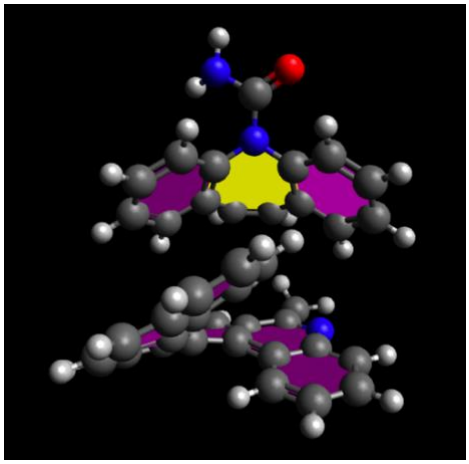

Figure 70: visual representation of the interacting Dimer 6. For the AB CBZ-Aza[6]helicene case.

5.7. Aza[6]helicene: comparing the AA vs. the AB pairs

As one can see by comparing Figure 71 against Figure 72 in the AB case there is a significant shift towards more negative interaction energy values (W_i). As I have done in the previous case of Aza[5]helicene (section 5.5), I use the Boltzmann distribution data to plot in Figure 73 the probability (f_i) as a function of the interaction energy (W_i). As expected, the plot has an exponential behavior, but it allows to quickly judge the interaction differences between AA and BB.

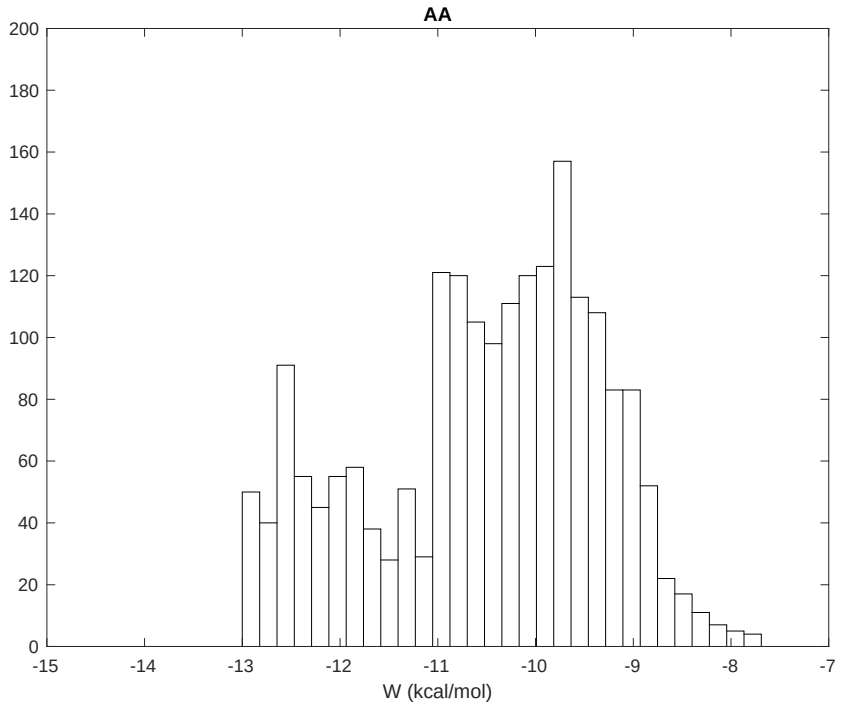

Figure 71: histogram of the OBSolv simulated dimers as a function of he interaction energy (W) for the case AA CBZ-Aza[6]Helicene. The histogram was obtained on a data set of 2000 independent simulations.

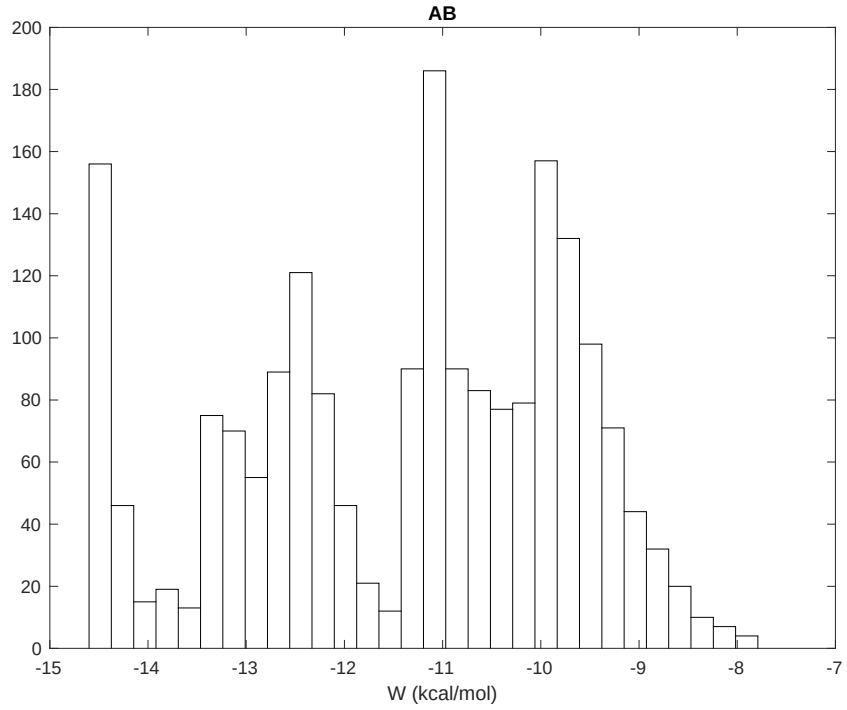

Figure 72: histogram of the OBSolv simulated dimers as a function of he interaction energy (W) for the case AB CBZ-Aza[6]Helicene. The histogram was obtained on a data set of 2000 independent simulations.

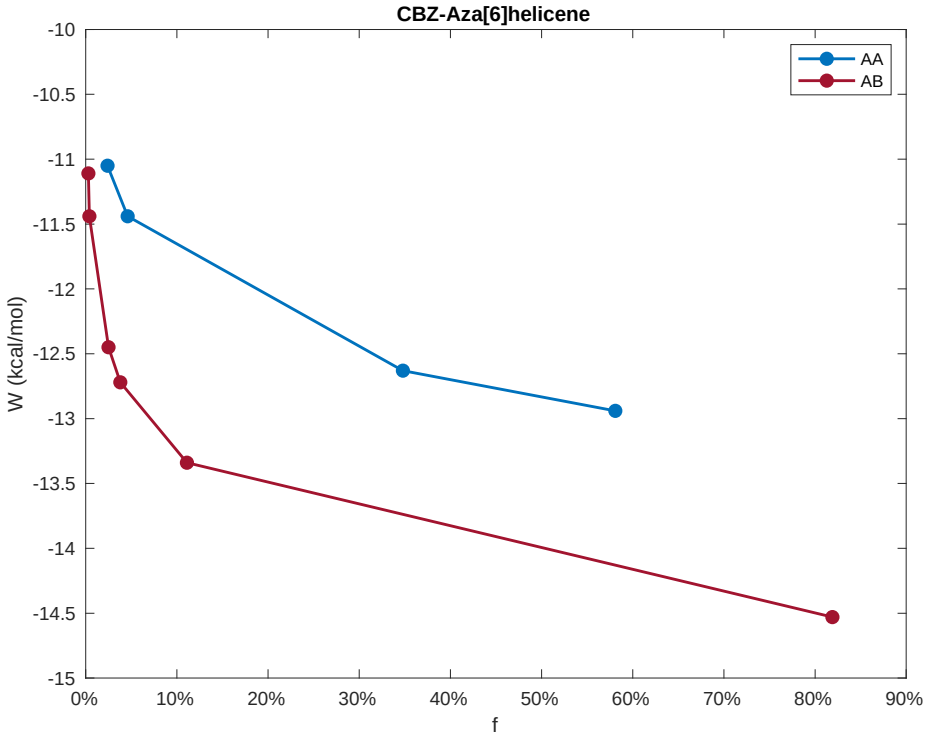

Figure 73: graph of the Boltzmann distribution (f) as a function of the interaction energy (W) for AA and AB in the case of CBZ-Aza[6]helicene.

