Ca' Foscari Venezia

Department of Molecular Sciences and Nanosystems

Master Thesis

Modeling and analysis of the $(H_2C_2)_n \cdot (HCN)_m$ gas-phase clusters by ab initio and DFT computational methods for astrochemical applications

Author:
Sebastiano Gaiotti

Supervisor:
Andrea Pietropolli Charmet

Co-supervisor:
Paolo Stoppa

Academic Year 2020/2021

A thesis submitted in fulfilment of the requirements
for the degree of Doctor in Chemistry

in the

Department of Molecular Sciences and Nanosystems

"*An expert is a person who has made all the mistakes that can be made in a very narrow field.*"

Niels Bohr

CA' FOSCARI VENEZIA

Abstract

Sustainable Chemistry and Technologies
Department of Molecular Sciences and Nanosystems

Doctor in Chemistry

Modeling and analysis of the $(H_2C_2)_n \cdot (HCN)_m$ gas-phase clusters by ab initio and DFT computational methods for astrochemical applications

by Sebastiano Gaiotti

Among the different purposes of astrochemistry, there is the study of chemical processes in the Interstellar Medium (ISM), and the corresponding reaction pathways that may lead to the formation of biological molecules. Within this framework, in recent years some gas-phase clusters have been investigated since they may constitute the first steps involved in these reaction networks. The study presented in this thesis aims at illustrate different types of clusters involving hydrogen cyanide and acetylene, which present a lot of interesting interactions including Van der Waals and dispersion forces. For each cluster, the geometry, energy and other parameters have been optimized throughout many *ab initio* and Density Function Theory (DFT) methods (also including some of the semi-empirical Grimme's density functionals). Finally, non-covalent interactions were investigated through the Quantum Theory of Atoms in Molecules (QTAIM) to identify the nature of the interactions that lead to the stability of $(H_2C_2)_n$, $(HCN)_m$ and mixed structures $(H_2C_2)_n \cdot (HCN)_m$.

Contents

List of Figures

List of Tables

Physical Constants

Speed of Light in vacuum	$c_0 = 2.997\,924\,58 \times 10^{8}\,\mathrm{m\,s^{-1}}$ (exact)
Hartree Energy	$a.u. = 4.359\,744\,722\,207\,1(85) \times 10^{-18}\,\mathrm{J}$
Electron Mass	$m_e = 9.109\,383\,701\,5(28) \times 10^{-31}\,\mathrm{kg}$
Proton Mass	$m_P = 1.672\,621\,923\,69(51) \times 10^{-27}\,\mathrm{kg}$
Elementary charge	$e = 1.602\,176\,634 \times 10^{-19}\,\mathrm{C}$ (exact)
Planck constant	$h = 6.626\,070\,15 \times 10^{-34}\,\mathrm{J \cdot s}$ (exact)
Reduced Planck constant	$\hbar = 1.054\,571\,817 \times 10^{-34}\,\mathrm{J \cdot s}$ (exact)
Solar Mass	$M\odot = 1.988\,55(24) \times 10^{30}\,\mathrm{kg}$

The values of the fundamental physical constants reported were taken from the site "The NIST Reference on Constants, Units and Uncertainty" (https://physics.nist.gov/cuu/Constants/). Last Update to Data Content: May 2019.

Chapter 1

Introduction and Aim of this Thesis

1.1 Astrochemistry

Astrochemistry can be seen as the study of composition and reactions of atoms, molecules and ions in space. It has developed in recent decades to understand the chemical evolution of the universe, from the simplest atoms formed after the Big Bang to complex organic molecules. The study of the possible formation pathways of complex molecules and how they can react to give more complex molecular systems are the greatest challenges of astrochemists. Specifically, the aim is to understand the origin of biomolecules' precursors that form the basis of DNA. The first step in studying the chemistry of the universe is to detect chemical species in certain regions, such as the Interstellar Medium (ISM) and Circumstellar Shells (CS). According to The Cologne Database for Molecular Spectroscopy (CDMS) [1], more than 200 molecules have been detected between ISM and CS at date (April 2021), demonstrating the variety of chemistry in the astronomical environment. For these reasons, molecular spectroscopy and computational chemistry are key tools in the hands of astrochemists and play a crucial role in the understanding of the composition and chemical reactivity of the universe. The development of radio telescopes, such as the Atacama Large Millimeter/submillimeter Array (ALMA), and the launch of satellite telescopes for the near-infrared (NIR) and visible regions of the electromagnetic spectrum, have led to an increase in the amount of data available for galaxy studies. This implies the need to improve the reference data that can be collected in Earth laboratories.

The discovery of prebiotic and biological molecules in some meteorites precipitated on Earth [2], prompted astrochemists to search for these molecules also in the ISM, without any success to date. So in the laboratories, they started simulating interstellar conditions owing to understand what were the possible processes that could lead to molecules of biological interest, which would help understand how life in the universe develops. The investigation of astronomical complex organic molecules (astroCOMs), molecules composed of at least six atoms, between carbon, hydrogen and other heavy elements such as oxygen and nitrogen, is greatly supported by quantum-chemical studies [3].

1.2 Birth and death of a star

"A cloud that veils one of nature's secret places. This is a stellar nursery, a place where stars are born. They condense by gravity from gas and dust until their temperatures become so high that they begin to shine. Such clouds mark the births of stars as others bear witness to their deaths." - Carl Sagan, Cosmos: A Personal Voyage

Infancy Stage It all started in the middle of the interstellar clouds, where gases and small solid particles wander to form stellar dust. These regions of space are abundant in the matter, mostly hydrogen, but they are also less dense than the artificial vacuum that humans can create. When a large number of particles, atoms and molecules condense together, the gravity in these points increase and attract more matter to form a dense interstellar cloud. The first small condensed cloud increases its size and the more it grows, the more its

gravitational attraction force attracts other matter to it until reaches a stellar mass (0.1 $M\odot$) in a few hundred thousand years. In the early stages of contraction, the matter is rarefied and transparent, so the heat produced by the compression of the gas is rapidly dissipated, but with increasing density it becomes opaque and the heat remains inside the cloud. The gas reaches a temperature high enough, allowing the fusion of hydrogen nuclei to give helium. Thus begins the longest and most stable life phase of the star. When it manages to reach a condition of equilibrium between the gravitational force, that tends to compress the gases, and the gases pressure.

Adult Stage Nuclear fusion reactions maintain a constant temperature and pressure within the stellar mass, if there were perturbations that led to a decrease or an increase in the cadence of nuclear reactions, a natural mechanism would intervene to re-establish the equilibrium. If the nuclear fusion reactions decrease, the temperature drops, and with it also the pressure leading to an imbalance of forces. Consequently, there will be an immediate compression of the gases, due to the predominant force of gravity, leading to an increment in the core temperature with an increase in the energy produced by the fusion reactions and vice versa. A star lives thanks to the consumption of hydrogen as a nuclear fuel, contained in the hottest part, in the centre, constantly producing energy. What happens when all the hydrogen in the central core has been turned into helium?

Old Age Stage The death of a star can happen in different ways, depending on the mass. For stars that have a mass of the order of 10 $M\odot$, when there is no fuel they lose the energy source that has kept the temperature and pressure high. The star begins to contract and the last supplies of hydrogen will keep producing energy, but the pressure is no longer sufficient to counterbalance the compression. This process continues until it reaches about 100 million degrees, enough to merge the helium nuclei which will give rise to heavier elements such as carbon and oxygen. While stars have a lower mass than 10 $M\odot$, the gravitational forces are weak and the gases pressure allows the expansion in space. At this stage the formation of the planetary nebula takes place. Some studies of the chemical composition of nebulae affirm that in many of them the main elements are helium, carbon, oxygen, nitrogen, and others [4, 5]. Now, the gases formed from the death of the stars will fill the space between them, forming what is called Interstellar Medium (ISM).

Figure 1.1: The Orion nebula is a stellar nursery – Credit: NASA/Hubble.

This section was written based on the following treaties: "L'universo alle soglie del 2000", Margherita Hack, Rizzoli [6] and "From Dying Stars to the Birth of Life - The new science of astrobiology and the search for life in the universe", Jerry L. Cranford, Nottingham University Press [7].

1.3 Interstellar Medium (ISM)

The Interstellar Medium (ISM) contains dust grains composed of C, Si, Fe, Mg, O with an abundance of 1% in mass but also contains in traces Polycyclic Aromatic Hydrocarbons (PAHs) and other bearing gaseous molecules. This composition is due to stellar explosions during the last stage of life, which involve the release of heavy elements to areas between stars. The dust is well-mixed with the gases that "fill" the ISM, they are mainly composed of hydrogen and helium in different states and heavy elements. The abundances of elements as number ratios relative to hydrogen in the ISM of our Galaxy, according to C. W. Allen [8], are 0.1 of helium, 6×10^{-4} of oxygen, 4×10^{-4} of carbon and 9×10^{-5} of nitrogen. Interstellar matter is not uniformly distributed over the galaxy, but is rather concentrated in clumps and filaments of various scales, called *clouds*. Figure 1.2 shows some representative classes of clouds based on temperature and density. The straight line represents the isobar zone, where certain clouds are in pressure equilibrium.

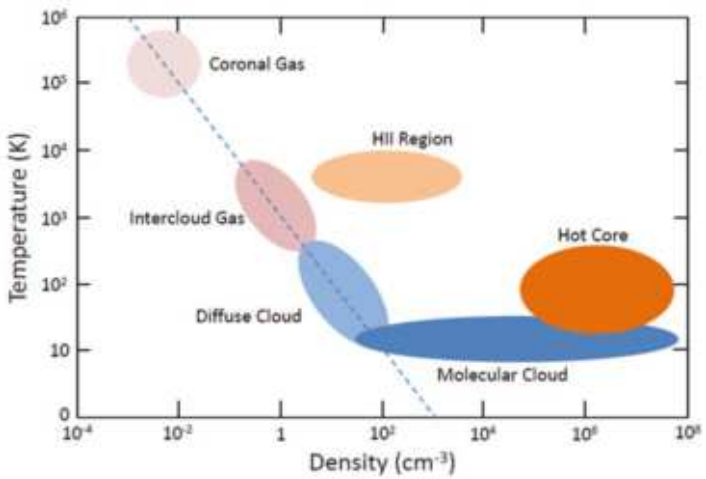


Figure 1.2: Temperature-density diagram of interstellar clouds [9].

According to Smith 2012 [10], it is possible to distinguish the regions of the ISM according to the values of temperature and particles density (n/cm^3).

The atomic phase contains neutral hydrogen atoms and can, in turn, be divided into cold and warm phase. The cold one has a temperature of 40-100 K and a density of a few tens of atoms over cm^3. The warm phase has a temperature of several thousand degrees, but with a lower density ($\sim$0.6 n/cm^3). There is also a warm ionized region at higher temperature and less dense than the neutral one, which contains ionic atomic hydrogen (H_{II}). It is estimated that both the warm regions contain as much matter as the cold one in our entire Galaxy. Atomic hydrogen and molecular hydrogen coexist in diffuse clouds, where the temperature decrease and the pressure increase, so these clouds become opaque to UV radiation. In regions protected by radiation from supernovae explosions and other sources, molecules are more stable and less reactive. Therefore, most astronomical observations focused on these areas to detect compounds of particular chemical complexity. Tab.1.1 shows the average structure of ISM.

Table 1.1 Summary of interstellar phases [11].

Phase	Temperature (K)	Density (n/cm^3)	Volume (%)	Mass (%)
Coronal gas or Hot Ionized Medium (**HIM**)	$10^{5.5}$-10^7	4×10^{-3}	50	<1
H_{II}	10^4	10^2-10^4	0.01	<0.01
Warm ionized medium	$\sim10^4$	0.3	10	23
Warm neutral medium	$\sim5\times10^3$	0.6	40	36
Cold neutral medium	40-100	$\sim$30	1	24
Diffuse molecular gas	30-100	30-300	0.1	0.17
Dense molecular cores	6-50	10^3-10^6	0.01	<0.05

Low-density clouds can be observed with optical instruments such as the Hubble Space Telescope, thanks to interactions of the matter with UV rays. The colder clouds, instead are completely blurred and can only be observed with radio telescopes such as Very Large Telescope (VLT) in U.S.A. and Australia.

Molecules such as hydrogen cyanide (HCN) [12, 13] and acetylene (H_2C_2) [14, 15] were detected in the 70's and 80's, but in recent years, also carbon based complex molecules such as propargylimine (C_3H_3N) [16] and benzonitrile (c-C_6H_5CN) [17] were discovered. That makes it possible to hypothesize the different mechanisms that led to these molecules.

1.4 Acetylene and hydrogen cyanide clusters

Since the discovery in meteorites precipitated on earth of biologically relevant molecules, such as vitamins B3, sugars, amino acids and others, it has begun to hypothesize the formation of the mentioned complexes in ionizing environments such as the ISM. Organic molecules containing nitrogen atoms have been identified in the gas phase ISM using radio astronomy, but detection of specific PAHs using radio astronomy is quite challenging. The computational study carried out by Tamar Stein *et al.* [18] showed that chemical bonds between hydrogen cyanide and acetylene molecule can be formed subsequent to ionization of neutral Van der Waals clusters. Ionization processes are often achieved in dense clouds by UV rays or in Titan's upper atmosphere where neutral and ionic species are observed by the CASSINI mission [19]. Neutral Van der Waals clusters of pure hydrogen cyanide or pure acetylene and the mixed ones may be the starting structures for more complex molecules, *e.g.* nucleic bases fragments.

1.5 Aim of this thesis

This thesis work aims to study the nature of interactions of hydrogen cyanide and acetylene clusters, through *ab initio*, DFT and semi-empirical methods. The fundamental step is the creation of a benchmark with the most common computational methods for the optimization of geometry and the energy of dimers. Simultaneously, the most accurate parameters of geometry and energy were calculated using the composite extrapolation procedure, to compensate the complete basis set limit error. The combination of method and basis set which give us the best result in according with the extrapolated value, it will be chosen to proceed with the second step of this work. In the second step, we begin the calculation and analysis of larger and more complex clusters, as trimers, tetramers and so on, using the most accurate

and fast method/basis set. This phase aimed to optimize ever larger structures as long as the resources at our disposal allow us. The third step aims to analyze the interactions nature that characterizes the clusters, with particular attention to most stable structures and Van der Waals forces.

Chapter 2

Theoretical Background

2.1 Wavefunction Theory (WFT)

Ab initio methods are useful tools in the hands of theoretical chemists to model molecular systems of various kinds. They are all based on the resolution of the electronic Schrödinger equation, which allows for the calculation of properties such as electron density, energy, geometry, etc. Using the Schrödinger equation in its most correct form, without approximation, we come across a highly complex problem. Fortunately, a less complex equation and an acceptable result can be obtained by several approximations. In particular, the *Hartree-Fock* (HF) method captures about 99% of the energy, while the remaining 1% is integrated with average values. Unfortunately, the latter fraction of energy is the most important from a chemical point of view. The post-HF methods developed try to best calculate the missing electron correlation energy.

2.1.1 Hartree-Fock Theory

Therefore, quantum mechanics is the best tool for geometric optimization and prediction of molecular properties. The Schrödinger equation provides a complete (or almost) mathematical description of a generic molecular system. The Hamiltonian operator describes the potential and the kinetic energy terms, while the ψ wavefunction represents everything we know about the system under analysis. The Hamiltonian ($\mathscr{H}$) is a mathematical operator which, when applied to the wavefunction, returns a value corresponding to the total energy of the system. The five terms that make up this operator refer to the kinetic energy of electrons and atomic nuclei and the potential energy given by the interaction between particles. Therefore, the time-independent, non-relativistic Hamiltonian of an isolated system can be expressed mathematically as follows:

$$\mathscr{H}_{tot} = \widehat{\mathbf{T}}_N + \widehat{\mathbf{T}}_e + \widehat{\mathbf{V}}_{Ne} + \widehat{\mathbf{V}}_{ee} + \widehat{\mathbf{V}}_{NN} = \tag{2.1}$$

$$= -\sum_i \frac{\hbar^2}{2m_e}\nabla_i^2 - \sum_k \frac{\hbar^2}{2m_k}\nabla_k^2 - \sum_i \sum_k \frac{e^2 Z_k}{r_{ik}} + \sum_i \sum_{j>i} \frac{e^2}{r_{ij}} + \sum_k \sum_{l>k} \frac{e^2 Z_k Z_l}{r_{kl}} \tag{2.2}$$

The first two terms refer to the kinetic energy of electrons and nuclei, defined by quantum mechanics as $\frac{p_i^2}{2m_i}$, with p_i as the operator of the momentum, $-i\hbar\nabla(i)$. While the third term represents the potential energy of attraction between electrons and nuclei due to Coulomb's forces. These contributions are negative because they lower the energy of the molecule. The remaining contributions include potential energies due to electrostatic repulsions, so they will raise the energy of the system. Considering the simplest case, without relativistic effects, electric fields and considering the time independence of the Hamiltonian, the Schrödinger equation remains unsolvable accurately. The cause arises from the correlation between the motions of particles, *i.e.* the non-independence. In order to dissolve this node, it is necessary to resort to an approximation based on the fact that electrons move much faster than atomic nuclei. The Born-Oppenheimer approximation of stationary nuclei considers electronic motion in a system of frozen nuclei. What we get is the calculation of an electronic function with a simplified Hamiltonian. In particular, the term concerning the kinetic energy of the nuclei is simplified, as they are considered fixed. The term nucleus-nucleus potential energy becomes

a constant additive to electronic energy. Whereas the term related to the nucleus-electron attraction becomes easily calculable considering now the independence between their two motions.

$$\mathscr{H}_{tot} = \widehat{\mathbf{T}}_e + \widehat{\mathbf{V}}_N\widehat{\mathbf{V}}_e + \widehat{\mathbf{V}}_{ee} + \widehat{\mathbf{V}}_{NN} = \tag{2.3}$$

$$= \mathscr{H}_{el} + \widehat{\mathbf{V}}_N \tag{2.4}$$

The Born-Oppenheimer approximation simplifies the Schrödinger equation with an irrelevant error since the simplified terms within the Hamiltonian are proportional to the reciprocal of the nuclei's mass. So the maximum error in energy calculation that can be obtained is given by hydrogen (proton) and is of the order of 10^{-4} Hartree. Another possible approximation to simplify the $\mathscr{H}$ operator is the approximation of independent electrons. The term related to interelectronic repulsion vanishes. In this way, it is possible to neglect the correlation of electronic motions. The total electronic Hamiltonian can be described as the sum of n monoelectronic Hamiltonians, h_i:

$$\mathscr{H}_{el} = h_1 + h_2 + \cdots + h_n \qquad h_i = -\frac{\hbar^2}{2m_e}\nabla_i^2 - \sum_k \frac{e^2 Z_k}{r_{ik}} \tag{2.5, 2.6}$$

$$\mathscr{H}_{el}\psi_{el} = \boldsymbol{E}_{el}\psi_{el} \tag{2.7}$$

In this case, the approximation of the independent electrons has led to a simplification of the Schrödinger equation which makes unacceptable errors. Without this simplification, however, the equation cannot be solved. The solution to this problem is to add corrective terms to the monoelectronic Hamiltonians by creating what are called Fock operators.

$$f_i = -\frac{\hbar^2}{2m_e}\nabla_i^2 - \sum_k \frac{e^2 Z_k}{r_{ik}} + \sum_{j\neq i}\left[\boldsymbol{J}(|\psi_j|^2) - \boldsymbol{K}(\psi_j, s_j)\right] \tag{2.8}$$

The implementation of f_i operators is the basis of the *Hartree-Fock* computational method. Fock operators consider two key energy contributions. **J**, the Coulombian operator, is the sum of the average repulsive potentials exerted on the individual electron by the others. **K**, the exchange operator, is the term that considers the exchange energy of electrons. The use of the Coulombian operator is also an important approximation since the interelectronic correlations are neglected. This explains the low accuracy of the HF method.