As one can see in Figure 73, the energy difference between the two more stable conformer for the two different cases (AA and AB) is of the order of 1.5 kcal/mol. This is an appreciable difference, that strongly suggest that chiral recognition of CBZ using Aza[6]helicene may be possible. Is also interesting to underline that for the most stable dimers of the two cases it is $f_{AB} \approx 80\%$ and $f_{AA} \approx 60\%$. Hence for those cases I can assume that the SERS signal will be mainly due to the more stable conformers. Chiral recognition is strongly suggested from such data, but to confirm this behavior, one should find appreciable differences in the Raman spectra.

5.8. Interaction energy distribution: the difference between Aza[5]helicene & Aza[6]helicene

By comparing the Aza[6]helicene interactions with respect to the Aza[5]helicene interactions, the shift of the interaction energy is at the opposite for the two cases AA and AB. For the AA case, CBZ-Aza[6]helicene is shifted towards higher interaction energy values with respect to CBZ-Aza[5]helicene (Figure 74). However, for the AB case, CBZ-Aza[6]helicene is shifted towards more negative interaction energy values with respect to CBZ-Aza[5]helicene (Figure 74).

For CBZ-Aza[5]helicene the two cases AA and AB almost overlap, and the only difference is the probability (f_i) for the two most stable structures (as discussed in section 5.5). Instead, for CBZ-Aza[6]helicene the two cases AA and AB significantly split, as discussed above. Therefore, one can deduce that the additional aromatic ring of Aza[6]helicene increases the π-stacking interaction, which affects the competition among the possible intermolecular interactions. This leads the formation of different geometries that are more stable for the AB case and less stable for the AA case. For this reason, the two cases split in the energy *vs*. probability diagram.

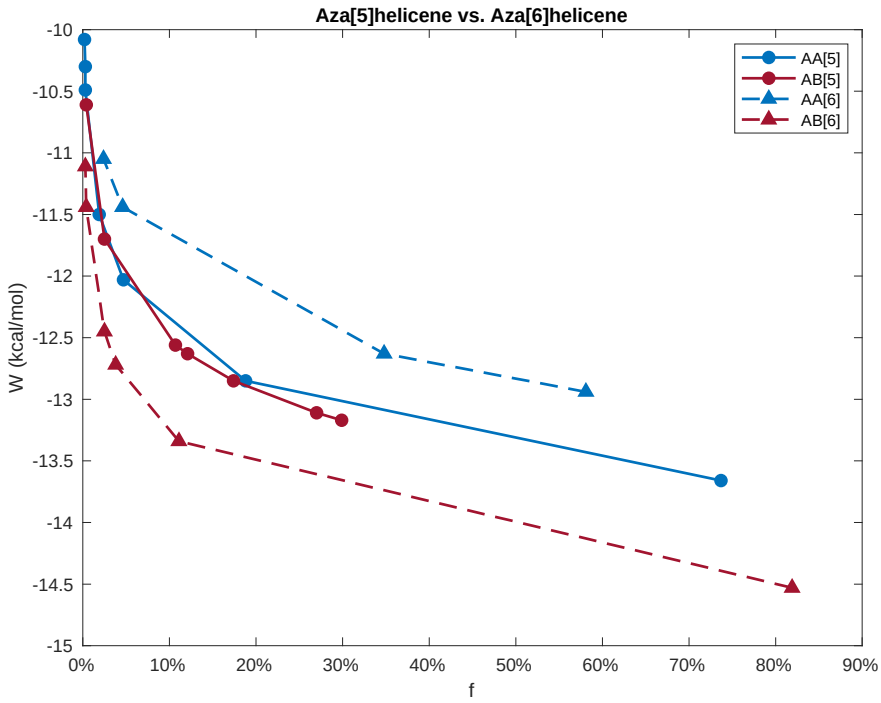

Figure 74: graphical comparison of the Boltzmann distribution (f) as a function of the interaction energy (W) between the two cases CBZ-Aza[5]helicene and CBZ-Aza[6]helicene.

6. Simulation of Raman & UV-Vis spectra of selected interacting dimers

6.1. Minimum energy geometries

The aim of this chapter is to study and compare the UV-Vis and Raman spectra, calculated by Density Functional Theory – DFT (B3LYP/6-31G(d,p) method, including Grimme's D3 correction for Van Der Waals interactions), of selected CBZ-Azahelicene dimers to understand if there are appreciable differences between the cases and therefore if it is possible to achieve chiral recognition.

First of all, for the sake of simplicity, I have selected the lowest energy dimers of each case and I have used them as the starting point for the subsequent DFT calculations. The geometric differences between the dimers simulated by OBSolv and those calculated by DFT are shown in the following images and tables.

Those differences are due to the fact that OBSolv use a Molecular Mechanics approach where the quantum aspects are neglected. However, DFT is a microscopic quantum theory for the study of many electron systems. DFT uses as the fundamental quantity the electronic charge density, that is the square of the wave function (Hohenberg 1964).

I begin with the AA and AB cases of Aza[5]helicene. In Figure 75 and Figure 76 are reported, on the left the OBSolv simulated dimers, and on the right the equilibrium geometries computed by DFT.

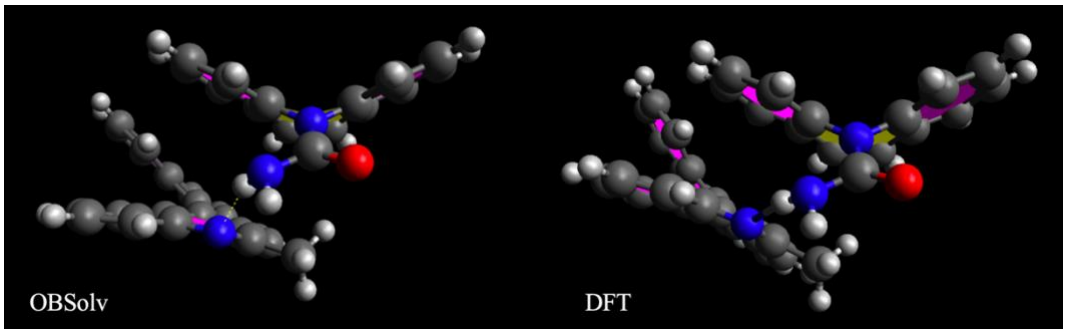

Figure 75: visual representations of the interaction dimer for the case AA CBZ-Aza[5]helicenen. Simulated, on the left with OBSolv and on the right with DFT.

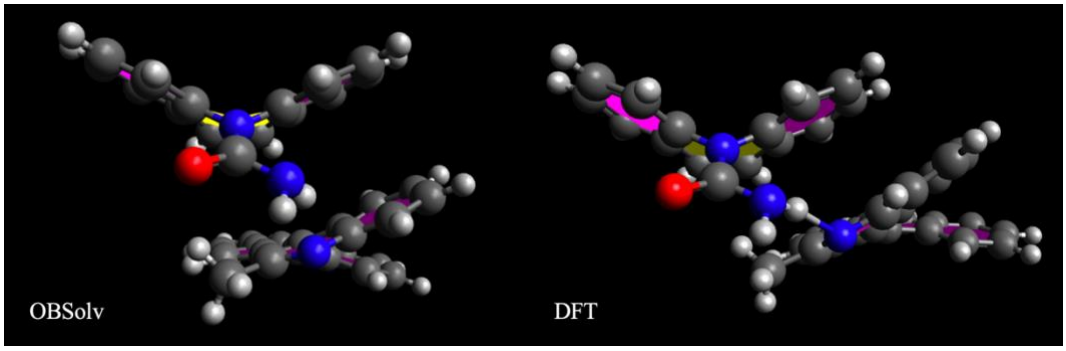

Figure 76: visual representations of the interaction dimer for the case AB CBZ-Aza[5]helicenen. Simulated, on the left with OBSolv and on the right with DFT.

The compared geometries are very similar but not identical. The DFT calculation tend to favor the hydrogen bond formation. To compare in more details the two geometries I measured the N…H distance and N…HN angle also for the DFT geometries.

Another interesting parameter for DFT is the $NH_{DFT\ bonded}$ which is the bond distance of the hydrogen-bonded NH atoms of the NH_2 group of CBZ (Figure 77).

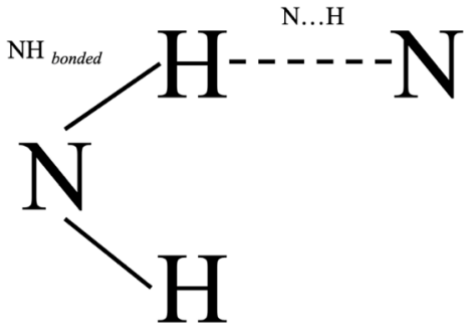

Figure 77: schematic representation of the N…H distance and the NH bonded distance.

The data are reported in Table 27.

Table 27: Boltzman distribution (f_i), interaction energy (W_i), ID and geometric parameters calculate by OBsolv and DFT of the minimum energy dimer configuratio for the two cases AA and AB in the CBZ-Aza[5]helicene system. The ID code is reported for archive purposes.

	ID	*W_i (kcal/mol)*	*N...H (Å)*	*N...HN (°)*	*$N...H_{DFT}$ (Å)*	*$N...HN_{DFT}$ (°)*	*$NH_{DFT\ bonded}$ (Å)*	*f_i*
AA	141695	-13.66	1.943	157	1.937	170	1.026	74%
AB	61953	-13.17	2.066	152	1.992	169	1.022	30%

For the AA case, the difference in the N…H distance computed by OBSolv and the distance computed by DFT is almost negligible. For the AB case the difference between the two methods is larger, around 0.074 Å, but still small.
In both cases (AA, AB) the N…HN angle computed by DFT (~170°) is larger than the one computed by OBSolv. In other words, DFT leads to a higher directionality of the hydrogen bond.

As we can also see, from Table 27, the $NH_{DFT\ bonded}$ distance is similar for both the cases. Ideally the best condition for hydrogen bond is achieved when the N…H and the $NH_{DFT\ bonded}$ are similar. So, one can say that for AA the interaction is higher than for AB, because the two values are more similar. As will be discussed later on, this different strength of the hydrogen bond produces a different spectral shape in the NH stretching region (see section 6.2).

I made the same comparisons also for the AA and AB cases of Aza[6]helicene (Figure 78, Figure 79 and Table 28). Also, in such cases the geometries from DFT are very similar but not identical to those obtained by OBSolv, and the comparisons of the hydrogen bonding parameters (NH & NHN) is useful.

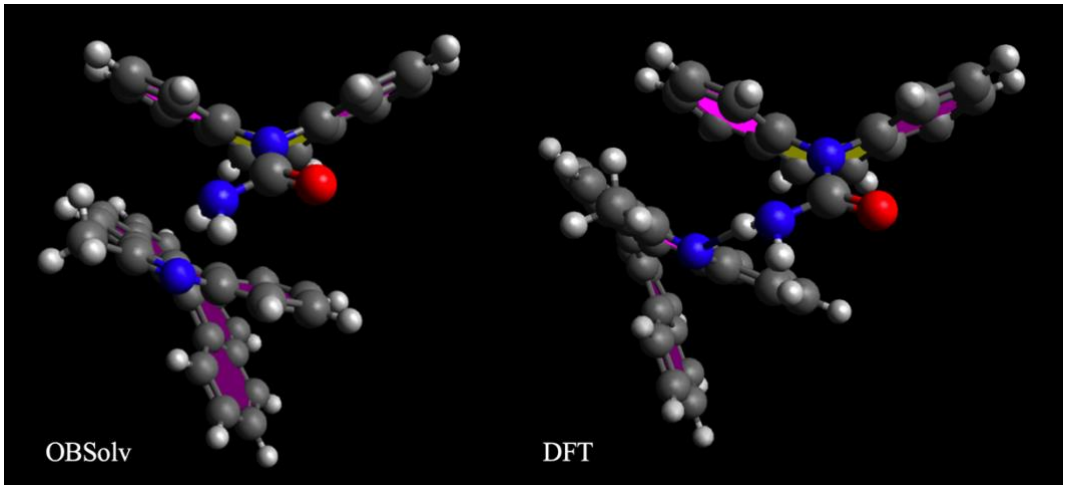

Figure 78: visual representations of the interaction dimer for the case AA CBZ-Aza[6]helicenen. Simulated, on the left with OBSolv and on the right with DFT.

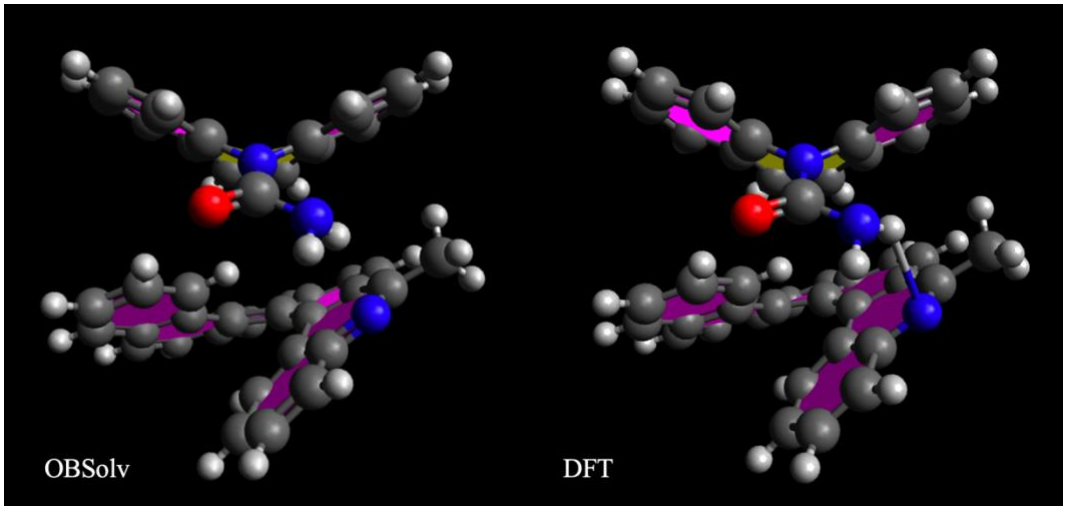

Figure 79: visual representations of the interaction dimer for the case AB CBZ-Aza[6]helicenen. Simulated, on the left with OBSolv and on the right with DFT.

Table 28: Boltzman distribution (f_i), interaction energy (W_i), ID and geometric parameters calculate by OBsolv and DFT of the minimum energy dimer configuratio for the two cases AA and AB in the CBZ-Aza[6]helicene system. The ID code is reported for archive purposes.

	ID	*W_i (kcal/mol)*	*N...H (Å)*	*N...HN (°)*	*$N...H_{DFT}$ (Å)*	*$N...HN_{DFT}$ (°)*	*$NH_{DFT\ bonded}$ (Å)*	*f_i*
AA	118867	-12.94	2.026	157	1.993	170	1.023	58%
AB	13591	-14.53	2.151	143	2.316	128	1.011	82%

The difference of the N…H distances between OBSolv and DFT is around 0.033 Å for AA and it is around -0.165 Å for AB. Not only this variation is very high for AB compared to the other cases but also leads to a N…H_{DFT} value that is higher than N…H. As for the N…HN angle, in the AA case it is 170° (as the previous cases); on the contrary, for AB, the N…HN angle decreases from 143° calculated by OBSolv to 128° calculated by DFT. Even $NH_{DFT\ bonded}$ is much smaller for AB. All such data suggest that the hydrogen bond contribution for AB will be smaller than in the other cases. Hence, I expect a lower shift in the NH stretching peak of upon the formation of the hydrogen bond in the CBZ-Aza[6]helicene AB case as can be seen in section 6.2.

Indeed, this anticipation is verified by the simulated Raman spectra reported in the next Section.