A Fock operator can be associated with the corresponding monoelectronic wavefunction ψ_i and the corresponding energy eigenvalue ε_i. ψ_i and ε_i represent the molecular orbitals and the orbital energies respectively.

$$f_i\psi_i = \varepsilon_i\psi_i \tag{2.9}$$

The Fock operator f_i depends on all the occupied Molecular Orbitals (MOs), due to the Coulomb and exchange operators. Thus, a specific orbital can only be determined if all the other occupied orbitals are known, introducing us to the Self-Consistent Field (SCF) Theory. The SCF procedure introduces arbitrary initial molecular orbitals $\psi_i^{(0)}$ (guess function), from which the initial Fock operators $f_i^{(0)}$ are derived. These operators are used to solve the Schrödinger equation to obtain new molecular orbitals $\psi_i^{(1)}$ with the respective energies $\varepsilon_i^{(1)}$ and so on. The iterative process is repeated until the energy variation between one cycle and the next are minimal. When the calculation converges it will be possible to calculate the energy corresponding to the most stable geometry.

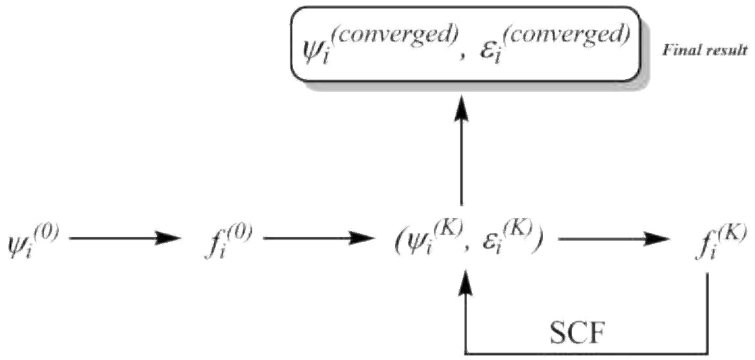


Figure 2.1: Self-Consistent Field procedure.

Only for the hydrogen atom (or He^+), there are orbitals exact eigenfunctions of the complete electronic Hamiltonian. As long as we consider molecules near their equilibrium geometry, the HF method often provides a good starting point for more elaborate theoretical methods such as many-body perturbation theory. However, the approximation involved in the derivation of HF theory (many-electron wavefunction constructed of one-electron wavefunctions) has a serious consequence. The treatment of electron-electron repulsion in an average way leads to a loss of electron correlation data.

2.1.2 Basis set

The set of mathematical functions by which a wavefunction is constructed is called *basis set*. The mathematical expressions used to describe orbital functions are mainly of two types. *Slater Type Orbitals* (STOs) and *Gaussian Type Orbitals* (GTOs):

$$\phi_{abc}^{STO}(x,y,z) = Nx^a y^b z^c e^{-\zeta r} \tag{2.10}$$

$$\phi_{abc}^{GTO}(x,y,z) = Nx^a y^b z^c e^{-\zeta r^2} \tag{2.11}$$

N is a normalization constant, a, b and c are coefficients that controls the angular momentum. ζ indicates the width of the orbital, where small ζ value gives a diffuse function. The STOs are proportional to e^{-r}, where r is the distance from the nucleus. Although these functions are more accurate, they make computational calculation time-consuming and resource-intensive. A solution proposed in 1950 by Boys was to use a linear combination of GTOs to simulate an STO function. This procedure is called Contracted GTOs (CGTOs), often written as STO-nG.

$$\phi_{abc}^{CGTO}(x,y,z) = N\sum_{i=1}^{n} c_i x^a y^b z^c e^{-\zeta_i r^2} \tag{2.12}$$

Hence, have a *double-*ζ basis sets means that there are two functions for each atomic orbital. Among the most common basis sets there are split valence basis sets, which usage *single-*ζ functions for *core* atomic orbitals and *n-*ζ functions for *valence* orbitals. For example, the 3-21G basis set uses three *single-*ζ GTOs functions for each core atomic orbital, meanwhile for each valence atomic orbital uses two *double-*ζ plus one GTOs functions.

$$\psi_i = \sum_i a_i \mathbf{\Phi}_i = \sum_i a_i \sum_g c_g (GTO)_g \tag{2.13}$$

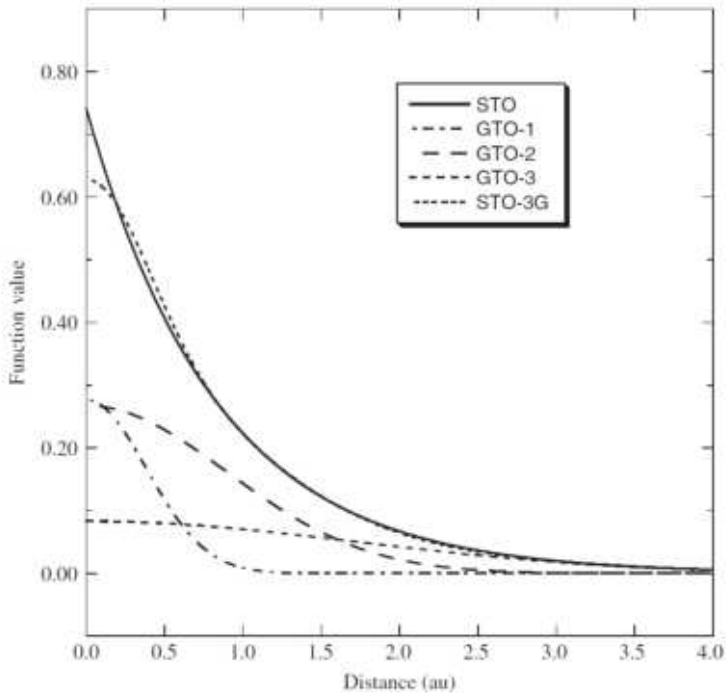


Figure 2.2: A 1*s*-STO modeled by a linear combination of three GTOs, *i.e.* STO-3G basis set.

During the SCF process, the c_g coefficients remain fixed, while a_i coefficients are optimized for each cycle.

In a molecular context, atomic orbitals fail to fully describe the chemical bond. So often polarized functions are needed to best describe the internuclear interaction. The resulting molecular orbitals will be localized more accurately by adding functions of a higher quantum angular number l.

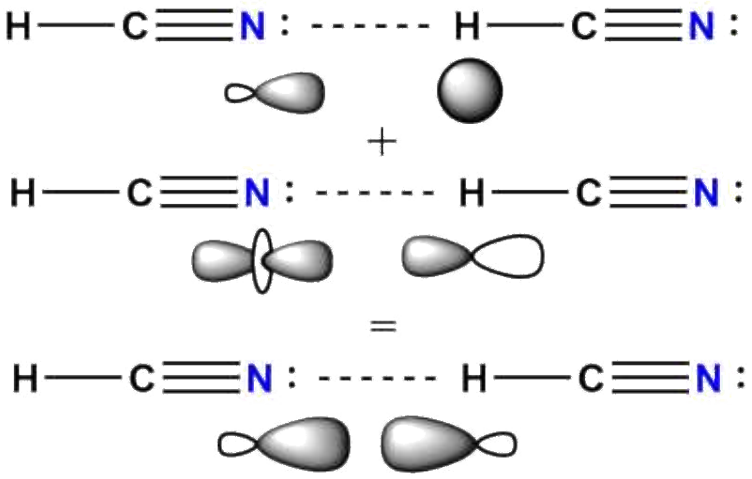


Figure 2.3: Example of applications of polarizing functions d_{z^2} and p_z to improve the representation of interaction between the nitrogen and hydrogen orbitals.

Generally, the polarizing functions d are used for elements of the second period, while the p functions are useful for hydrogen. The application of p GTOs functions in the hydrogen atom can lead to a better description of non-covalent interactions. In Pople notation, polarizing functions are indicated by the *star* symbol (*), *e.g.* the basis set **6-31G***, adds a set of d functions to the heavier atoms only (2^{nd} period onwards). While **6-31G**** basis set adds p

GTOs functions to light atoms and *d* functions to heavy atoms. As for Dunning's Correlation-Consistent basis sets, polarizing functions are indicated within the acronym **cc-pV*n*Z**, by the letter *n*, *e.g.* **cc-pVDZ** (*double-ζ*) inserts a set of *p* functions for light atoms and *d* functions for heavy ones.

In order to describe extended molecular orbital, diffuse functions are used. Some properties, such as dispersion forces, act at points in space far from the atomic nucleus. The diffuse functions have a small ζ, see Eq. 2.12, which means they cancel out less quickly in function of the distance from the atom. According to Pople's notation, the addition of diffuse functions such as *s* and *p* for heavy atoms is indicated with a "+" symbol, *e.g.* **6-31+G**. While if there are diffuse functions of the type *s* for the hydrogen atom it will also write **6-31++G**. In Dunning notation, the presence of diffuse functions is indicated by the prefix *augmented*, *e.g.* **aug-cc-pVTZ** which has diffuse *f, d, p* and *s* functions on heavy atoms and diffuse *d, p* and *s* functions on hydrogen atoms.

Diffuse functions are not always needed and their use can lead to increased computation times and problems of convergence. Truhlar *et al.* [20], presented a perspective on the usage of diffuse functions for electronic structure calculations. Therefore, considering the basis set **aug-cc-pV*n*Z** with the most diffuse functions, a set of basis sets has been devised in which the last set of diffused functions is removed. In particular, the *aug* prefix is interpreted as *august* and as the months go back, more and more diffuse functions are removed, thus lightening the set of functions to be used. In Tab. 2.1 are shown the calendar series for basis set with *quadruple-*, *triple-* and *double-ζ*.

Table 2.1 Calendar basis sets series [21].

Month	Basis set	Diffuse functions for heavy atoms	Diffuse functions for light atoms
August	**aug-cc-pVQZ**	*s p d f g*	*s p d f*
July	**jul-cc-pVQZ**	*s p d f g*	*None*
June	**jun-cc-pVQZ**	*s p d f*	*None*
May	**may-cc-pVQZ**	*s p d*	*None*
April	**apr-cc-pVQZ**	*s p*	*None*
August	**aug-cc-pVTZ**	*s p d f*	*s p d*
July	**jul-cc-pVTZ**	*s p d f*	*None*
June	**jun-cc-pVTZ**	*s p d*	*None*
May	**may-cc-pVTZ**	*s p*	*None*
August	**aug-cc-pVDZ**	*s p d*	*s p*
July	**jul-cc-pVDZ**	*s p d*	*None*
June	**jun-cc-pVDZ**	*s p*	*None*

jun-cc-pVDZ, **may-cc-pVTZ** and **apr-cc-pVQZ** basis sets are also called minimal augmented (*maug*-cc-pV*n*Z).

2.1.3 Hartree-Fock Limit

Lack of using a complete basis set in practice, due to the increase in computational effort, limits calculations to a small number of basis functions. For most systems, this leads to large errors in energy. This is especially upsetting when calculating weak forces, *e.g.* Van der Waals complexes. The system converges as the complexity of the basis set increases to a limit value, called HF limit. The limit energy value is not the true value, because of the approximations implemented by the HF method, such as the electron correlation energy. A solution often used to achieve the most accurate value is the use of composite extrapolation methods. In principle, many systems can be calculated with high accuracy by using a CCSD(T) method and performing a series of calculations with larger basis series to extrapolate to the basis set limit. In practice, this method is computationally expensive. Thus many approximate procedures have been developed for estimating the infinite basis set value. These models consist in the fact that different proprieties converge with different rates as the level of sophistication increases and that effects from the extending basis set to a certain degree are additive. The extrapolation procedure performed in this thesis work will be illustrated in chapter 3.

2.2 Multi-configuration Methods

The *Hartree-Fock* method generates solutions to the Schrödinger equation where the real electron-electron interaction is replaced by an average value. Electronic correlation energy refers to the difference in energy between HF and the lowest achievable value for a given basis set. The approximation of the correlation energy by the HF method leads to an underestimation of the electron repulsion. For example, in the case of a geometric optimization calculation, the bond lengths will be slightly shorter. As the HF solution usually gives ∼99% of the correct value, electron correlation methods normally use the HF wave function as a starting point to describe the occupancy of the molecular orbitals and the spin state of the system. The basic idea is to create a wave function that is the combination of several different electronic configurations with the same multiplicity. This would lead to a better description of the system under consideration by offsetting some of the correlation energy.

$$\Psi^{MC} = a_0\Phi_{HF} + \sum_{i=1} a_i\Phi_i \tag{2.14}$$

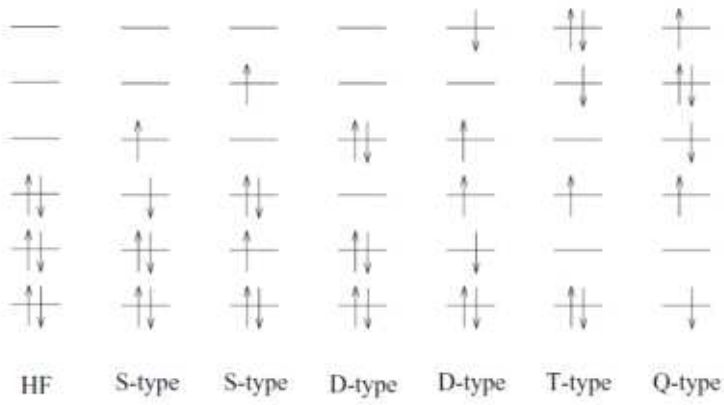


Figure 2.4: Different Slater Determinants of exited states - Singles (S), Doubles (D), Triples (T) and Quadruples (Q).

The most common multi-configuration methods include those based on the Møller-Plesset (MP) and Coupled Cluster (CC) theory. Both are based on the combination of different configurations of the system under consideration, to improve the description of the molecular electronic structure, but with a different mathematical approach. The MP perturbation theory adds to the Hamiltonian operator a perturbative factor that becomes the exact electron correlation potential. While the CC theory constructs multi-electron wavefunctions using the exponential cluster operator to account for electron correlation. This thesis is not intended to

deepen the mathematical treatment of these theories, but only to demonstrate their usefulness in the field of computational chemistry. The number of combinations of electronic configurations for a molecule is extremely high. Therefore, these methods are limited to combining only some of them. For example, methods based on MP theory can consider, in addition to the ground state, single and double electronic excitations, MP2. Or, MP3 if you intend to consider also the triple excitations. Similarly, CC methods can be called CCSD if they consider the combination of the basic electronic state with that corresponding to the single and double electronic excitation. The so-called "gold-standard" method CCSD(T) consider both the single and double electronic excitations and in addition, the triple excitations are calculated via perturbation theory, not to be confused with the CCSDT method, which calculates the triple excitations through the usual CC theory. Multi-configuration methods are very efficient and can correct much of the correlation energy. However, given the greater number of electronic configurations to consider compared to other methods, they are computationally expensive. To make them less demanding, a general approximation is employed which considers only valence electrons in the energy calculation, neglecting those of the nucleus. The *frozen core (fc)* approximation reduces the number of electronic configurations to be calculated, based on the lower impact that core electrons have on the system's energy. A multi-configurational method, having to handle multiple wave functions at the same time, is more sensitive to the characteristics of the basis set than the HF method. This led to *correlation consistent* basis sets such as those developed by Dunning [22], *e.g.* cc-pVTZ.

2.3 Density Function Theory (DFT)

Density functional theory is conceptually and computationally very similar to Hartree–Fock theory, but provides much better results and has consequently become a very popular method. Hartree-Fock methods exactly treat exchange-correlation but have difficulties recovering dynamic electron correlation while DFT has an exact form for dynamic electron correlation but since DFT is not quantum mechanical, it must approximate exchange-correlation. Several hybrid DFT functional were chosen based on their ability to describe non-covalent complex proprieties. The fact that they are hybrids allows considering different contributions to both exchange energy (E_x) and correlation energy (E_c).

2.3.1 B3LYP

This functional uses a different mixing scheme involving three parameters for the exchange energy. Specifically, the local spin density approximation (LSDA) and the B88 functionals, are added to the exact HF exchange energy. The correlation energy term is computed using the gradient-corrected correlation functional named Lee-Yang-Parr (LYP) and a small part of Vosko-Wilk-Nusair (VWN) functional [23].

$$\boldsymbol{E}_{xc}^{B3LYP} = a_0\boldsymbol{E}_x^{HF} + (1-a_0)\boldsymbol{E}_x^{LSDA} + a_x\boldsymbol{\Delta E}_x^{B88} + a_c\boldsymbol{E}_c^{LYP} + (1-a_c)\boldsymbol{E}_c^{VWN} \tag{2.15}$$

2.3.2 B2PLYP

Double-hybrid density functionals are based on a mixing of standard generalized gradient approximations (GGAs) for exchange and correlation with Hartree-Fock (HF) exchange and a perturbative second-order correlation part (PT2) [24].

$$\boldsymbol{E}_{xc}^{B2PLYP} = a_x\boldsymbol{E}_x^{HF} + (1-a_x)\boldsymbol{E}_x^{GGA} + (1-a_c)\boldsymbol{E}_c^{GGA} + a_c\boldsymbol{E}_c^{PT2} \tag{2.16}$$

The B2PLYP seems to work well for more strongly bound systems with shorter intermolecular distances, but for very weakly bound the GGA parts are mainly repulsive and the perturbation theory term, an attractive contribution, is too small.