6.2. Raman spectra simulated by DFT

In this Section I compare the Raman spectra simulated by DFT for the different interactions between CBZ and Azahelicenes, so as to assess any appreciable spectroscopic difference that could be used to provide chiral recognition for our systems. To investigate in detail the vibrational modes and assign them in the spectra is beyond the purpose of this thesis. I have used Matlab to plot the DFT data for the most stable dimers of Aza[5]helicene and Aza[6]helicene with CBZ in the AA and AB cases. I have considered two spectroscopic ranges, the first in the fingerprint region (600-1800 cm^{-1}) and the second in the CH/NH stretching region (2800-4000 cm^{-1}), where we expect to detect the spectroscopic effects of hydrogen bonding.

The following figures are made up of four quadrants that display the spectra, to simplify the comparison. From top to bottom, in the first and second quadrants there are respectively the AA spectra and the AB spectra (purple line), together with the "fictitious" spectrum resulting from the algebraic sum of the spectrum of CBZ and Azahelicene (yellow line). The third and fourth quadrant show respectively the spectrum of Azahelicene (red) and CBZ (blue) computed for the isolated molecules.

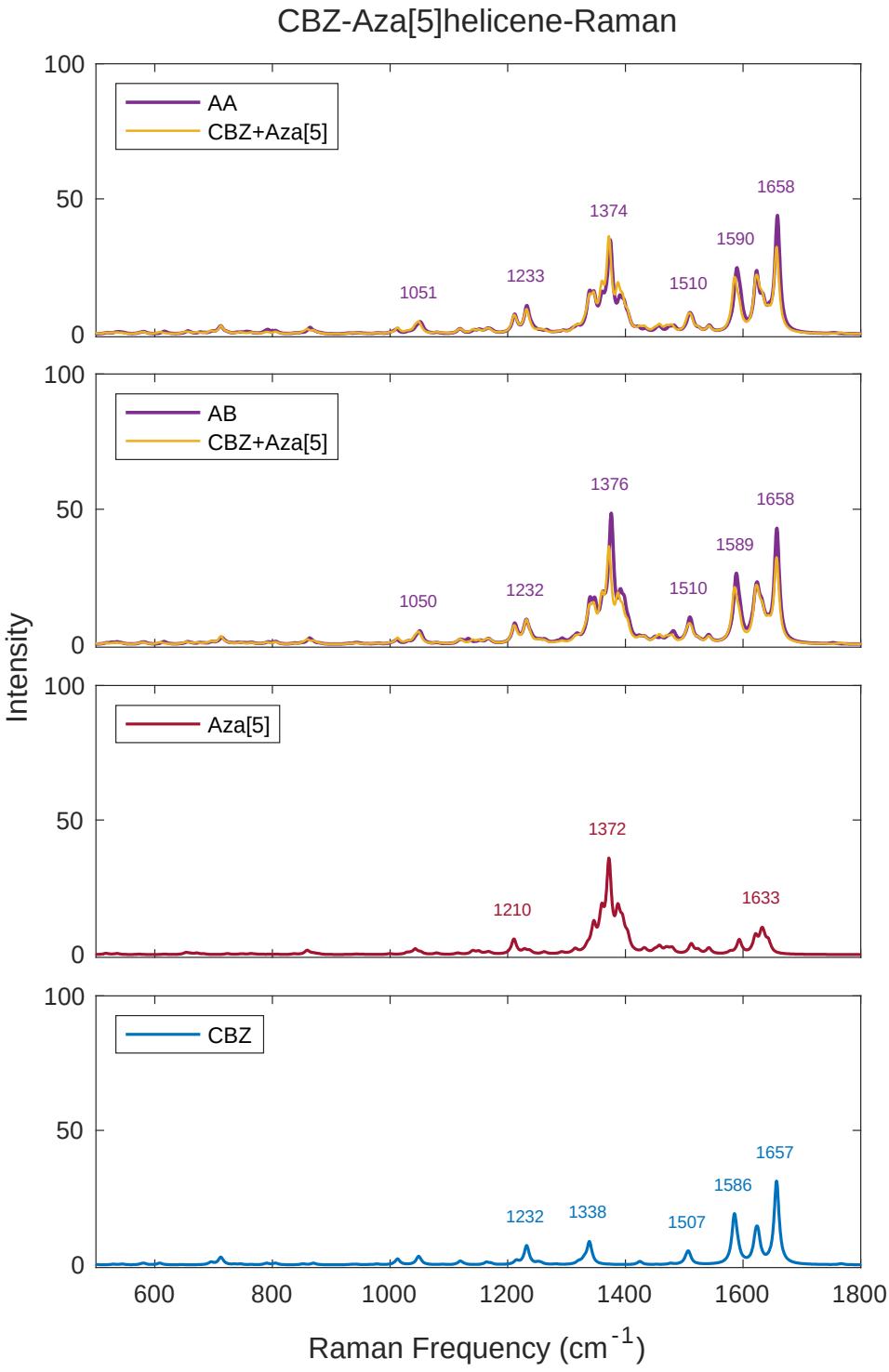

Figure 80: Raman spectra of the case CBZ-Aza[5]helicene, from top to bottom: comparison of case AA and the sum of the two isolated molecules, comparison of case AB and the sum of the two isolated molecules, isolated Aza[5]helicene, isolated Carbamazepine.

By examining Figure 80 (CBZ-Aza[5]helicene) one can see that the two cases AA and AB are very similar to each other. The only appreciable difference is the intensity of the peak at 1374 cm^{-1}, which is higher in the AB case than in the AA case.