2.3.3 Minnesota functional M06-2X

The Minnesota family are a set of four highly parameterized and exchange-correlation energy approximated functionals. They are based on the *meta*-GGA, where the exchange-correlation

potentials depend on the second derivative of electron density. Therefore, this procedure implies more demanding time resources. The M06 functional, developed by Truhlar D. G. and Zhao Y. in 2006 [25], uses a parameterized exchange-correlation potential plus a portion of HF exchange potential term. The M06-2X functional consider double the HF exchange contribution than the common M06, 27% of E_x^{HF} for M06 and 54% of E_x^{HF} for M06-2X.

2.4 Semi-empirical Methods

The main problem of HF and multiconfigurational methods is the high computational resources demanding, which rapidly increase with the system complexity. Semi-empirical methods allow keeping a quantum mechanical approach, but with less time and hardware demands. Several semi-empirical methods are considered an approximation of HF theory, in which some parameters are replaced with experimental data with the purpose of drastically reduce computational times. Due to the development of DFT methods, semi-empirical methods have also evolved as approximations of density function theory, which has given rise to a vast set of models called tight-binding DFT (DFTB). Grimme *et al.* [26, 27] present an extended semi-empirical model, designed for the fast calculation of structures and non-covalent interaction energies for molecular systems with approximately 1000 atoms. The geometries-frequencies-non-covalent interactions tight-binding model, GFN2-xTB, consists of a minimal valence basis set atoms centred, STO-nG. Polarization functions for heavy atoms are also employed. GFN2-xTB represents the first broadly parametrized tight-binding method which includes electrostatic and exchange-correlation Hamiltonian terms. It's parametrized for all elements up to radon (Z=86) and the parameters are fitted to yield reasonable structures, vibrational frequencies and non-covalent interactions.

2.5 Composite Methods

Grimme *et al.* [28–30] developed many DFT functionals and an HF method where three computational corrections for structures, vibrational frequencies and non-covalent interaction energies were applied. Taking the HF-3c method as an example, three terms are added to correct the Hartree-Fock energy in order to include the London dispersion forces, account the BSSE and correct for basis set overestimation of bond lengths. Equation 2.17 shows the three corrective terms, the first term, $\boldsymbol{E}_{disp}^{D3BJ}$, correct the dispersion energy by the D3BJ scheme [31]. The second term, $\boldsymbol{E}_{BSSE}^{gCP}$, estimated the Basis Set Superposition Error (BSSE) through the *geometric counterpoise (gCP)* correction [32]. Finally, the last term, $\boldsymbol{E}_{SRB}$ is a short-ranged correction for the covalent bond lengths overestimation.

$$\boldsymbol{E}_{tot}^{HF-3c} = \boldsymbol{E}_{tot}^{HF} + \boldsymbol{E}_{disp}^{D3BJ} + \boldsymbol{E}_{BSSE}^{gCP} + \boldsymbol{E}_{SRB} \tag{2.17}$$

2.6 Functional Comparison

Calculation times in computational chemistry are influenced by the number of atoms to be optimized, the number of functions of the basis set and the level of theory used. Semi-empirical methods are approximated using experimental parameters but maintain a quantomechanical approach. This reduces the calculation time, while ab initio methods are more expensive since all parameters need to be calculated from the beginning. In particular, post-HF methods, which rely on multiple electronic configurations, require many steps and therefore more time and computational resources. Considering the same level of theory used in the calculation, what makes the difference in timing is the number of basis set functions used. An HF method coupled with a *triple-ζ* basis set will be more expensive than the same method using a *double-ζ* basis set. Consequently, the increase of atomic nuclei in the molecular system will also increase the functions of the basis set to be used and thus the calculation time. The figures below illustrate the time comparison between different functionals and basis sets. The calculations performed are single-point energy on acetylene clusters on the work station *Jupiter*, where the specifications are described in further chapter 3. In Fig. 2.5 is shown the time difference between using an aug-cc-pVDZ basis set and an aug-cc-pVTZ with different methods. For

all the methods tested, the *triple-ζ* basis set require more time to complete the calculations, due to the increase of functions number. The aug-cc-pVDZ basis set has 9 functions for each hydrogen atom and 23 functions for each carbon atom, while the aug-cc-pVTZ basis set has 23 functions for each hydrogen atom and 46 functions for each carbon atom. The same plot but for the acetylene isolated molecule is reported in Fig. 2.6 to illustrate the time difference between monomer and dimer calculation.

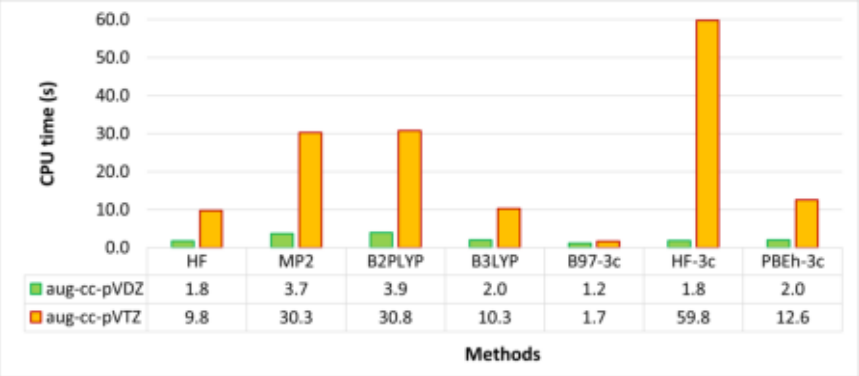


Figure 2.5: Time comparison of single-point energy calculation on acetylene dimer with different methods and both *double-* and *triple-ζ* basis set.

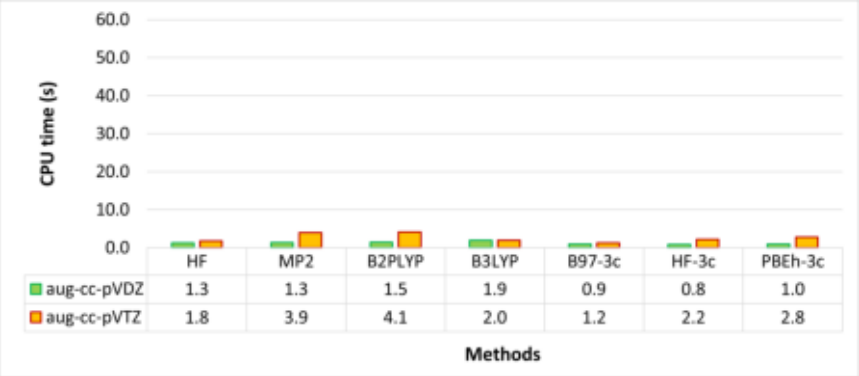
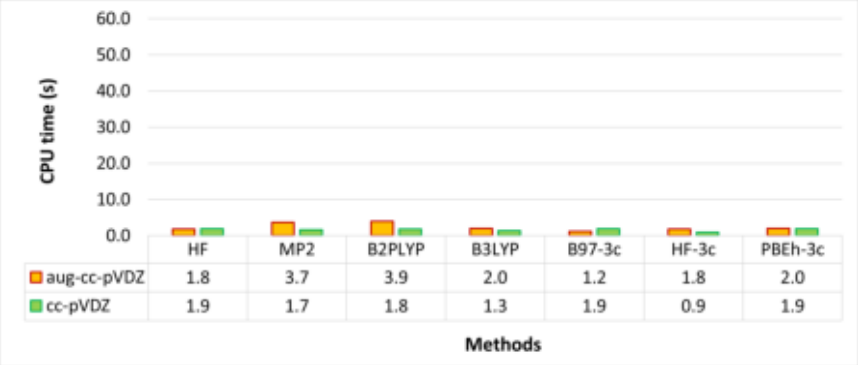
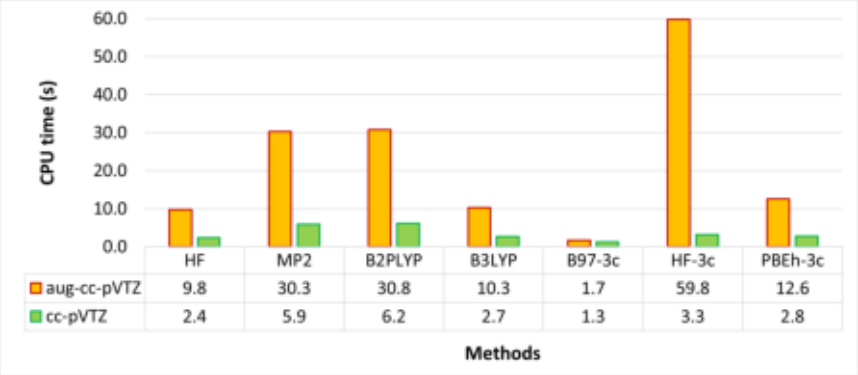


Figure 2.6: Time comparison of single-point energy calculation on acetylene molecule with different methods and both *double-* and *triple-ζ* basis set.

The figures 2.7 show the difference of single-point energy calculation time between a basis set with diffuse functions and the same basis set without them. Using diffuse functions in the basis set complicates calculations and adds more functions, so it is plausible to expect an increase in computation times due to their implementation.



(a)



(b)

Figure 2.7: Time comparison of single-point energy calculation on acetylene dimer with different methods, first with diffuse functions (**a**) then without diffuse functions (**b**).

Fig. 2.8 shows how increasing the number of acetylene units increases the time of the calculation carried out with the xTB method. Due to the high parameterization of this method, the calculation times and the differences between clusters are not appreciable. However, this trend is reflected in any kind of functional.

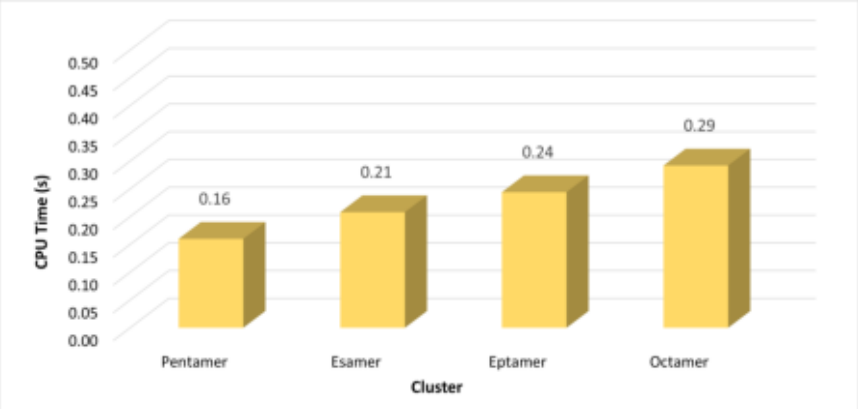


Figure 2.8: Time comparison of single-point energy calculation with xTB method of different acetylene clusters.

2.7 Non-covalent interactions

Ionic, covalent and metallic bonds are atomic strong forces which arise from Coulombic forces or the overlap of electron wave functions. Interactions of dipolar nature such as the strong hydrogen bond or the weak Van der Waals (VdW) forces, arise from dipole-dipole interactions. Hydrogen bond is a strong type of dipole-dipole interaction, due to the high polarization of electron density in the molecule, *e.g.* hydrogen atom bonded to oxygen, form a strong dipole due to oxygen electronegativity. Van der Waals interactions are medium-ranged forces and the interaction length can be of few Angstroms or hundreds. Van der Waals interactions are significant in fluidic systems, but also for interactions between microscopic bodies. VdW forces can be divide into three groups:

- Dipole-dipole force: Molecules having permanent dipoles will interact by dipole-dipole interaction;
- Dipole-induced dipole force: The field of a permanent dipole induces a dipole in a non-polar molecule;
- Dispersion or London force: Due to charge fluctuations in a non-polar molecule. At certain moment, a dipole takes form and induces a dipole in another non-polar molecule. Thus, two non-polar molecules attract each other due to their temporary dipoles.

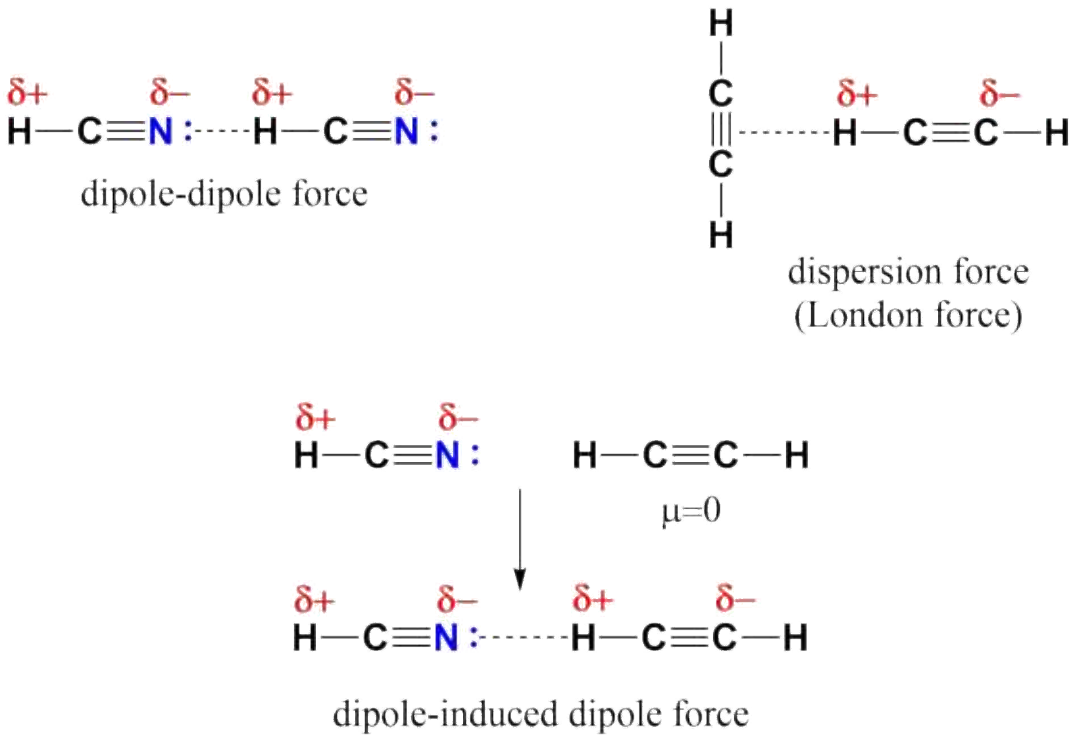


Figure 2.9: Examples of Van der Waals forces.

2.8 Quantum Theory of Atoms in Molecules (QTAIM)

The Quantum Theory of Atoms in Molecules, developed by Richard F. W. Bader and coworkers [33], allows to describe the molecular structure from a quantomechanical point of view. This theory is widely used for the study of covalent and non-covalent interactions, from the smallest molecules to DNA. The electron density $\rho(\mathbf{r}, \mathbf{X})$ in an observable quantity useful to study the nature of chemical bond. Atomic nuclei are called attractors of electron density, because of the positive charge, and provide the force field around each atom. What we expect is a higher electron density close to the nuclei.

$$\rho(\boldsymbol{r}, \boldsymbol{X}) = N_e \int \psi^*(\boldsymbol{r}, \boldsymbol{X}) \psi(\boldsymbol{r}, \boldsymbol{X}) d\tau \tag{2.18}$$

A *critical point* (CP) is a point in space at which the gradient of the electron density goes to zero.

$$\nabla\rho(\boldsymbol{r}) = \boldsymbol{i}\frac{d\rho}{dx} + \boldsymbol{j}\frac{d\rho}{dy} + \boldsymbol{k}\frac{d\rho}{dz} \rightarrow \begin{cases} = \vec{0} & \text{at CPs and at } \infty. \\ \neq \vec{0} & \text{at other points.} \end{cases} \tag{2.19}$$

To locate in the space which is the minimum, maximum and saddle points, it is necessary to consider the matrix of a CP located at $\boldsymbol{r}_c$. The matrix can be diagonalized, being real and symmetrical, thus yielding three eigenvalues.

The three eigenvalues λ_1, λ_2 and λ_3 are the curvatures of the electron density. By using these three λ, each critical points is characterised in terms of the *rank* (ω) and *signature* (σ). The first is the number of non-zero eigenvalues λ_i of a CP; only the critical points with $\omega = 3$ are chemically interesting. The signature σ is the algebraic sum of the signs of the curvatures, *i.e.* $\sigma = -3, \ldots, +3$.

$$\nabla^2\rho(\boldsymbol{r}) = \frac{\partial^2\rho(\boldsymbol{r})}{\partial x^2} + \frac{\partial^2\rho(\boldsymbol{r})}{\partial y^2} + \frac{\partial^2\rho(\boldsymbol{r})}{\partial z^2} \tag{2.20}$$

Table 2.2 Classification of critical points by (ω, σ) notation.

($\boldsymbol{\omega}$,$\boldsymbol{\sigma}$)	**Description**	**Critical point**
(3, −3)	Three negative curvatures: ρ is a local maximum	Nuclear Critical Point (NCP)
(3, −1)	Two negative curvatures: ρ is a maximum along one coordinate	Bond Critical Point (BCP)
(3, +1)	Two positive curvatures: ρ is a minimum along one coordinate	Ring Critical Point (RCP)
(3, +3)	Three positive curvatures: ρ is a local minimum	Cage Critical Point (CCP)

In Fig. 2.10 is reported a pentamer cluster optimized in this thesis work, which illustrates an example of different critical points.

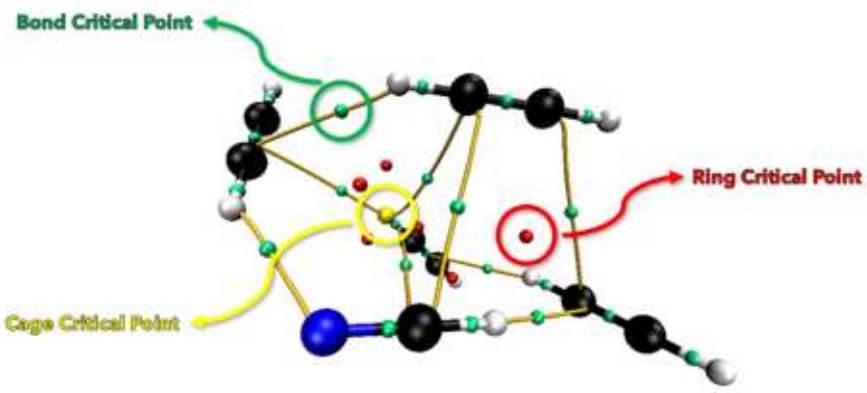


Figure 2.10: $(H_2C_2)_5{\cdot}(HCN)$ cluster which present BCP (green) along each bond, RCP (red) in each faces of the structure and a CCP (yellow) in the middle of the structure. Every atom nucleus present a NCP covered by the graphical representation.