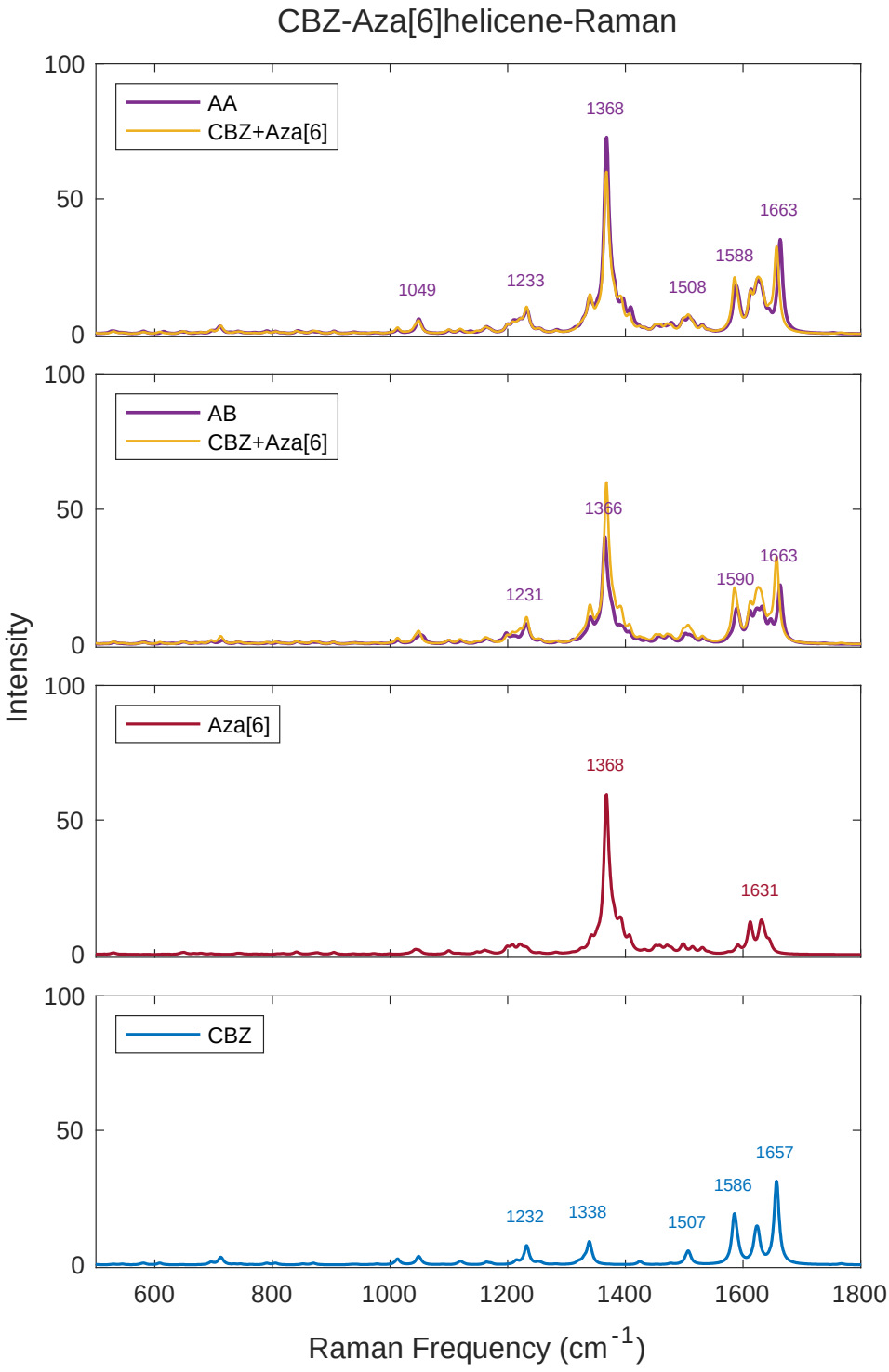

Figure 81: Raman spectra of the case CBZ-Aza[6]helicene, from top to bottom: comparison of case AA and the sum of the two isolated molecules, comparison of case AB and the sum of the two isolated molecules, isolated Aza[6]helicene, isolated Carbamazepine.

In Figure 81 (CBZ-Aza[6]helicene) one observes, for both the AA and AB case, a shift of the peaks at 1590 cm^{-1} and 1663 cm^{-1} towards higher wavenumber with respect to the yellow spectra that are given by the sum of the spectra of isolated CBZ and Aza[6]helicene. This shift reflects the interaction mechanisms active in CBZ-Aza[6]helicene that are stronger than in CBZ-Aza[5]helicene (where the effect was not evident). Also, the peak at 1366 cm^{-1} is noticeable, as it is more intense in the AA case than the AB case.

It is also interesting to observe that for Aza[5]helicene the peak around 1374 cm^{-1} has an intensity variation between the two cases (AA & AB) which is opposite to the intensity variation observed for the Aza[6]helicene peak around 1366 cm^{-1}. This matches the different behavior of the two Azahelicenes that was discussed in section 5.8.

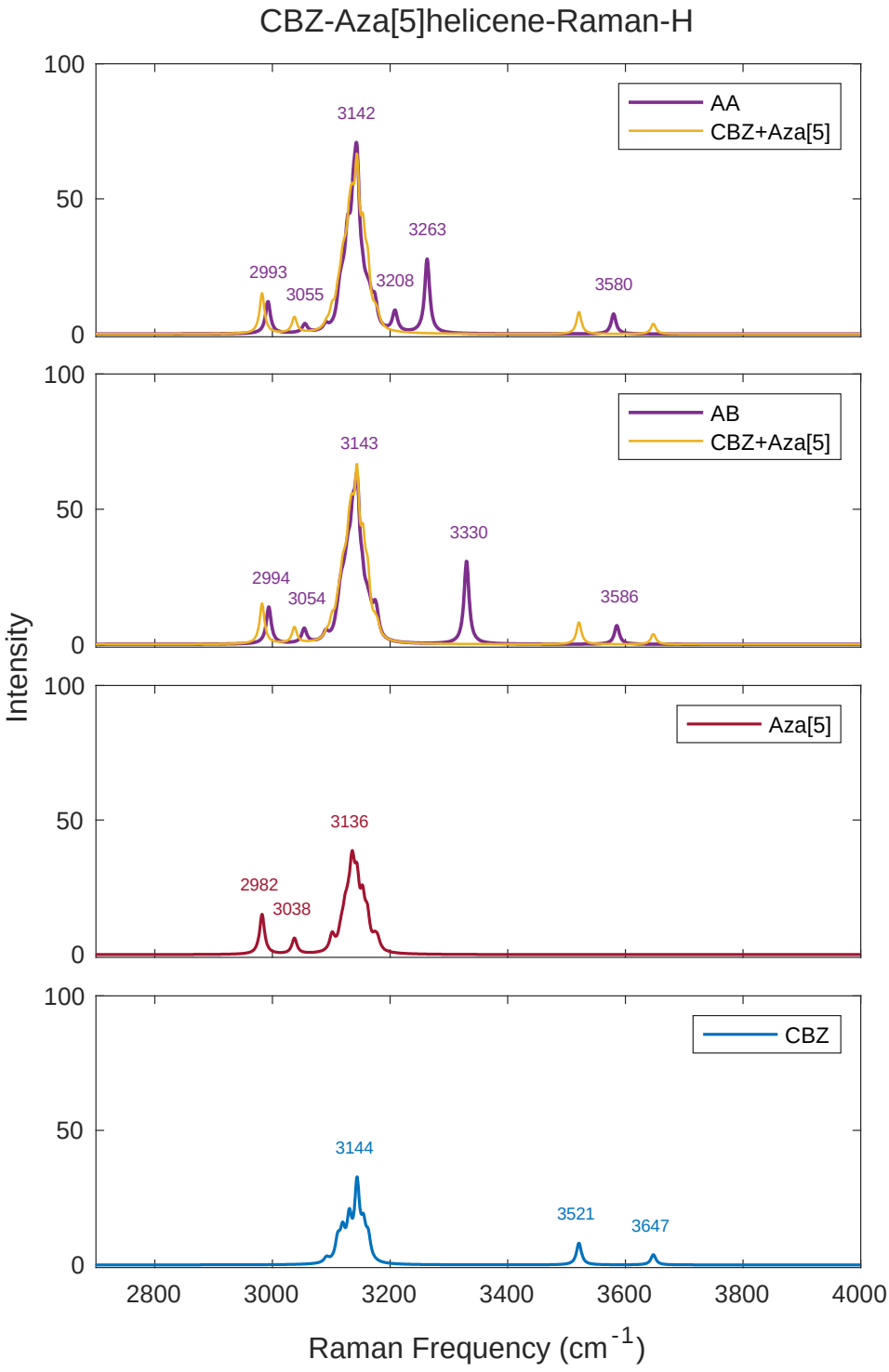

Figure 82: Raman spectra in the higher frequency range of the case CBZ-Aza[5]helicene, from top to bottom: comparison of case AA and the sum of the two isolated molecules, comparison of case AB and the sum of the two isolated molecules, isolated Aza[5]helicene, isolated Carbamazepine.

In Figure 82 (CBZ-Aza[5]helicene) we are looking at the spectroscopic effects of hydrogen bond in our systems. As one can see, in both cases (AA & AB) there is an appreciable shift of the NH stretching peaks of CBZ (located at 3647 cm^{-1} and 3521 cm^{-1}) this confirms the hydrogen bond interaction. The first peak (3647 cm^{-1}) in both cases, is shifted towards 3585 cm^{-1}. More interestingly the second peak (3521 cm^{-1}) is shifted more towards low wavenumber for the AA case (3263 cm^{-1}) than for the AB case (3330 cm^{-1}). The peak observed at 3208 cm^{-1} in AA is not present for AB. It is probably shifted towards low wavenumber and disappears within the features located at 3143 cm^{-1}.

In principle, such appreciable spectroscopic differences could be used for chiral recognition.