Between one attractor and another, it is possible to distinguish a surface where the electron density gradient is cancelled, separating the space into regions controlled by the individual attractors. These regions can be described as in Fig. 2.11 by the red arrows, which represent the increase of electron density gradient.

(a) **(b)**

Figure 2.11: Vector map of electron density gradient – Hydrogen cyanide (**a**) and Acetylene (**b**).

The topology analysis considers many quantities for each bond critical point (BCP), called QTAIM descriptors. There is one BCP between each intramolecular atom and one for each intermolecular interaction. One descriptor is the electron density at the BCP which expresses the strength of chemical bond (bond order). The relation between the bond order (B.O.) and the electron density at the BCP (ρ_b) is given by:

$$\boldsymbol{B.O.} = e^{[A(\rho_b - B)]} \tag{2.21}$$

where A and B are two constants that depend on the nature of the bonded atoms. This expression is a simplified version and considers only two atoms involved in the interaction.

The Laplacian of electron density is another interesting parameter of BCP that can help us distinguish between attractive and repulsive interactions. As mentioned above, it consists of the sum of the three eigenvalues, at the nuclei, all of the eigenvalues are negative, while long the bond path, the λ_3 eigenvalue is positive, and the λ_1 and λ_2 eigenvalues represent the variation of electron density in the plane normal to λ_3. The eigenvalue λ_2 can be either positive and negative, depending on the interaction type. Hydrogen bond interactions are characterized by an accumulation of density perpendicular to the bond path, which means $\lambda_2 < 0$. Non-bonded interactions, such as steric repulsion, produce a density decrease around the bond, so $\lambda_2 > 0$. Interactions of Van der Waals type are characterized by a negligible,

negative value of λ_2. While the sign of λ_2 enables us to distinguish between the different interactions, the value of electron density gives us information on the interaction strength. Figure 2.12 summarizes the parameters for recognizing the NCI type.

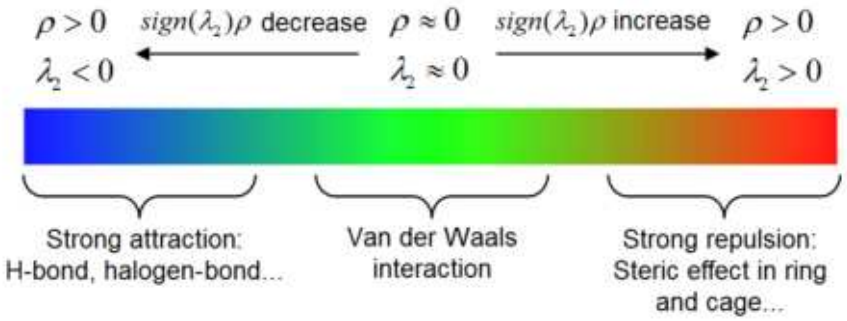
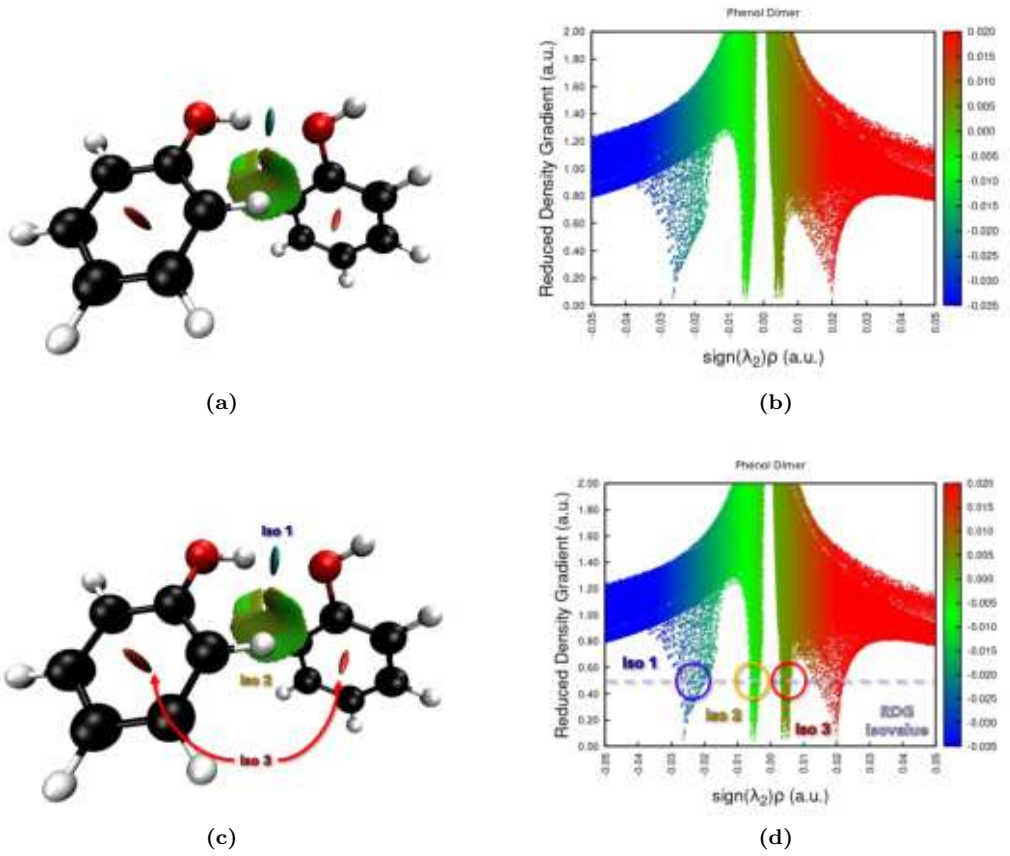


Figure 2.12: Colour scheme of NCI classification - The colours represent a scale of NCI strength. Credit: Multiwfn 3.8 Manual.

Non-covalent interactions can be visualized in a 3D plot of electron density times the sign of λ_2 versus the reduced density gradient (RDG). The colour scale is based on the value of sign$(\lambda_2)\rho$ and represent different interaction types, red for steric repulsion, green for Van der Waals and blue for interactions with more covalent character, such as hydrogen bond. RDG($\boldsymbol{\rho}$) parameter is a dimensionless quantity used to describe the deviation from a homogeneous electron distribution.

$$\boldsymbol{RDG}(\boldsymbol{\rho}) = \frac{1}{2(3\pi^2)^{\frac{1}{3}}} \frac{|\nabla\rho|}{\rho^{\frac{4}{3}}} \tag{2.22}$$

The behaviour of RDG($\boldsymbol{\rho}$) at low densities is dominated by ρ, thus RDG($\boldsymbol{\rho}$) tends to diverge except in the regions around the critical point, where the $\nabla\rho$ dominates and RDG($\boldsymbol{\rho}$) value goes to zero. The isosurface describes a region where there are points with the same value of reduced gradient, in our cases we choose an isosurface with RDG($\boldsymbol{\rho}$) = 0.5 a.u. The graphs are built around the non-covalent interaction critical points. The phenol dimer and his RDG plot are reported below as an example of the NCI index analysis functioning.

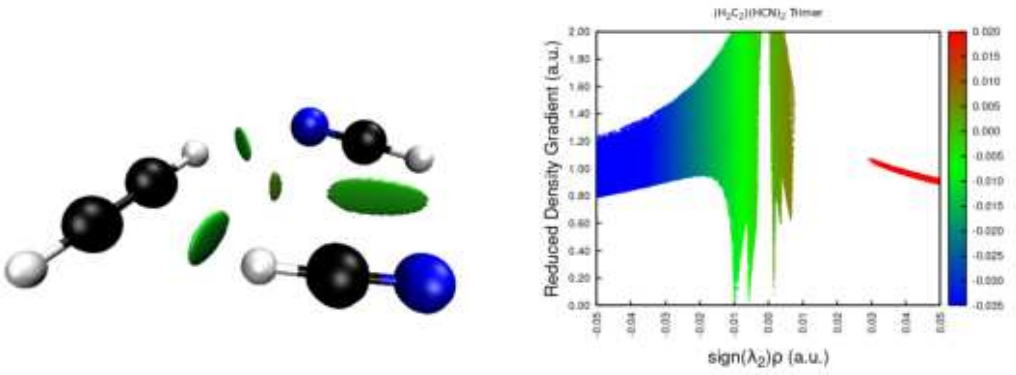


(a) (b)

(c) (d)

Figure 2.13: Isosurfaces (RDG($\boldsymbol{\rho}$) = 0.5) and Reduced Density Gradient vs sign(λ_2)·ρ plot of phenol dimer.

Fig. 2.13 shows many different types of interactions between the two phenol molecules. The points in the graph where the RDG($\boldsymbol{\rho}$) value goes to zero mean that there is a maximum or a minimum of electron density, and the value of the latter indicates the nature of the interaction. The blue peak in the range of -0.02 and -0.03 a.u. of electron density, suggests a strong bonding interaction corresponding to the hydrogen bond between the two hydroxyl groups, illustrated by the blue isosurface in Fig. 2.13a. The two signals near the zero indicate the presence of Van der Waals interactions, shown in Fig. 2.13a as the green surface between the two molecules, including both repulsive and attractive forces. Finally, the red peak at +0.02 a.u. in Fig. 2.13b displays the presence of steric repulsions, found at the phenol ring centre. The four peaks represent approximately the critical points positions in AIM theory.



(a) (b)

Figure 2.14: Isosurfaces (RDG($\boldsymbol{\rho}$) = 0.5) and Reduced Density Gradient vs sign(λ_2)·ρ plot of (H_2C_2)·$(HCN)_2$ trimer optimized at B97-3c/def2-TZVP level.

In Fig. 2.14 are illustrated the non-covalent interactions for one of the clusters optimized in this work, as a demonstration of the usefulness of this analysis for the study of our Van der Waals complexes. The structure in Fig. 2.14a shows three isosurfaces due to Van der Waals interactions and one ring type isosurface due to steric effect, as displayed also from the plot in Fig. 2.14b. From the RDG plot, we can assume that the steric repulsions are weak, thus the ring structure is quite open and flexible.

A descriptor related to the covalent nature of the interactions is the total electron energy density at the BCP. The Laplacian of electron density is related to the kinetic energy density and the potential energy density.

$$\left(\frac{\hbar^2}{4m}\right)\nabla^2\rho(\boldsymbol{r}) = 2\boldsymbol{G}(\boldsymbol{r}) + \boldsymbol{V}(\boldsymbol{r}) \tag{2.23}$$

We always have $\boldsymbol{G}(\boldsymbol{r}) > 0$ and $\boldsymbol{V}(\boldsymbol{r}) < 0$, so for $\nabla^2\rho_b < 0$ the interactions on the BCP are dominated by a local reduction of the potential energy. On the other hand, interactions for which $\nabla^2\rho_b > 0$ are dominated by a local excess in the Lagrangian kinetic energy ($\boldsymbol{G}(\boldsymbol{r})$). The total electron energy density ($\boldsymbol{H}_b$) is defined as the sum of the two terms, $\boldsymbol{H}_b$ is negative for interactions with significant covalent character.

$$\boldsymbol{H}_b = \boldsymbol{G}(\boldsymbol{r}) + \boldsymbol{V}(\boldsymbol{r}) \tag{2.24}$$

The ratio of absolute potential energy density $|V(\mathbf{r})|$ to Lagrangian kinetic energy density $G(\mathbf{r})$, at the NCI CP can be interpreted as a measure of covalent character. $\frac{|V(\mathbf{r})|}{G(\mathbf{r})} < 1$ corresponds to weak hydrogen bonds and large interaction distance.

Furthermore, applications of QTAIM descriptors and NCI index analysis will be explained down the road of this thesis.

This chapter was written based on the following treaties: "Approccio qualitativo alla Chimica computazionale" M. Bortoluzzi, ARACNE, 2009 [34], "Introduction to Computational Chemistry", F. Jensen, John Wiley & Sons, 2017 [35], "Essentials of Computational Chemistry: theories and models", C.J. Cramer, John Wiley & Sons, 2013 [36] and "The quantum theory of atoms in molecules: from solid state to DNA and drug design", A. Beck, John Wiley & Sons, 2007 [37].

Chapter 3

Methodology and Computational Resources

The first step of this work was to optimize four different dimers of hydrogen cyanide and acetylene, with various computational methods. The choice of the dimers was carried out due to the relevance of these dimers, as stated in the literature. For the $(HCN)_2$ dimer the linear structure was studied, where a hydrogen bond between the molecules take place. Many authors were able to isolate and stabilized that dimer and did measurements like vibrational frequencies [38, 39]. For the $(H_2C_2)_2$ dimer the T-Shaped structure was studied, as an example of the H-π interaction system. This and other structures were studied in literature by Karpfen [40], such as the cross-shaped and the slipped parallel. Karpfen demonstrated through his calculations that these last two structures are first-order saddle-points, while only the T-Shaped configuration is a minimum. The last two dimers studied in this work were the linear and the T-Shaped structures of the $HCN{\cdot}H_2C_2$ cluster. The former structure has a lone pair of electrons on the nitrogen atom that interacts with the electron-poor hydrogen of acetylene. The latter is stable due to the interaction between the electron-poor hydrogen of HCN and the π-bond of acetylene.

Both geometry and energy parameters were taken into consideration for these structures. Particularly important the bond length of Van der Waals interactions and the variations of the intramolecular bond length due to the non-covalent interaction. The binding energy (BE) of each structure was calculated considering the energy of the cluster and the energy of the isolated monomers.

$$\boldsymbol{B.E.} = \boldsymbol{E_{cluster}} - n\boldsymbol{E_{monomer}} \tag{3.1}$$

For a better description of a real system, the Zero Point Vibrational Energy (ZPVE) correction was used, adding the Zero Point Vibrational Energy term obtained from the harmonic calculation of the dimers.

$$\boldsymbol{ZPVBE} = (\boldsymbol{E_{cluster}} + \boldsymbol{ZPVE_{cluster}}) - n(\boldsymbol{E_{monomer}} + \boldsymbol{ZPVE_{monomer}}) \tag{3.2}$$

All calculations were performed using both GAUSSIAN 16 [41], revision C 0.1, and ORCA 4.2.1 [42] software. The harmonic force field was constructed by studying the stationary state points and the second derivatives for the identification of the minimum and saddle points. All geometric parameters, such as bond lengths and bond angles, of each molecule, were left free to change during each optimization.

The non-covalent interactions regions were studied with both Multiwfn - A Multifunctional Wavefunction Analyzer 3.8 [43] and NCIPLOT 4.0 [44–46] and graphically processed with VMD - Visual Molecular Dynamics [47].

3.1 Dimers's Benchmark

The benchmark was built from the most common DFT functionals, applying to each Grimme's correction for dispersion forces (D3BJ) [31]. For comparison, calculations without the D3BJ correction were also performed. In particular, the doubly hybrid B2PLYP and the hybrid functionals B3LYP. The calculations with the M06-2X Minnesota functional was also performed. Each functional was used with the triple-ζ calendar series basis set and a particular version of

may-cc-pVTZ, where a set of *d* functions for hydrogen is removed, *i.e.* may-cc-pVTZ(-dH). The double-ζ basis set SNSD from the work of Piccardo *et al.* [48] was also applied with B3LYP. Long-range corrected DFT functionals such as CAM-B3LYP and LC-ωPBE with the Jensen basis set *aug*-pc-3 were used, to obtain a better description of NCIs.

Grimme's recent work on "low-cost" methods pushed us to add HF-3c, B97-3c and PBEh-3c to the benchmark. These semi-empirical methods include the D3 correction and can be applied to hundreds of atoms, in particular the B97-3c. For these methods, the Ahlrichs type basis set def2-*n*ZVP were used, as recommended by Grimme [49]. An extended semiempirical tight-binding model (xTB), presented in chapter 2, was implemented in our benchmark. The xTB model was designed for the fast calculation of geometry and energy of structure and non-covalent interactions of roughly a thousand atoms. The calculations were carried out with the xTB software package [50].

In recent years, Martin and Kozuch [51, 52] developed a dispersion-corrected, spin-component-scaled, double hybrids (DSD) DFT functional. The combination between PBE DFT for the exchange energy and P86 for the correlation energy results as a very accurate and robust method, the DSD-PBEP86. Hence, we decided to implement it in our benchmark coupled with the def2-TZVP basis set. The revised version designed by Martin, namely the revDSD-PBEP86, was also used being considered very promising in terms of performances [53]. Both DSD-PBEP86 and revDSD-PBEP86, with dispersion forces correction, represent the top pier of Jacob's Ladder [54], where the last rung represent the most accurate hybrid DFT functionals for exchange and correlation energies.

Figure 3.1: Jacob's Ladder of DFT functionals [55].

The geometry optimization at coupled-cluster level including full account of single and double excitations, with perturbative estimate of triple excitations (CCSD(T)), in conjunction with a triple-ζ basis set within the frozen core approximation, was performed. The same method calculation but with an additional set of diffuse functions, *aug*-cc-pVTZ, was also implemented in our benchmark.

An extrapolation to the Complete Basis Set (CBS) was carried out, in such a way we could confront the benchmark results with parameters at the highest level of theory. The Composite Extrapolation Procedure is based on the "Cheap" Composite Approach to Non-Covalent Interactions, described extensively in the work of Barone *et al.* [56], here a brief overview of the "Cheap" extrapolation scheme is presented.

The extrapolation was build from the calculations of dimers geometry at HF level and cc-pV*n*Z basis set, with n = T, Q and 5. We used the 3-points extrapolation (Eq. 3.3) from Halkier *et al.*[57] to obtain the HF_{CBS} value of the parameter in analysis. Another brick to the extrapolation scheme was the fc-MP2 correlation energies term, computed employing the cc-pVTZ and cc-pVQZ basis sets, using the formula by Helgaker *et al.* [58] (Eq. 3.4). These terms were corrected by adding the diffuse function's contribution $\Delta\alpha$, calculated as

the difference between fc-MP2/aug-cc-pVTZ and fc-MP2/cc-pVTZ results. We also considered the core-valence contribution computed as the difference between ae-MP2/cc-pVTZ (*all electrons)* and fc-MP2/cc-pVTZ, for estimating the intershell correlation (core-valence) and the intrashell correlation (core-core). The last term added was the contribution of the perturbative treatments of triples at the CCSD(T)/cc-pVTZ level, denoted as ΔT. This one was calculated by the difference between the fc-CCSD(T)/cc-pVTZ and fc-MP2/cc-pVTZ parameters.

$$\boldsymbol{HF_{CBS}} = \frac{\boldsymbol{P_{TZ}P_{5Z}} - \boldsymbol{P_{QZ}^2}}{\boldsymbol{P_{TZ}} - \boldsymbol{2P_{QZ}} + \boldsymbol{P_{5Z}}} \tag{3.3}$$

$$\boldsymbol{\Delta P_{QZ-TZ}^{MP2}} = \frac{(\boldsymbol{P_{TZ}^{MP2}} - \boldsymbol{P_{TZ}^{HF}}) \cdot \boldsymbol{n_{TZ}^3} - (\boldsymbol{P_{QZ}^{MP2}} - \boldsymbol{P_{QZ}^{HF}}) \cdot \boldsymbol{n_{QZ}^3}}{\boldsymbol{n_{TZ}^3} - \boldsymbol{n_{QZ}^3}} \tag{3.4}$$

$$\boldsymbol{P^{CBS}} = \boldsymbol{HF_{CBS}} + \boldsymbol{\Delta P_{QZ-TZ}^{MP2}} + \boldsymbol{\Delta P_{ae-fc}^{MP2}} + \boldsymbol{\Delta T} + \boldsymbol{\Delta \alpha} \tag{3.5}$$

The extrapolation scheme allows us to determine the dimers' best geometry (Geom_{best}), calculating all the parameters with the methodology illustrated above, and also the best estimated binding energy. Since this procedure is computationally expensive, we estimated the binding energy also by performing single-point calculation at the "gold-standard" fc-CCSD(T) level, with cc-pVTZ basis set. The results thus found generally were in agreement with the ones obtained with by CBS-extrapolation (see later). The Absolute Percentage Error (APE%) respect the extrapolated values of intermolecular interaction distance was evaluated to determine the accuracy of the various functionals tested in the benchmark of each dimer. Finally, for a better understanding of the functionals accuracy, the Mean Absolute Percentage Error (MAPE%) over the dimers results was evaluated. The extrapolated results are unaffected by the Basis Set Superposition Error (BSSE) (see chapter 5 of "Introduction to Computational Chemistry", F. Jensen, John Wiley & Sons, 2017 [35]). The extrapolation scheme used does not include the calculation of the frequencies, therefore it was not possible to quantify the zero point vibrational binding energy. This has allowed us to save both time and computational resources.

Critical points of the structures and their properties were computed. We focused our interest on the properties of electron density described in Chapter 2, paragraph QTAIM. Based on the work of Emamian S. *et al.* [59], we explored the trend of binding energy with the different QTAIM descriptors of the NCI critical point under consideration.

3.2 Larger and more complex clusters

The geometries of clusters with three units (or more) have been computed using an algorithm designed to return all the possible combinations in terms of spatial distribution. After that, the algorithm removed all the similar structures. This algorithm together with xTB software given the first collection of optimized structures, which were later optimize using different levels of theory, according to their size. $(HCN)_3$, $(H_2C_2)_3$, $(H_2C_2)\cdot(HCN)_2$ and $(H_2C_2)_2\cdot(HCN)$ trimers have been optimized at first using the low-cost methods HF-3c, B97-3c and PBEh-3c, with the def2-TZVP basis set. Then we proceed with more challenging DFT functionals, such as the revDSD-PBEP86 method with aug-cc-pVDZ basis set. The geometries of $(HCN)_4$, $(H_2C_2)_4$, $(H_2C_2)\cdot(HCN)_3$, $(H_2C_2)_3\cdot(HCN)$ and $(H_2C_2)_2\cdot(HCN)_2$ tetramers were optimized at first with the B97-3c/def2-TZVP method. More demanding calculations were tested, but unfortunately the structures reach the convergence with difficulties.

Each cluster with five or more units was optimized only with the B97-3c/def2-TZVP method, due to an increase in the dimensions of the system.

The "Cheap" scheme is well tested for semi-rigid systems [60], therefore, as the degrees of freedom of the clusters increases, the extrapolation may become less and less accurate and more computationally challenging. For these reasons, the composite "Cheap" Scheme wasn't performed for clusters larger than dimers. Due to its computational cost, with respect to the available computational resources illustrated below, the extrapolation scheme was not used for larger clusters, *i.e.* other than dimers.

Topological analysis was used to study the complex structures calculated. The analysis of the nature of non-covalent interactions allowed us to investigate and explain the stability of the clusters.

Finally, energy contributions per monomer molecule were analysed in each cluster. The analysis was carried out using the equations 3.6, based on the study by Conradie *et al.* [61].

$$\frac{\boldsymbol{\Delta E_n}}{n} = \frac{\boldsymbol{B.E.}}{n} \quad \text{and} \quad \frac{\boldsymbol{\Delta E_n}}{n}\% = \frac{\boldsymbol{B.E.}}{n} * 100 \tag{3.6}$$

3.3 Computational resources

Several computational facilities were used for the calculations reported in this thesis work. Specifically, we employed the High-Performance Computing (HPC) *Galileo Cluster* from the CINECA consortium and the two work stations *Mars* and *Jupiter* at Ca'Foscari's Laboratory. During the thesis internship, we faced an unscheduled shutdown of the computational resources available on Galileo Cluster starting from the second half of March; thus we were not able to complete some high-level calculations, *i.e.* revDSD-PBEP86 for some trimers.

CINECA-Galileo Cluster – GAUSSIAN 16 calculations

- Project name: CINECA Project CHIMERAS, Grant No. HP10CUY70O;
- CPU: Intel Broadwell 2x Intel Xeon E5-2697@2.3 GHz, 18 cores each;
- Total cores: 36792;
- Total nodes: 1022;
- RAM per node: 128 GB.

Mars Work Station – ORCA 4.2.1 calculations

- CPU: Intel (R) Core(TM) i7-2600@3.4 GHz;
- RAM: 8 GB.

Jupiter Work Station – ORCA 4.2.1 calculations

- CPU: Intel(R) Core(TM) i7-4790S@3.20GHz;
- RAM: 16 GB.

Chapter 4

Results and Discussion

This chapter will discuss the extrapolations and benchmark results of the four selected dimers and the monomers that compose them. The extrapolation procedure allowed us to calculate the best geometry ($Geom_{best}$) of the dimers, while the benchmark permitted to observe the evolution of the different computational methods used for the clusters $(H_2C_2)_n \cdot (HCN)_m$. By evaluating the Absolute Percentage Error (APE%) about the parameters of the $Geom_{best}$, all the DFT functions used were compared. Specifically, we consider the intermolecular bond length as more significant parameter for our analysis. The best DFT method selected, which has the better evaluation of the bond length, was used for the geometry optimization of clusters with more than two monomers. Our focus has been on the most economical calculation methods for computational resources, in anticipation of the more demanding calculations of the larger clusters.

4.1 $(HCN)_2$ Linear Dimer

The linear hydrogen cyanide was the first analyzed. The two molecules of HCN interact through the intermolecular bond between the nitrogen lone pair of electrons and the electron-poor hydrogen. As shown from the Total Electrostatic Potential (ESP) map reported in Fig. 4.1, the HCN molecule presents a negative potential on the nitrogen pole and a positive potential on the hydrogen pole. Hence, a couple of HCN molecules represents a dipole-dipole system.

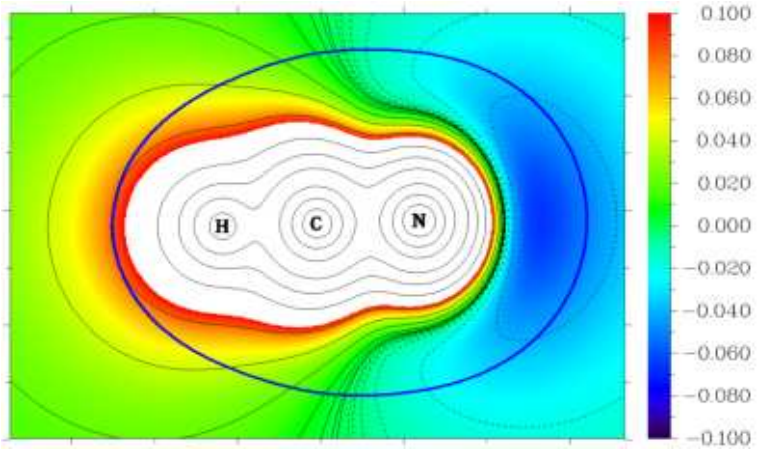


Figure 4.1: Total Electrostatic Potential map of hydrogen cyanide. The blue line represent the Van der Waals surface, the colour scale represent the electrostatic potential value.

Figure 4.2 shown the electrostatic potential change from the isolated molecule. In particular, the nuclear sphere of the hydrogen acceptor is reduced, due to the interaction with the nitrogen electron doublet.

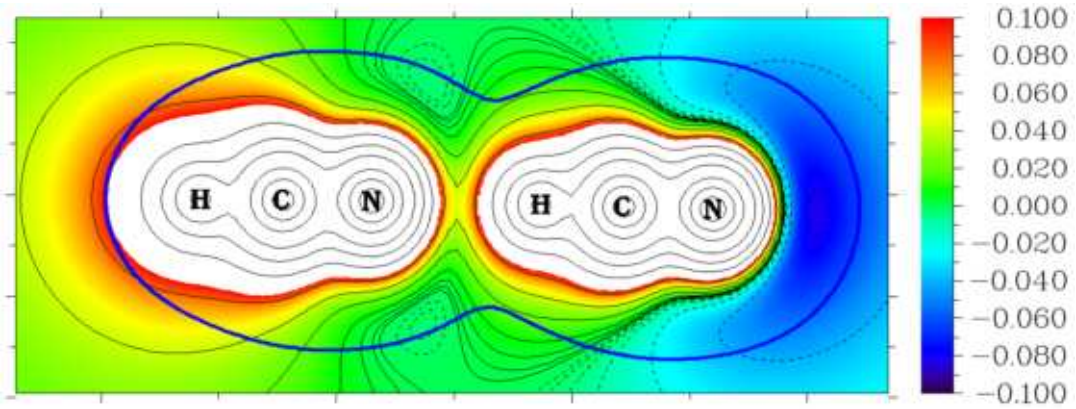


Figure 4.2: Total Electrostatic Potential map of hydrogen cyanide linear dimer. The blue line represent the Van der Waals surface, the colour scale represent the electrostatic potential value.

4.1.1 Best Geometry

The results of the extrapolation procedure for the linear dimer of hydrogen cyanide shown a linear structure with a 180° angle between the two units, as expected from our knowledge. The bond length among the hydrogen and the carbon atoms of the donor molecule are slightly shorter than the mean result provides from the DFT methods (∼1.07$\AA$). The carbon-nitrogen triple bond length decrease when interacts with the acceptor molecule. This is due to the sharing of electrons from nitrogen, so there is a depletion of electrons on the nucleus that lead to an increase of the C≡N bond strength. The bond length is around 1.15$\AA$ for the extrapolated value but generally also for the other DFT results. Regarding the acceptor molecule of HCN, the H-C bond is elongated respect to the monomer analogue. This is due to the interaction with the electron doublet of the nitrogen and the consequent weakening of the bond. Meanwhile, the intermolecular interaction has a value of 2.17$\AA$, shorter respect other results from DFT calculations. The extrapolated structure calculations of the HCN molecule shown a strong dipole moment of 3.23 D (Debye), against the experimental value 2.98 D [62], which explain the powerful interaction of two or more HCN molecules aligned together.

In Tab. 4.1 are reported the extrapolated dimer bond lengths and their deviation from the extrapolated monomer ones. In blue is represented the donor molecule, in red the acceptor molecule and in orange the intermolecular interaction.

Table 4.1 Extrapolated geometric parameters of linear hydrogen cyanide dimer and percentage deviation (PD%) from extrapolated geometric parameters of hydrogen cyanide monomer.

Bond type	**H–C**	**C≡N**	**N- - -H**	**H–C**	**C≡N**
Bond lengths ($\AA$)	1.06	1.15	**2.17**	1.07	1.15
PD% monomer	0.1%	-0.2%	–	0.6%	0.1%

The binding energy of the best geometry was computed by extrapolation at CBS and also by single-point energy calculation at fc-CCSD(T)/cc-pVTZ level on $Geom_{best}$, both calculations result in the binding energy of -23.0 kJ/mol (-5.5 kcal/mol) in agreement with the average dipole-dipole interactions value. That means the single-point calculation on the $Geom_{best}$ leads to an accurate BE value, despite the BSSE. This trend was also found in the other dimers studied.

4.1.2 Benchmark results

The linear dimer of HCN was optimized at the "gold-standard" level taken into account the *frozen core* approximation, fc-CCSD(T). The calculations were carried out with aug-cc-pVTZ basis set with a total of 23 atomic functions for each hydrogen atom and 46 functions for each heavy atom (carbon and nitrogen). In Tab. 4.2 are reported the geometric parameters and their deviation from extrapolated geometry. The calculation with fc-CCSD(T)/aug-cc-pVTZ leads to a binding energy value of -21.6 kJ/mol (-5.2 kcal/mol).

Table 4.2 Geometric parameters of linear hydrogen cyanide dimer calculated at fc-CCSD(T)/aug-cc-pVTZ level and percentage error (PE%) from extrapolated geometric parameters.

Bond type	**H–C**	**C≡N**	**N- - -H**	**H–C**	**C≡N**
Bond lengths ($\AA$)	1.07	1.16	**2.20**	1.07	1.16
PE% $\mathbf{Geom}_{best}$	0.46%	0.71%	**1.38%**	0.46%	0.68%

Both B2PLYP and B3LYP methods with the triple-ζ and double-ζ SNSD basis sets and considering the D3BJ correction, achieve Absolute Percentage Error (APE%) values inferior to 1% for the evaluation of intermolecular bond. As expected, values of APE% without the D3BJ correction are greater than the corrected ones, due to underestimation of dispersion forces. From the graphs below it can be concluded that the use of the D3BJ correction is necessary even in a dimer where dipole-dipole forces are predominate.

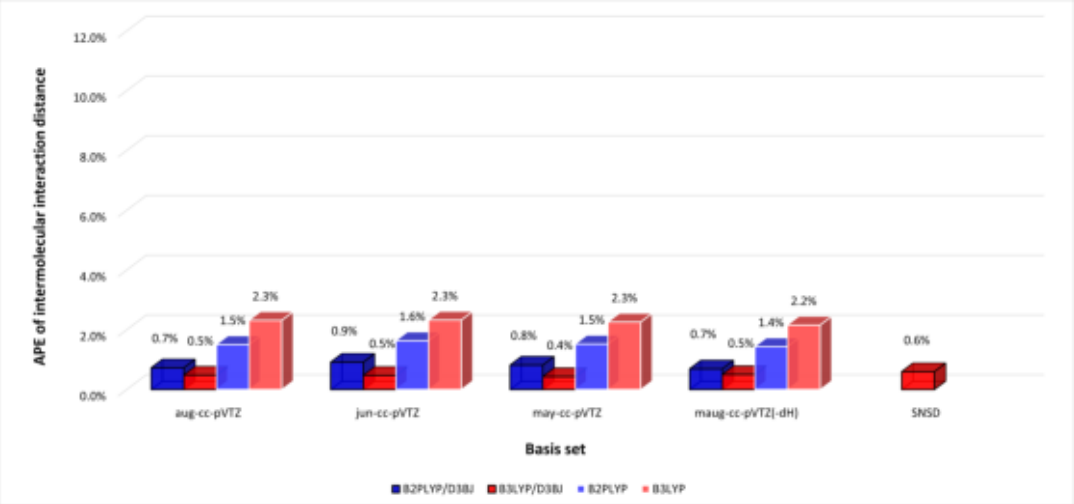


Figure 4.3: Absolute Percentage Error (APE%) of intermolecular distance parameters of hybrid DFT functionals B2PLYP e B3LYP with calendar series of basis set. Both functionals were tested with and without the D3BJ correction of dispersion forces.

The "low-cost" semi-empirical methods xTB, HF-3c, PBEh-3c and B97-3c return good intermolecular interaction length APE%, the lowest value of 1.2% is obtained by the xTB method with the minimum basis set for each atom.

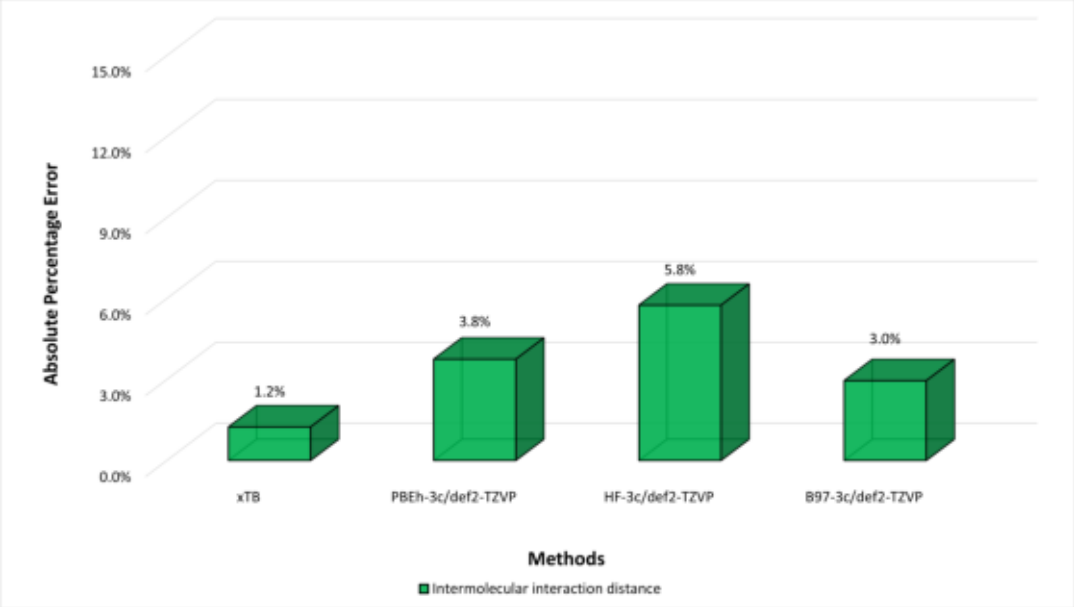


Figure 4.4: Absolute Percentage Error (APE%) of intermolecular distance parameter of "low-cost" semi-empirical methods.

The Minnesota functional M06-2X proved to be sufficiently accurate in the calculation of bond lengths related to this system. Functionals such as the CAM-B3LYP, LC-ωPBE, DSD-PBEP86 and its revised version, accompanied by triple-ζ basis sets, proved to be very accurate, for geometry calculations. Specifically, the CAM-B3LYP/aug-pc-3 results are among the most accurate. Nevertheless, these methods are too expensive for clusters with more units than the dimer under examination.