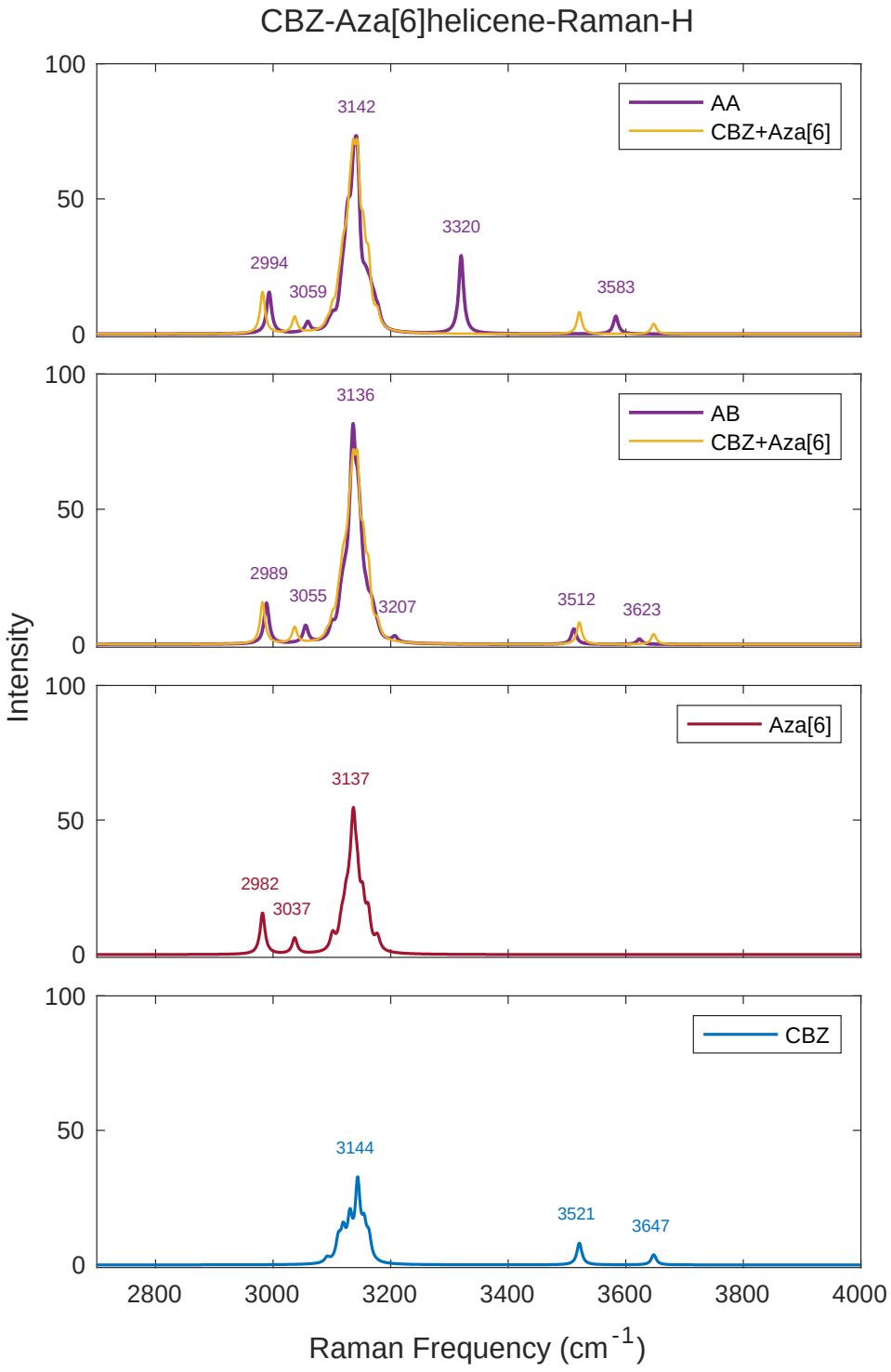

Figure 83: Raman spectra in the higher frequency range of the case CBZ-Aza[6]helicene, from top to bottom: comparison of case AA and the sum of the two isolated molecules, comparison of case AB and the sum of the two isolated molecules, isolated Aza[6]helicene, isolated Carbamazepine.

As expected, one can see in Figure 83 (CBZ-Aza[6]helicene) that for the AB case the shift of the peaks at 3623 cm^{-1} and 3512 cm^{-1} is very small compared to the AA case. This is in line with the data in Table 28 that tell us that for this peculiar case the hydrogen bond is almost negligible. Hence, chiral recognition is very probably achievable for this kind of configuration by using Aza[6]helicene.

6.3. UV-Vis spectra simulated by DFT

I compare in this section the UV-Vis spectra simulated by DFT for the different interactions of CBZ with Aza[5]helicene and Aza[6]helicene to establish if there are appreciable spectroscopic difference that could be used to provide chiral recognition for our systems. Similar to the case of the Raman spectra, in the following Figures one can see four quadrants that show different spectra for comparison. From the top to the bottom, in the first and second quadrants there are respectively the simulated spectra of the AA and AB dimers (purple line), together with the "fictitious" spectra resulting from the algebraic sum of the spectra of the isolated CBZ and Azahelicenes spectra (yellow line). In the third and fourth quadrant there are respectively the spectra of Azahelicene (red) and CBZ (blue) calculated in the isolated condition. The spectra are studied in the 200-500 nm range, which is the one most accessible to ordinary spectrometers.

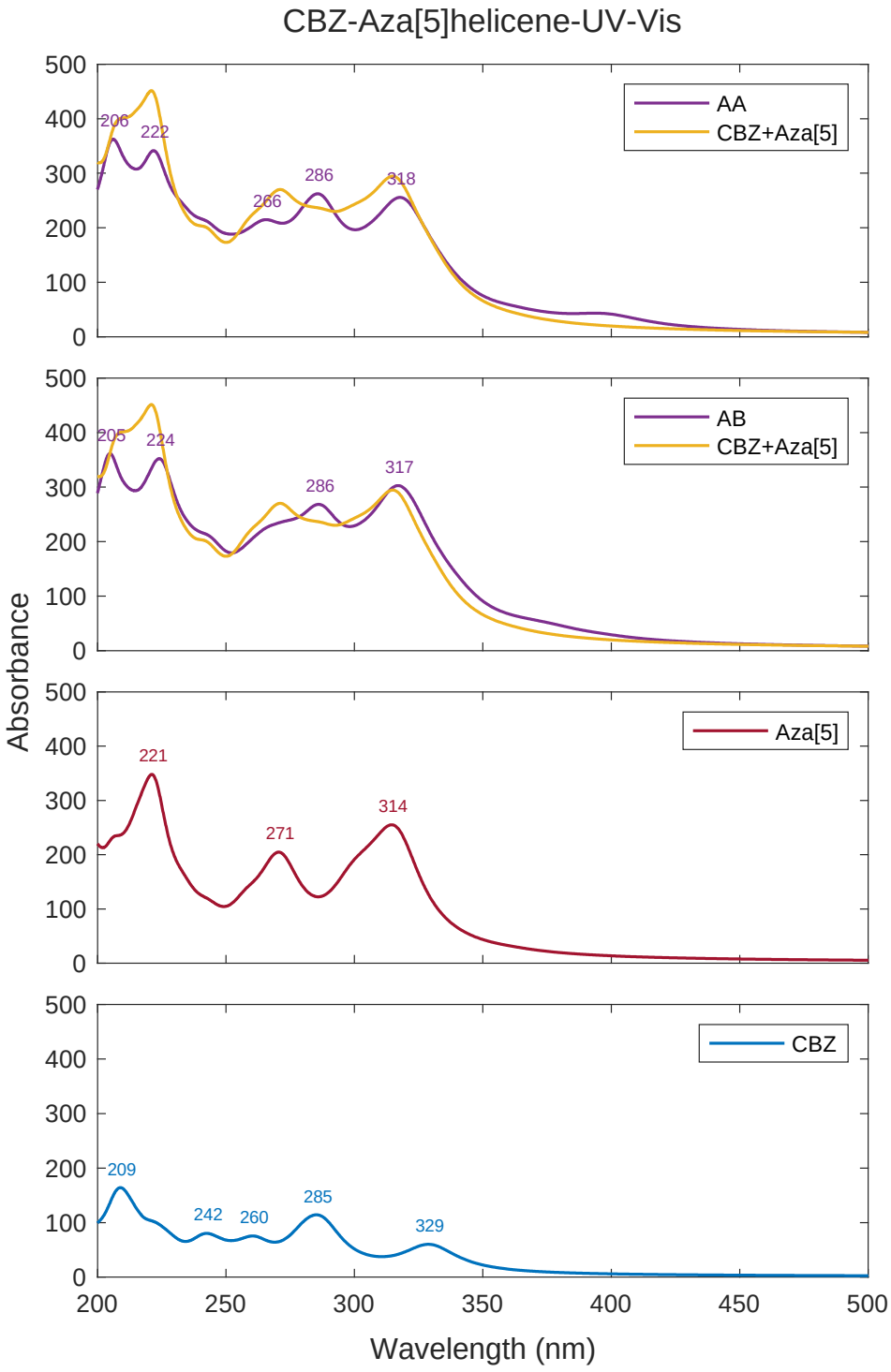

Figure 84: UV-Vis spectra of the case CBZ-Aza[5]helicene, from top to bottom: comparison of case AA and the sum of the two isolated molecules, comparison of case AB and the sum of the two isolated molecules, isolated Aza[5]helicene, isolated Carbamazepine.