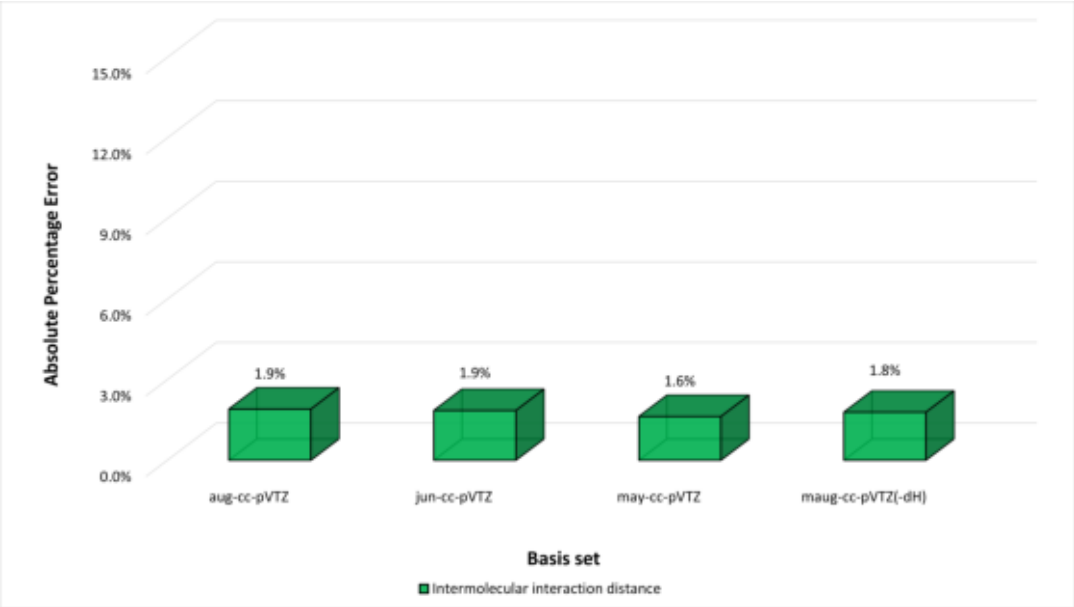


Figure 4.5: Absolute Percentage Error (APE%) of intermolecular distance parameter of M06-2X functional with different calendar series of basis set.

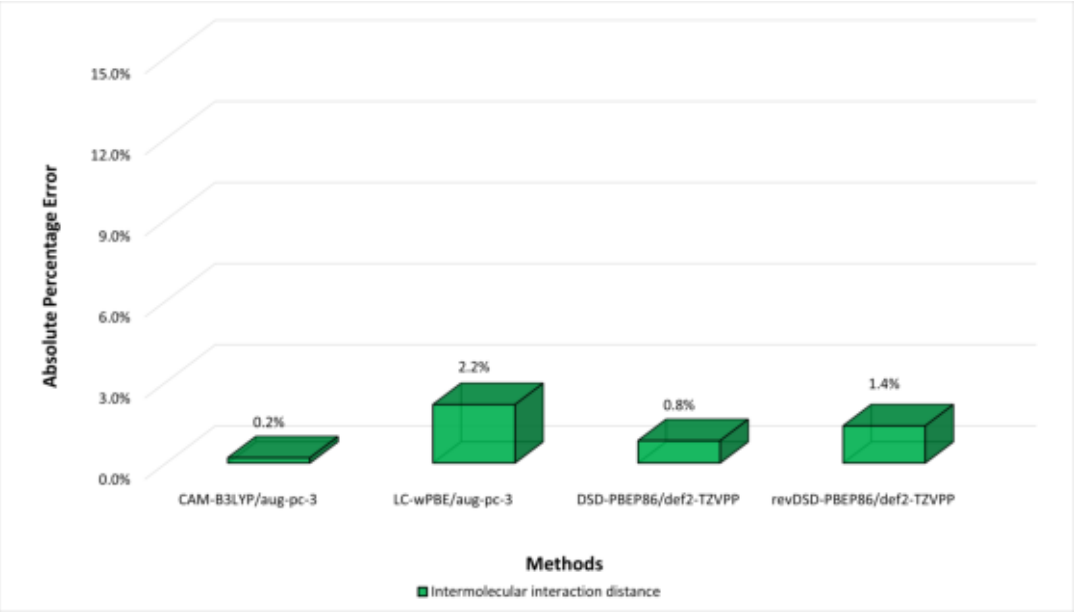


Figure 4.6: Absolute Percentage Error (APE%) of intermolecular distance parameter of CAM-B3LYP, LC-ωPBE, DSD-PBEP86 and rev-DSD-PBEP86 with appropriated triple-ζ basis sets.

Fig. 4.7 resume the absolute percentage errors for each method tested for HCN linear dimer. Generally, all these methods lead to a good valuation of NCI bond length, but focusing our attention on semi-empirical functionals, it can be observed that the xTB method goes well for geometric parameters calculation. Also, the other three "low-cost" methods tested goes quite well if we consider their practicality.

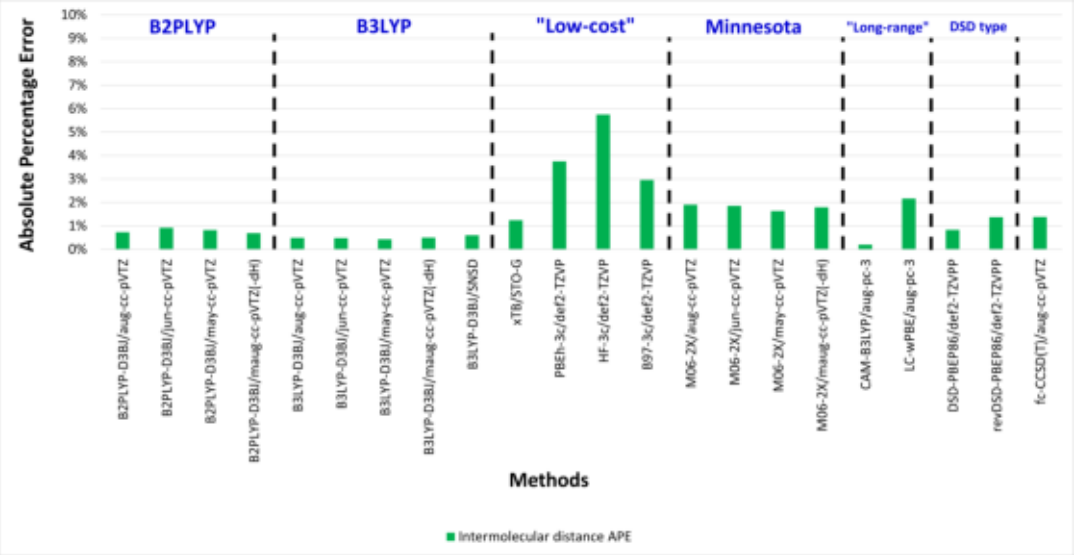
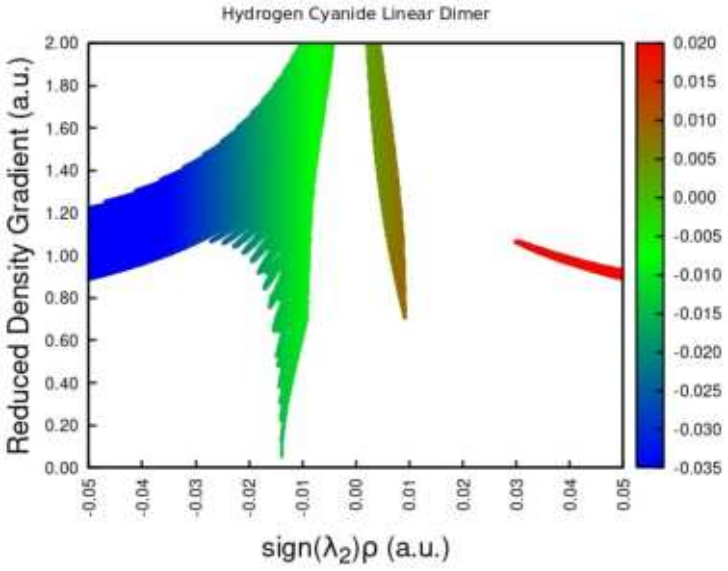


Figure 4.7: Benchmark resume for linear hydrogen cyanide dimer.

4.1.3 Topology Analysis

The topology analysis, as expected, has identified a Bond Critical Point between the nitrogen and the hydrogen of the two monomers. The Fig. 4.8b shows a strong interaction between the two attractors N- - -H, where in Fig. 4.8a the NCI is illustrated by the green isosurface.

(a)

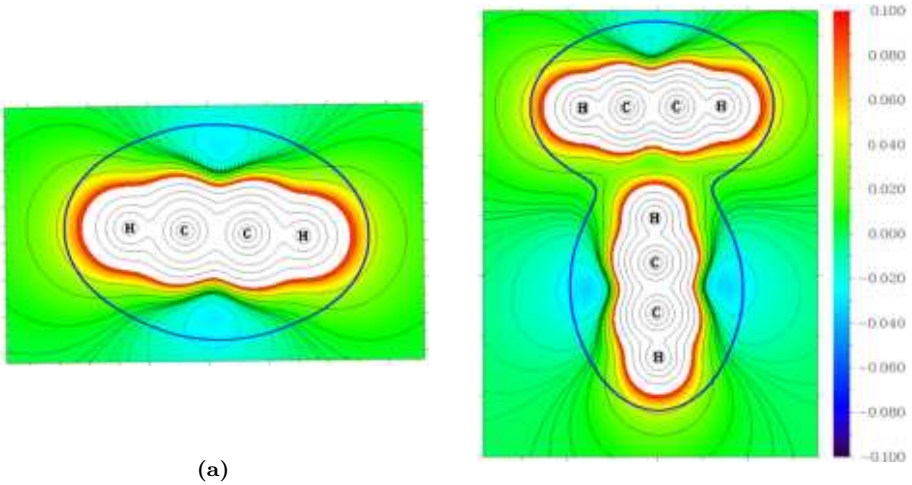


(b)

Figure 4.8: Isosurfaces (RDG($\boldsymbol{\rho}$) = 0.5) and Reduced Density Gradient vs sign(λ_2)·ρ plot of hydrogen cyanide dimer.

4.2 $(H_2C_2)_2$ T-Shaped Dimer

The T-Shaped dimer of H_2C_2 interaction is dominated by the dispersion forces due to the π-bond. The acetylene molecule is perfectly symmetric, and also its electron density. As shown in Fig. 4.9a, the hydrogen atoms are quite positive and can interact with the π-electrons of the other molecule.



(a) **(b)**

Figure 4.9: Total Electrostatic Potential map of acetylene monomer and dimer. The blue line represent the Van der Waals surface, the colour scale represent the electrostatic potential value.

In Fig. 4.9b, it is possible to notice how the electronic clouds of the two acetylene molecules have slightly changed. The region around the hydrogen of the non-covalent bond has an inappreciably lower potential than the hydrogen of the isolated molecule, indicating an interaction with the π system.

4.2.1 Best Geometry

Extrapolated geometry parameters for the T-Shaped acetylene dimer are quite similar to the monomer. Probably, due to the weak interaction of dispersion forces which are predominant in dimer structure, the intramolecular bonds don't deform as much as in the case of dipole-dipole interactions. The intermolecular distance has a value of 2.62 $\AA$ and is higher than the other dimers, the weakness of dispersion forces is emphasised by the large bond length.

In Tab. 4.3 are reported the extrapolated dimer bond lengths and their deviation from the monomer ones. In blue is represented the donor molecule, in red the acceptor molecule and in orange the intermolecular interaction. Since, the donor molecule interacts only with the π bond, the two H–C bonds are perfectly symmetric and are reported only once in table.

Table 4.3 Extrapolated geometric parameters of T-Shaped acetylene dimer and percentage deviation (PD%) from extrapolated geometric parameters of acetylene monomer.

Bond type	**H–C**	**C≡C**	**π- - -H**	**H–C**	**C≡C**	**H–C**
Bond lengths ($\AA$)	1.06	1.20	**2.62**	1.06	1.20	1.06
PD% monomer	0.0%	0.0%	–	0.2%	0.0%	0.0%

The extrapolated binding energy value of acetylene dimer is -8.2 kJ/mol (∼2.0 kcal/mol), a very weak interaction as expected from a dimer of two non-polar molecules. Probably, the π-electrons induce a dipole moment of the acceptor molecule and established a bond. The binding energy related to this dimer is lower than the others, proving the weakness of London forces.

4.2.2 Benchmark results

The acetylene dimer was optimized at the "gold-standard" level taken into account the *frozen core* approximation, fc-CCSD(T). The calculations were carried out with aug-cc-pVTZ basis set with a total of 23 atomic functions for each hydrogen atom and 46 functions for each heavy atom (carbon). In Tab. 4.4 are reported the geometric parameters and their deviation from extrapolated geometry. The calculation with fc-CCSD(T)/aug-cc-pVTZ leads to a binding energy value of -7.5 kJ/mol (-1.8 kcal/mol).

Table 4.4 Geometric parameters of T-Shaped acetylene dimer calculated at fc-CCSD(T)/aug-cc-pVTZ level and percentage error (PE%) from extrapolated geometric parameters.

Bond type	**H–C**	**C≡C**	**π- - -H**	**H–C**	**C≡C**	**H–C**
Bond lengths ($\AA$)	1.07	1.21	**2.64**	1.07	1.21	1.06
PE% Geom$_{best}$	0.5%	0.7%	**1.0%**	0.4%	0.7%	0.4%

The results obtained with the hybrid DFT functionals B2PLYP and B3LYP tested are acceptable for the calculation of bond lengths. The use of D3BJ correction is crucial for a better understanding of dispersion forces, especially for the acetylene dimer dominated from London forces.

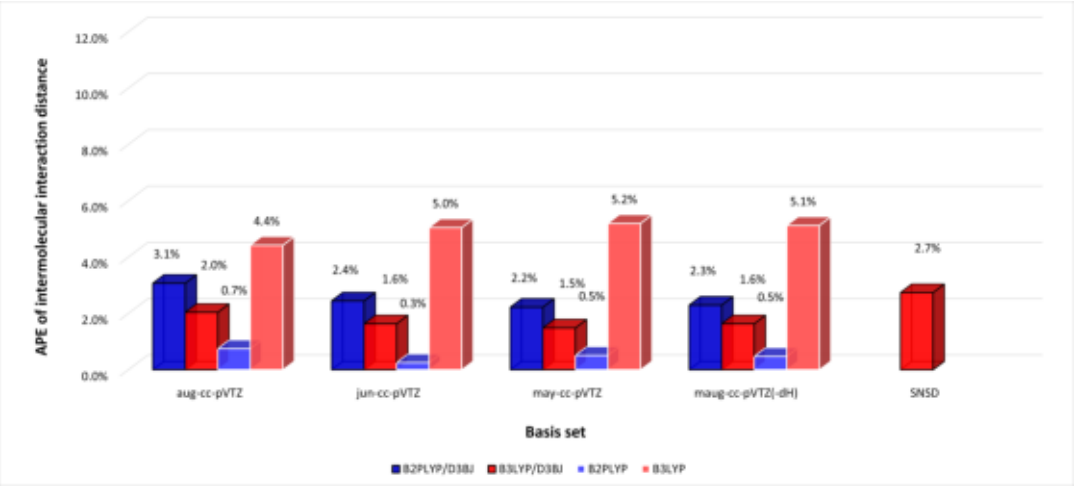


Figure 4.10: Absolute Percentage Error (APE%) of intermolecular distance parameters of hybrid DFT functionals B2PLYP e B3LYP with calendar series of basis set. Both functionals were tested with and without the D3BJ correction of dispersion forces.

In this case, the "low-cost" functionals B97-3c and PBEh-3c proved to be the most accurate between semi-empirical methods describing the NCIs. The APE% values for the intermolecular distance are very accurate and the B97-3c/def2-TZVP method remains one of the most accurate methods along with the B3LYP.

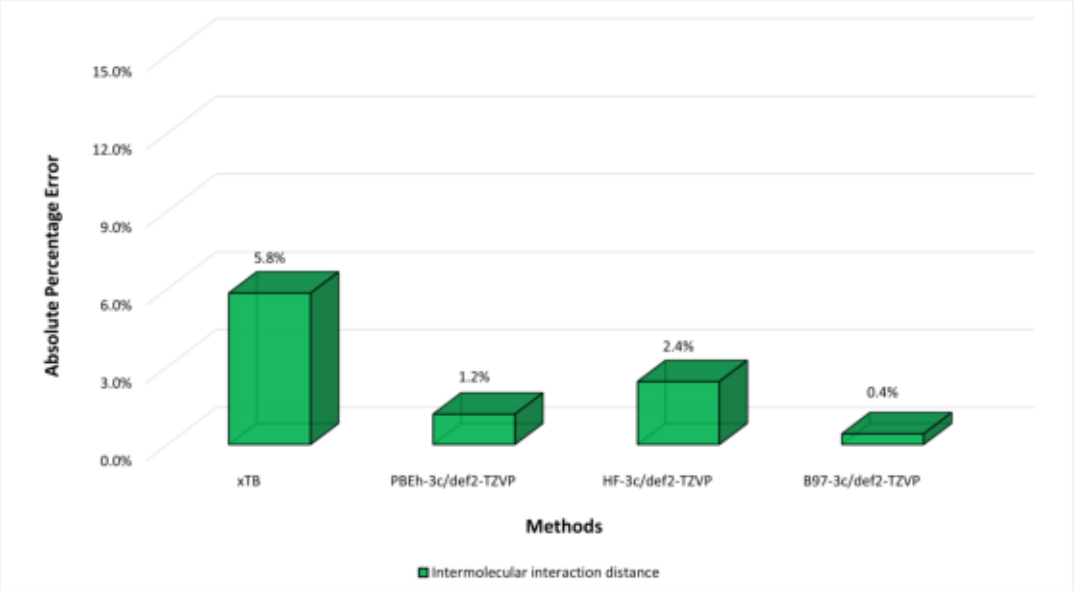


Figure 4.11: Absolute Percentage Error (APE%) of intermolecular distance parameter of "low-cost" semi-empirical methods.

The M06-2X functional with the calendar series of basis set have an accurate description of the bond length. Generally, long-range functionals were able to describe fairly accurately the bond length π- - -H.

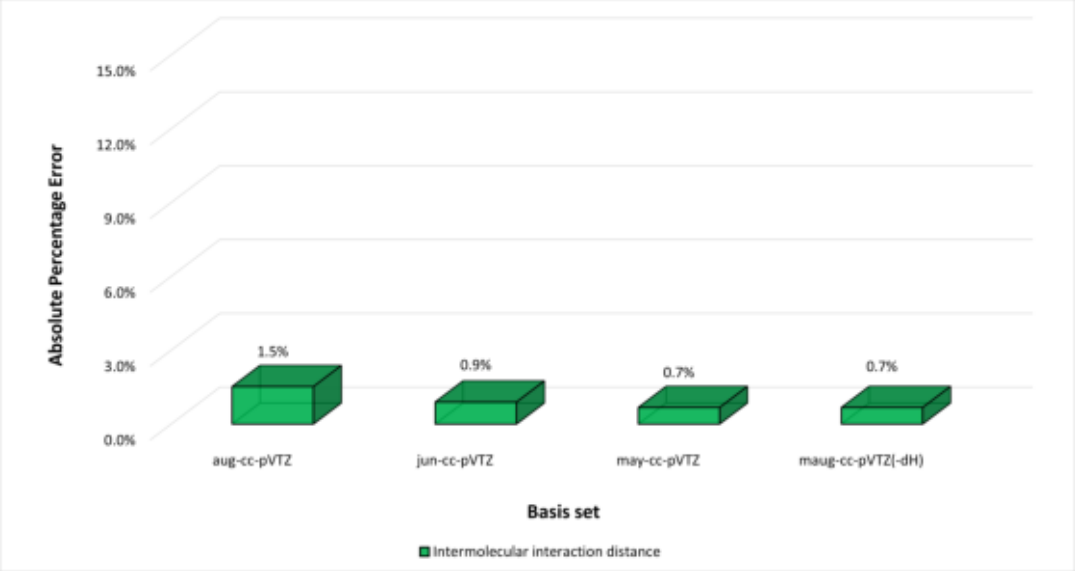


Figure 4.12: Absolute Percentage Error (APE%) of intermolecular distance parameter of M06-2X functional with different calendar series of basis set.

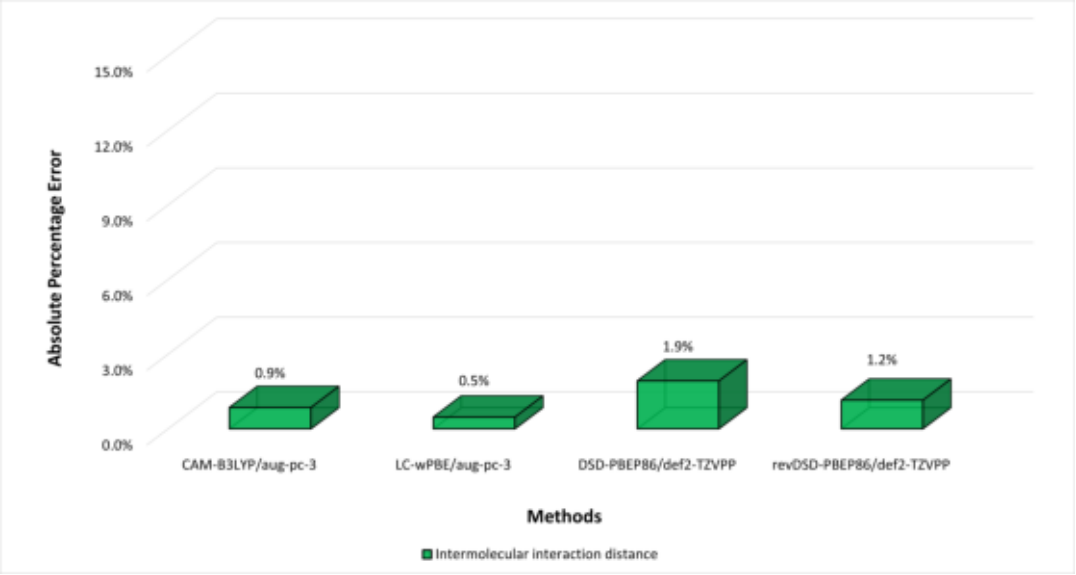
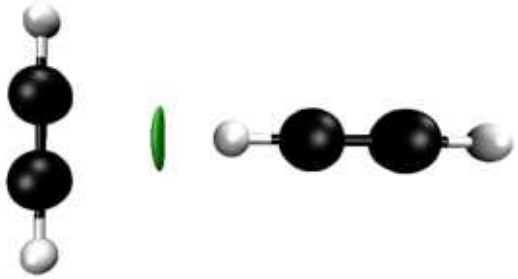


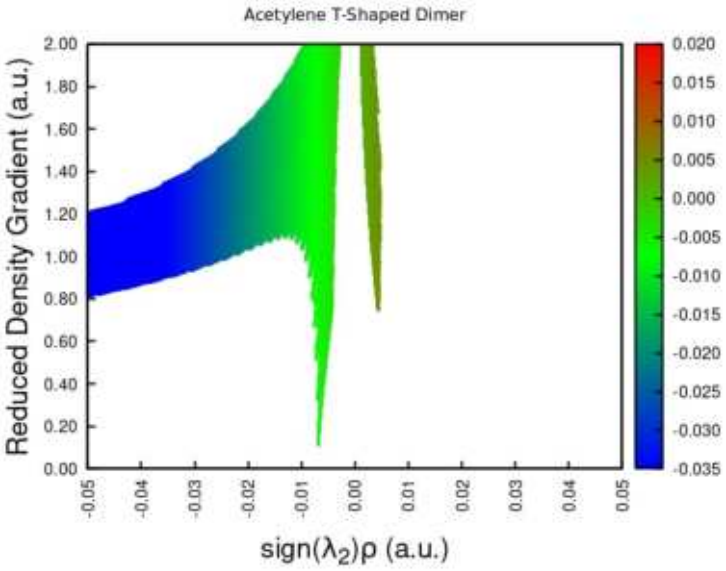
Figure 4.13: Absolute Percentage Error (APE%) of intermolecular distance parameter of CAM-B3LYP, LC-ωPBE, DSD-PBEP86 and rev-DSD-PBEP86 with appropriated triple-ζ basis sets.

4.2.3 Topology Analysis

Fig. 4.14a shows the reduced density gradient isosurface between the two molecules along the π- - -H bond path. The RDG plot below display a weak interaction between zero and -0.01 a.u. value of sign(λ_2)·ρ.



(a)



(b)

Figure 4.14: Isosurfaces (RDG($\boldsymbol{\rho}$) = 0.5) and Reduced Density Gradient vs sign(λ_2)·ρ plot of acetylene dimer.

4.3 (HCN)·(H_2C_2) Linear Dimer

The third dimer analyzed has been the mixed cluster of hydrogen cyanide and acetylene. The interaction between the nitrogen of HCN molecules and the hydrogen of H_2C_2 molecule leads the structure to have a linear configuration. Fig. 4.15 shows as the electrostatic potential is more negative in the NCI and the π regions, compared to the others dimers. This is probably due to the polarization of the electron cloud of acetylene due to the approach of the electron doublet of nitrogen, with the formation of a temporary dipole on acetylene.

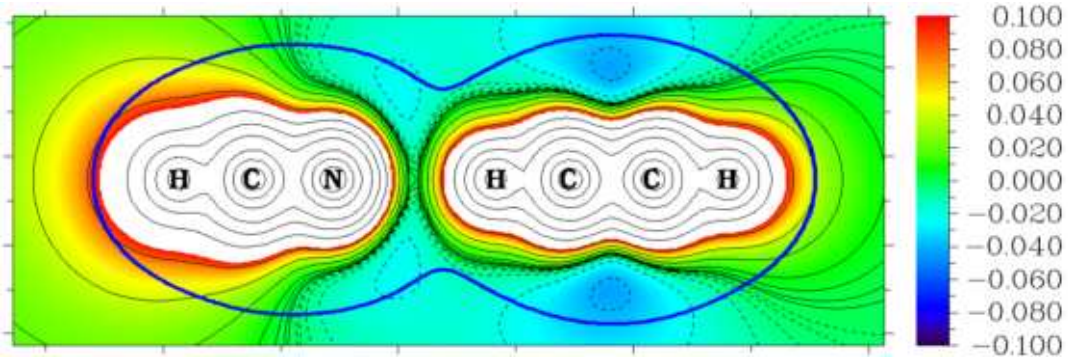


Figure 4.15: Total Electrostatic Potential map of (HCN)·(H_2C_2) linear dimer. The blue line represent the Van der Waals surface, the colour scale represent the electrostatic potential value.

4.3.1 Best Geometry

The bond length of hydrogen cyanide molecule are shorter than the monomer, especially the H–C bond which is 0.7% shorter than the analogue bond in HCN monomer. A shortening in bond length leads an increase in bond strength. While, the length of H–C bond of acetylene is slightly elongated due to the interaction with HCN. The approach of non-bonding electronic doublet of nitrogen on acetylene molecule, leads an increase of electron potential around the π bond. The NCI length is lightly greater than the HCN linear dimer, which is the strongest, and has a value of 2.31 $\AA$. In Tab. 4.5 are reported the extrapolated dimer bond lengths and their deviation from the monomer ones. In blue is represented the donor molecule, in red the acceptor molecule and in orange the intermolecular interaction.

Table 4.5 Extrapolated geometric parameters of linear mixed dimer of HCN and H_2C_2 and percentage deviation (PD%) from extrapolated geometric parameters of their monomers.

Bond type	**H–C**	**C≡N**	**N- - -H**	**H–C**	**C≡C**	**C–H**
Bond lengths ($\AA$)	1.05	1.15	**2.31**	1.06	1.20	1.06
PD% monomer	-0.7%	-0.1%	–	0.3%	0.1%	0.0%

The extrapolated binding energy value is of -12.3 kJ/mol (-2.9 kcal/mol), as expected from an hydrogen bond.

4.3.2 Benchmark results

The linear heteromolecular dimer was optimized at the "gold-standard" level taken into account the *frozen core* approximation, fc-CCSD(T). The calculations were carried out with aug-cc-pVTZ basis set with a total of 23 atomic functions for each hydrogen atom and 46 functions for each heavy atom (carbon and nitrogen). In Tab. 4.6 are reported the geometric parameters and their deviation from extrapolated geometry. The calculation with

fc-CCSD(T)/aug-cc-pVTZ leads to a binding energy value of -12.3 kJ/mol (-2.9 kcal/mol), coincident with the $Geom_{best}$ binding energy value.

Table 4.6 Geometric parameters of linear acetylene and hydrogen cyanide dimer calculated at fc-CCSD(T)/aug-cc-pVTZ level and percentage error (PE%) from extrapolated geometric parameters.

Bond type	**H–C**	**C≡N**	**N- - -H**	**H–C**	**C≡C**	**C–H**
Bond lengths ($\AA$)	1.07	1.16	**2.34**	1.07	1.21	1.06
PE% $Geom_{best}$	1.3%	0.7%	**1.4%**	0.5%	0.6%	0.5%

The DFT functionals B2PLYP and B3LYP with the correction D3BJ return very accurate bond length values with an error, compared to the "Cheap" Scheme, around 1%. The absence of D3BJ correction, in general, leads to high error except for the B2PLYP functional, which thanks to the solid PT2 correlation energy part is able to capture quite well the NCI length parameter.

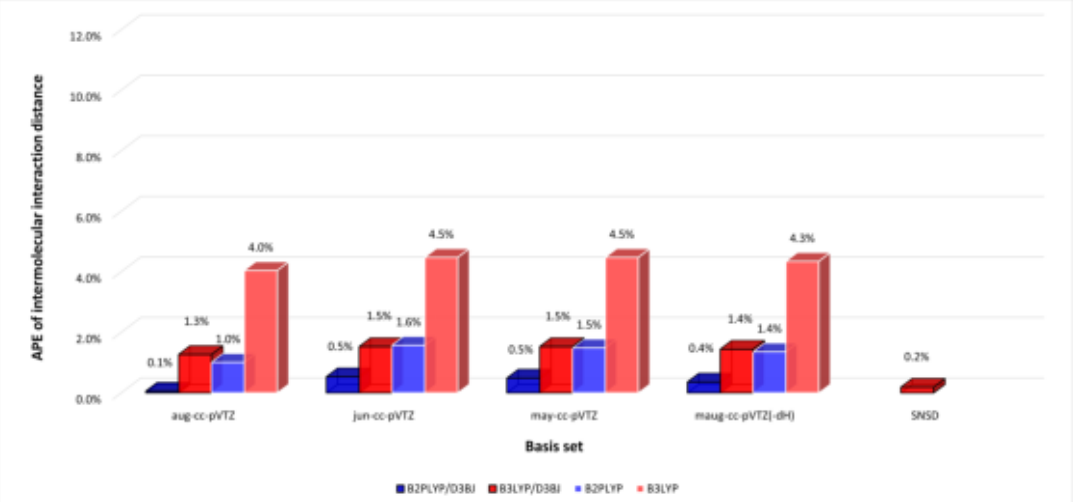


Figure 4.16: Absolute Percentage Error (APE%) of intermolecular distance parameter of hybrid DFT functionals B2PLYP e B3LYP with calendar series of basis set. Both functionals were tested with and without the D3BJ correction of dispersion forces.

The semi-empirical methods, compared with hybrid DFT functionals, leads to good APE% values, especially xTB method. Considering the low computational resources required by semi-empirical methods, these results should be considered despite the high APE% values compared to DFT methods.

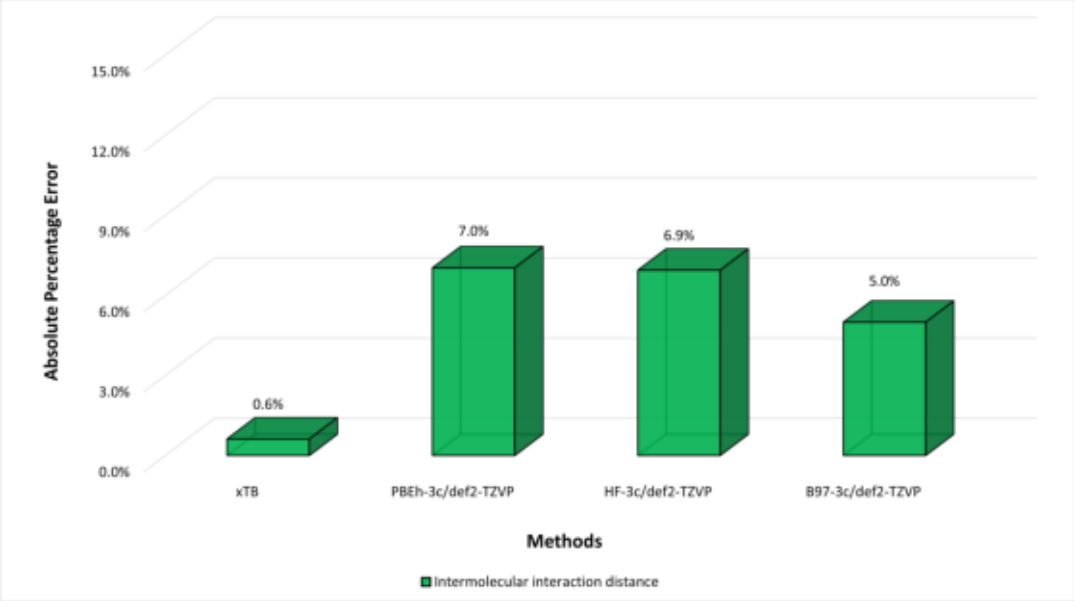


Figure 4.17: Absolute Percentage Error (APE%) of intermolecular distance parameter of "low-cost" semi-empirical methods.

The M06-2X functional describe the bond length less accurately than the other DFT methods. CAM-B3LYP/def2-TZVP method leads to an accurate calculation of bond length, but it is pretty time expensive.

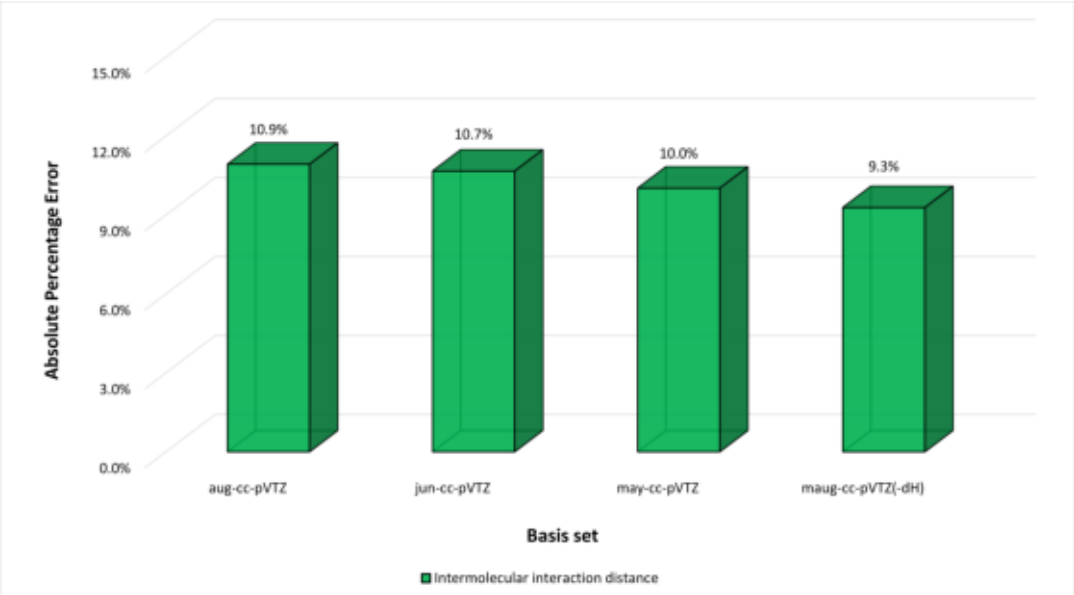


Figure 4.18: Absolute Percentage Error (APE%) of intermolecular distance parameter of M06-2X functional with different calendar series of basis set.