As reported in Figure 84 (CBZ-Aza[5]helicene) the AA and AB cases are quite similar. There are no major differences in the relevant peaks, and the intensity pattern are very similar. The only difference can be seen around 400 nm where only for AA a very wide and low intensity peak can be noticed.

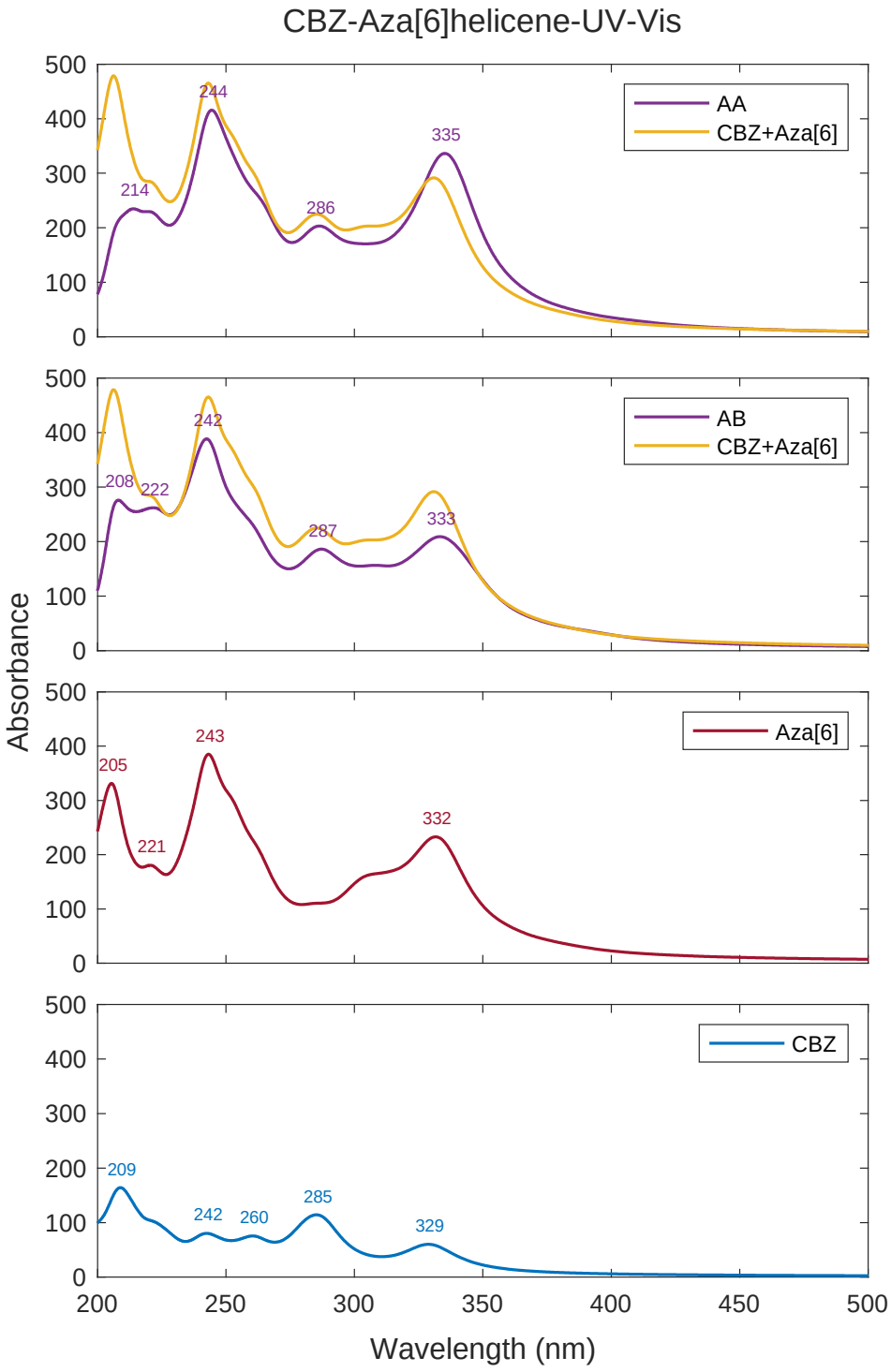

Figure 85: UV-Vis spectra of the case CBZ-Aza[6]helicene, from top to bottom: comparison of case AA and the sum of the two isolated molecules, comparison of case AB and the sum of the two isolated molecules, isolated Aza[6]helicene, isolated Carbamazepine.

The CBZ-Aza[6]helicene case is reported in Figure 85: also here the simulated spectra of the two cases (AA & AB) are quite similar. The only difference is the variation in intensity of the peak at 335 nm. Also, contrary to CBZ-Aza[5]helicene, there are no observable effects in the 400 nm region.

In conclusion, chiral recognition by using UV-Vis measurements could be possible for the case of Aza[5]helicene considering the 400 nm region.

7. Conclusion and perspectives

In my thesis I investigated in details the intermolecular interaction between two Azahelicenes (namely, Aza[5]helicene and Aza[6]helicene) and Carbamazepine (CBZ). In particular, I focused my attention on the possible chiral recognition by Aza[5]helicene and Aza[6]helicene interacting with the two geometric enantiomers of CBZ. I have established a theoretical approach for studying such interactions, and highlight their possible spectroscopic signatures, along the following steps:

1. definition of the possible conformers of the selected molecules;
2. use of a molecular mechanics-based Monte Carlo code (OBSolv) to simulate a large number of interacting dimers for each selected conformers.
3. inspection of the most stable dimers and selection of the intermolecular geometric parameters (lengths, angles) that are associated with the interaction mechanisms;
4. representation of the interaction energy as a function of the selected parameters to identify local minima in the potential energy landscape which describe dimers of interest for assessing the intermolecular interaction mechanism and geometry;
5. compute the Boltzmann distribution of such local minima geometries to identify the dimers that are more likely to form at room temperature;
6. finally, use DFT to simulate the Raman and UV-Vis spectra of the selected stable dimers.

Clearly, the approach sketched above is modular and it can be easily adapted to different application scenarios. For instance, the definition of the parameters describing intermolecular interaction can be easily adapted to other molecules.

Another important result on the methodological side is that I have checked the capability of OBSolv to correctly deal with different conformational isomers. In particular the AA and BB interaction pairs (as well as AB and BA) are computed as degenerate by OBSolv, as expected.

Form the distribution of interaction energies (section 5.7) and simulated spectra (section 6.2) of the CBZ-Aza[6]helicene pair, I have deduced the possibility of chiral recognition, with observable spectroscopic signatures of intermolecular interaction. As for the case of CBZ-Aza[5]helicene, chiral recognition is less evident, and it could be easier to detect intermolecular interaction signatures in UV-Vis spectroscopy (section 6.3) rather than Raman.

In general, the spectroscopic effects due to chirality-specific intermolecular interaction may be difficult to detect due to low signal, especially in low-concentration molecular sensing applications. This issue can be addressed by using Azahelicene-functionalized gold SERS substrates, as described in (Zanchi 2018). Moreover, the functionalization of the SERS substrates with Azahelicenes can increase the Raman enhancement mechanism for CBZ by adding the contribution of chemisorption to the total enhancement thanks to the formation of stable interaction between the two molecules (i.e., hydrogen bonding, and π-stacking that may lead to excitonic effects, possibly coupled to the plasmonic resonances of the SERS substrate). Of course, experimental studies are required to assess the validity of such theoretical concepts.

Furthermore, the theoretical approach developed here can be used to screen a range of possible candidate analytes (e.g., AEDs) before focusing on the synthesis of the most suitable molecule to be used for functionalizing the plasmonic substrates. This kind of theoretically assisted molecular design would be useful to optimize the number of experiments and the time required.

References

Agarwal, N.R. et al. 2014. "SERS activity of silver and gold nanostructured thin films deposited by pulsed laser ablation." *Applied Physics A* 347–351.

Albrecht, MG., Creighton, JA. 1977. "J. AM. Chem. Soc." 99: 5215.

Anslyn, E. V. 2004. *Modern Physical Organic Chemistry.* university science books.

Bazzini, C. et al. 2005. "Synthesis and Characterization of Some Aza[5]helicenes." *European Journal of Organic Chemistry* (7): 1247–1257.

Beghi, E. 2004. "Efficacy and tolerability of the new antiepileptic drugs: comparison of two recent guidelines." *The Lancet Neurology* 3 (10): 618-621.

Biedermann, P. U. et al. 1999. "Inversion Barrier of Corannulene. A Benchmark for Bowl-to-Bowl Inversions in Fullerene Fragments." *The Journal of Organic Chemistry* 64 (10): 3655–3662.

EMA, European Medicines Agency. 2012. "Perampanel (Fycompa): assessment report (EMA/424476)." https://www.ema.europa.eu/en/documents/assessment-report/fycompa-epar-public-assessment-report_en.pdf.

Emsley, J. 1980. "Very strong hydrogen bonding." *Chemical Society Reviews* 9 (1): 91.

Fleischmann, M., Hendra, PJ., McQuillan, AJ. 1974. "Chem. Phys. Lett." 26: 123.

Gross, A. S. 1998. "Best practice in therapeutic drug monitoring." *British journal of clinical pharmaology* 95-99.

Halgren, T. A. 1995. "Merck Molecular Force Field. II. MMFF94 van der Waals and Electrostatic Parameters for Intermolecular Interactions." *Journal of Computational Chemistry* 17 (5 & 6): 520-552.

Hanwell, M. D. et al. 2012. "Avogadro: an advanced semantic chemical editor, visualization, and analysis platform." *Journal of Cheminformatics* 4 (1).

Hohenberg, P., Kohn, W. 1964. "Inhomogeneous Electron Gas." *Physical review* 136 (3B): B864-B871.

Hunter, C. A. et al. 2001. "Aromatic interactions." *Journal of the Chemical Society, Perkin Transactions* 2 (5): 651–669.

Hunter, C. A., Sanders, J. K. M. 1990. "The nature of π–π interactions." *Journal of the American Chemical Society* 112 (14): 5525-5534.

Jallon, P. 1997. "Epilepsy in Developing Countries." *Epilepsia* 38 (10): 1143-1151.

Jaworska, A. et al. 2016. "Potential of Surface Enhanced Raman Spectroscopy (SERS) in Therapeutic Drug Monitoring (TDM). A Critical Review." *Biosensors* 6 (3): 47.

Jeanmaire, DL., Duyne, RPV. 1977. "J. Electroanal. Chem." 84: 1.

Jensen, F. 2006. *Introduction to Computational Chemistry.* John Wiley & Sons Ltd.

Johannessen, S. I. et al. 2003. "Therapeutic Drug Monitoring of the Newer Antiepileptic Drugs." *Therapeutic Drug Monitoring* 25 (3): 347-363.

Kang, J. S., Lee, M. H. 2009. "Overview of Therapeutic Drug Monitoring." *The Korean Journal of Internal Medicine* 24 (1): 1-10.
Kneipp, K. 2001. "Lectures." *M.I.T.*
Kryger, G. et al. 1999. "Structure of acetylcholinesterase complexed with E2020 (Aecicept): implications for the design of new anti-Alzheimer drugs." *Struccture* 7 (3): 297-307.
Larson, J. W., McMahon, T. B. 1984. "Gas-phase bihalide and pseudobihalide ions. An ion cyclotron resonance determination of hydrogen bond energies in XHY- species (X, Y = F, Cl, Br, CN)." *Inorganic Chemistry* 23 (14): 2029–2033.
Mark, A. E., van Gunsteren, W. F. 1994. "Decomposition of the Free Energy of a System in Terms of Specific Interactions Impications for Theoretical and Experimental Studies." *Journal of Molecular Biology* 240: 167-176.
Mercolini, L. et al. 2010. "Simultaneous HPLC-F analysis of three recent antiepileptic drugs in human plasma." *Journal of Pharmaceutical and Biomedical Analysis* 53 (1): 62-67.
Meyer, A. et al. 2010. "Global disparities in the epilepsy treatment gap: a systematic review." *Bulletin of the World Health Organization* 88 (4): 260-266.
Minisci, F. et al. 1993. "Polar Effects in Free-Radical Reactions. Homolytic Heteroaromatic Substitutions by Alkyl Bromides." *Journal of Organic Chemistry* 58: 4207-4211.
Morrison, C. A., Siddick, M. M. 2003. "Determining the Strengths of Hydrogen Bonds in Solid-State Ammonia and Urea: Insight from Periodic DFT Calculations." *Chemistry - A European Journal* 9 (3): 628–634.
Nicholls, A. et al. 2008. "Predicting Small-Molecule Solvation Free Energies: An Informal Blind Test for Computational Chemistry." *Journal of Medicinal Chemistry* 51 (4): 769-779.
O'Boyle, N. M. et al. 2011. "Open Babel: An open chemical toolbox." *Journal of Cheminformatics* 3 (1).
Pistaffa, M. 2017. "SERS Assessment of Carbamazepine in Methanol, Chloroform and Patient Samples using Au sensors with optimized surface nanostructure." *Master Thesis. Politecnico di Milano.*
Quintas, R. et al. 2012. "Psychosocial difficulties in people with epilepsy: A systematic review of literature from 2005 until 2010." *Epilepsy & Behavior* 25 (1): 60-67.
The Open Babel Package, version 2.3.1,. 2011. Oct. http://openbabel.org.
Thomas, A. et al. 2002. "Aromatic Side-Chain Interactions in Proteins. II. Near- and Far-Sequence Phe-X Pairs." *Proteins: Structure, Function, and Bioinformatics* 48: 635–644.
Thurman, D. J. et al. 2011. "Standards for epidemiologic studies and surveillance of epilepsy." *Epilepsia* 52 (Suppl. 7): 2-26.
Tommasini, M. et al. 2019. "Laser-Synthesized SERS Substrates as Sensors toward Therapeutic Drug Monitoring." *nanomaterials* 9 (5): 677.
Wakelin, L. P. G. 1986. "Polyfunctional DNA intercalating agents." *Medicinal Research Reviews* 6 (3): 275-340.

Zanchi, C. et al. 2019. "Evaluation of Molecular Polarizability and of Intensity Carrying Modes Contributions in Circular Dichroism Spectroscopies." *Applied Sciences* 9 (21): 4691.
Zanchi, C. et al. 2018. "Functionalization of nanostructured gold substrates with chiral chromophores for SERS applications: The case of 5-Aza[5]helicene." *Chirality* 30 (7): 875–882.
Zanchi, C. et al. 2016. "Laser tailored nanoparticle arrays to detect molecules at dilute concentration." *Applied Surface Science* 396: 1866-1874.
Zhao, Y. et al. 2015. "Conformational Preferences of π–π Stacking Between Ligand and Protein, Analysis Derived from Crystal Structure Data Geometric Preference of π–π Interaction." *Interdisciplinary Sciences: Computational Life Sciences* 7 (3): 211–220.