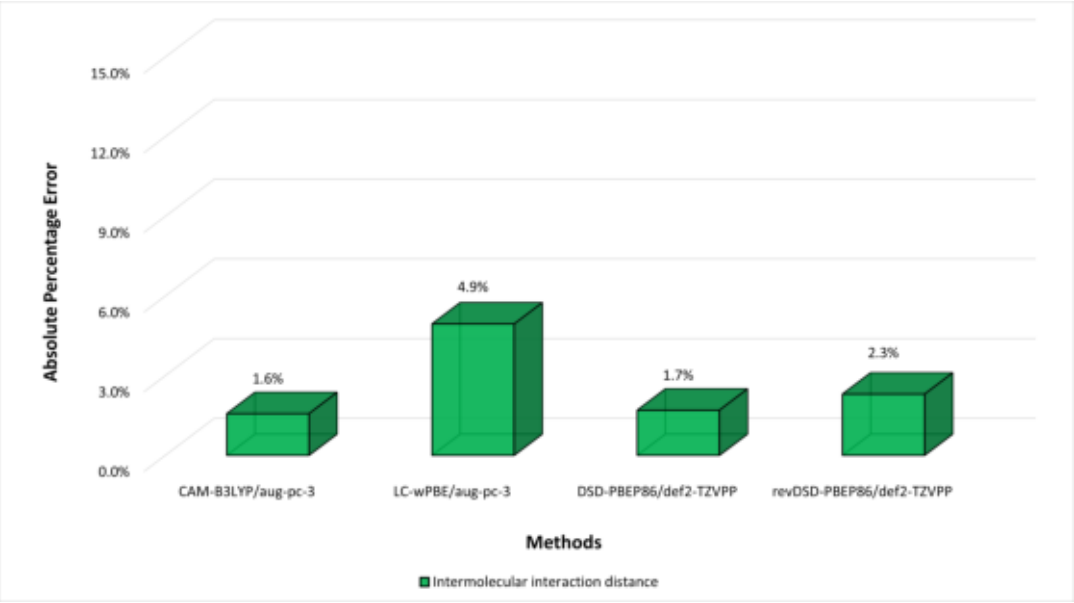


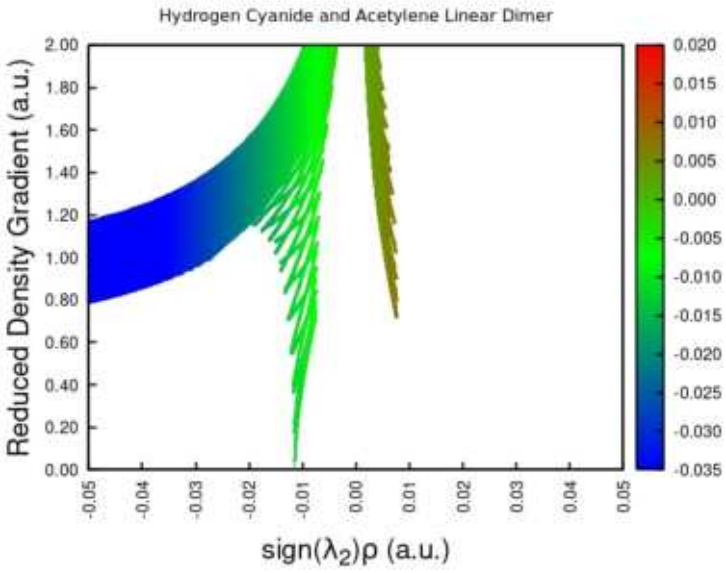
Figure 4.19: Absolute Percentage Error (APE%) of intermolecular distance parameter of CAM-B3LYP, LC-ωPBE, DSD-PBEP86 and rev-DSD-PBEP86 with appropriated triple-ζ basis sets.

Generally, all these methods lead to a good valuation of NCI bond length. Focusing our attention on semi-empirical functionals, it can be observed that the xTB method goes well for geometric parameters calculation, while the other three "low-cost" methods tested go quite well if we consider them from the point of view of the computational time.

4.3.3 Topology Analysis

Fig. 4.20a display the RDG isosurface among the NCI path, indicating the donation of σ electrons from the nitrogen to the acetylene hydrogen. The RDG plot shows a strongest interaction than the acetylene dimer due to the inducted dipole temporary established in acetylene molecule.

(a)



(b)

Figure 4.20: Isosurfaces (RDG($\boldsymbol{\rho}$) = 0.5) and Reduced Density Gradient vs sign(λ_2)·ρ plot of the HCN and H_2C_2 linear dimer.

4.4 $(H_2C_2)\cdot(HCN)$ T-Shaped Dimer

The interaction between the two monomers is weak. As can be seen from the Fig. 4.21, the electrostatic potential around the NCI region highlights the dominance of the π electrons in the bond.

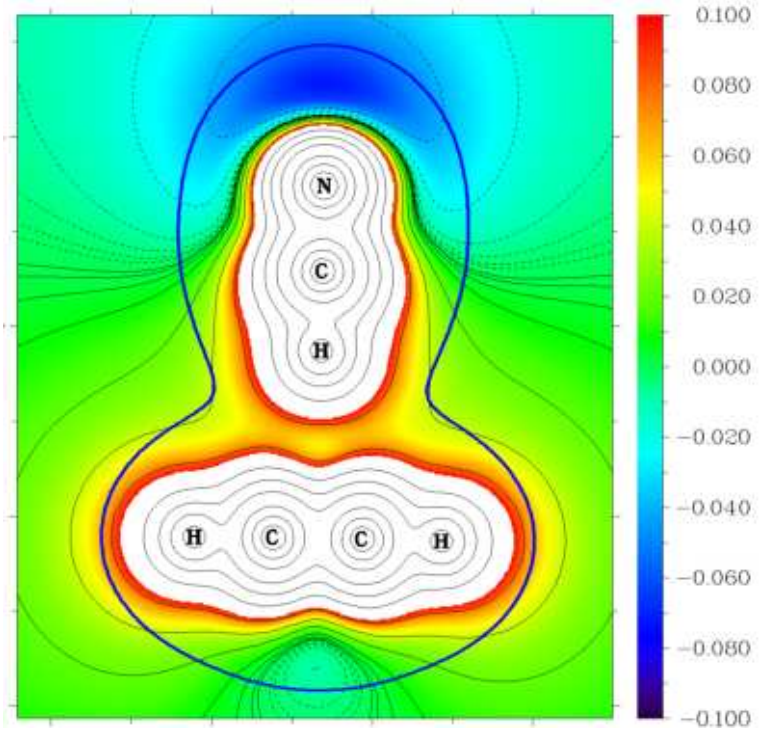


Figure 4.21: Total Electrostatic Potential map of $(H_2C_2)\cdot(HCN)$ T-Shaped dimer. The blue line represent the Van der Waals surface, the colour scale represent the electrostatic potential value.

4.4.1 Best Geometry

In the case of the non-linear mixed dimer, the bond length are quite similar to the isolated monomers, except for the H–C bond of hydrogen cyanide molecule, which is slightly elongated due the interaction with π system. The intermolecular bond length has a value of 2.50 $\AA$ and is a little larger than the mixed linear structure, due to the dispersion forces nature. In Tab. 4.7 are reported the extrapolated dimer bond lengths and their deviation from the monomer ones. In blue is represented the donor molecule, in red the acceptor molecule and in orange the intermolecular interaction. Since, the donor molecule interacts only with the π bond, the two H–C bonds are perfectly symmetric and are reported only once in table.

Table 4.7 Extrapolated geometric parameters of T-Shaped mixed dimer of HCN and H_2C_2 and percentage deviation percentage deviation (PD%) from extrapolated geometric parameters of their monomers.

Bond type	**H–C**	**C≡C**	**π- - -H**	**H–C**	**C≡N**
Bond lengths ($\AA$)	1.06	1.20	**2.50**	1.07	1.15
PD% monomer	0.1%	0.0%	–	0.6%	0.0%

The binding energy of this structure has a value of -13.4 kJ/mol (-3.2 kcal/mol) and is a little bit more stable than the linear analogue.

4.4.2 Benchmark results

The T-Shaped heteromolecular dimer was optimized at the "gold-standard" level taken into account the *frozen core* approximation, fc-CCSD(T). The calculations were carried out with aug-cc-pVTZ basis set with a total of 23 atomic functions for each hydrogen atom and 46 functions for each heavy atom (carbon and nitrogen). In Tab. 4.8 are reported the geometric parameters and their deviation from extrapolated geometry. The calculation with fc-CCSD(T)/aug-cc-pVTZ leads to a binding energy value of -11.7 kJ/mol (-2.8 kcal/mol), 12.5% in a reasonable agreement with the $Geom_{best}$ binding energy value of -13.4 kJ/mol.

Table 4.8 Geometric parameters of T-Shaped acetylene and hydrogen cyanide dimer calculated at fc-CCSD(T)/aug-cc-pVTZ level and percentage error (PE%) from extrapolated geometric parameters.

Bond type	**H–C**	**C≡C**	**π- - -H**	**H–C**	**C≡N**
Bond lengths ($\AA$)	1.07	1.21	**2.51**	1.07	1.15
PE% $Geom_{best}$	0.5%	0.7%	**0.7**%	0.3%	0.7%

The DFT methods B2PLYP and B3LYP are particularly accurate for bond length calculations, removal of D3BJ correction leads to higher errors. The B3LYP with the double-ζ basis set SNSD results more accurate than the others.

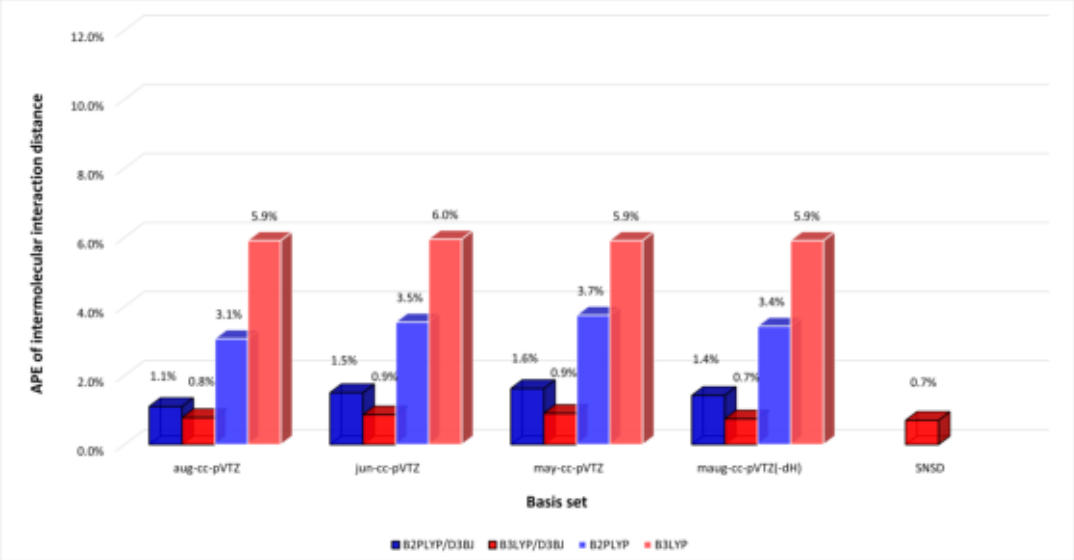


Figure 4.22: Absolute Percentage Error (APE%) of intermolecular distance parameter of hybrid DFT functionals B2PLYP e B3LYP with calendar series of basis set. Both functionals were tested with and without the D3BJ correction of dispersion forces.

"Low-cost" method xTB has a very accurate understanding of bond length. Generally, all semi-empirical methods tested led to good APE% values for the geometric parameter, where the worse method among them, HF-3c/def2-TZVP, leads to a 5.4% of APE, which is reasonable.

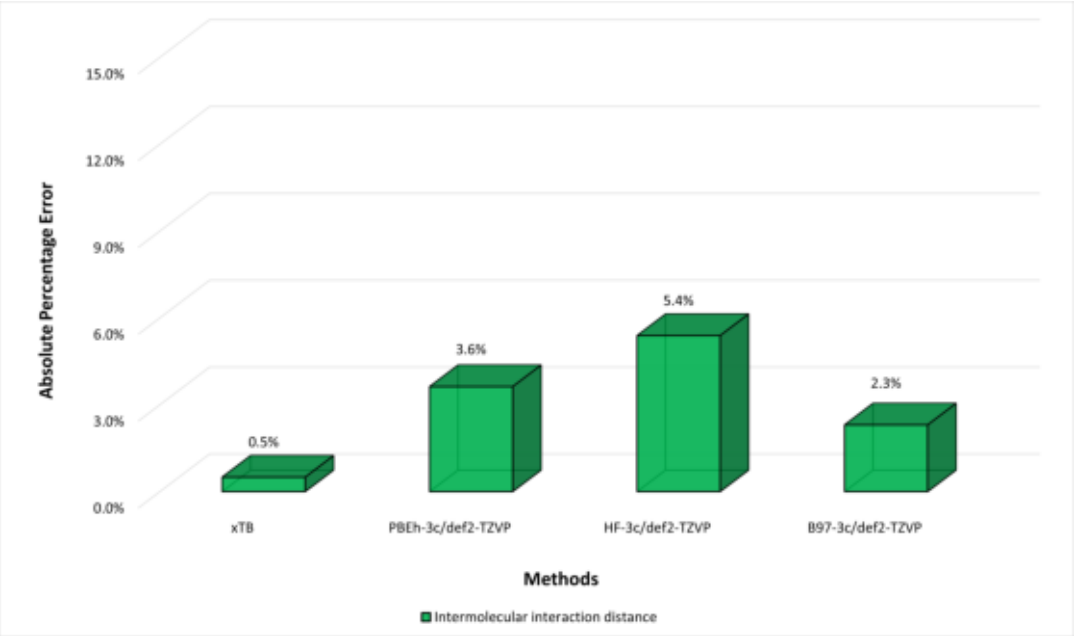


Figure 4.23: Absolute Percentage Error (APE%) of intermolecular distance parameter of "low-cost" semi-empirical methods.

The Minnesota functional M06-2X combined with the calendar basis sets has a good grasp for NCI length. The CAM-B3LYP/aug-pc-3 is the most accurate for binding energy calculation between the "long-range". Both DSD-PBEP86 and revDSD-PBEP86 goes well for geometry calculations.

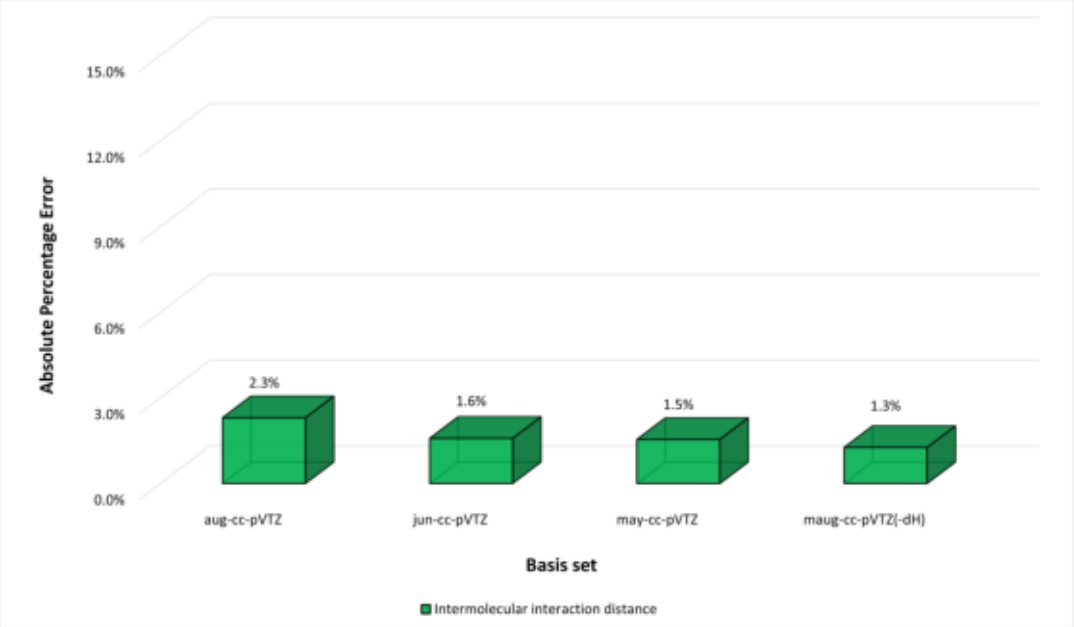


Figure 4.24: Absolute Percentage Error (APE%) of intermolecular distance parameter of M06-2X functional with different calendar series of basis set.

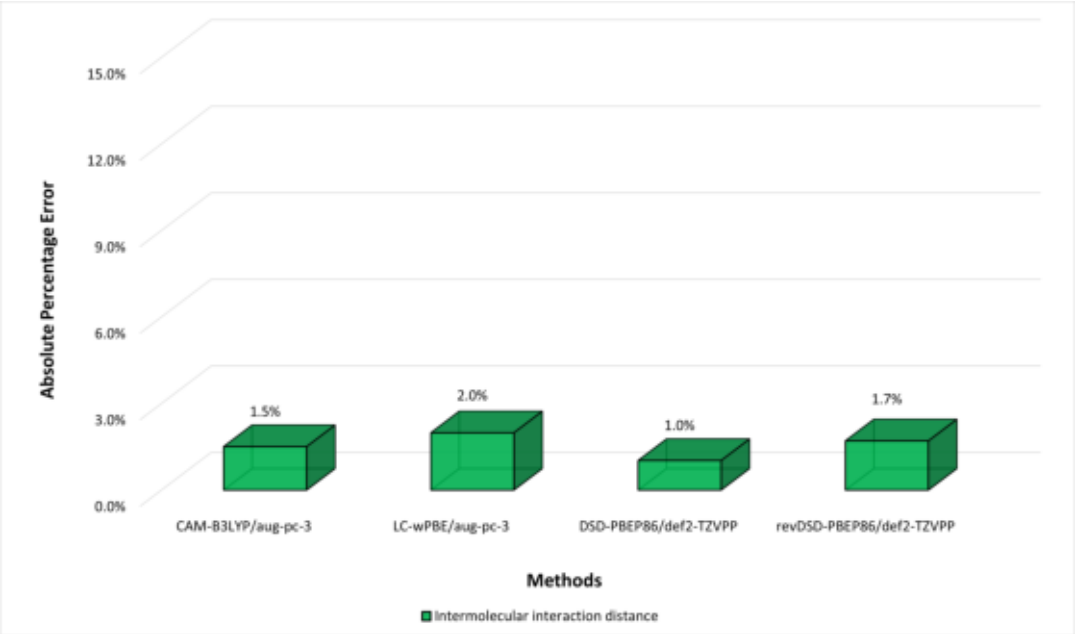
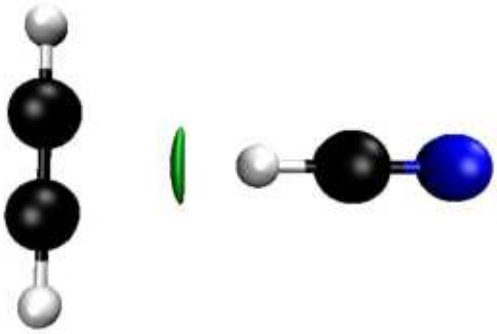


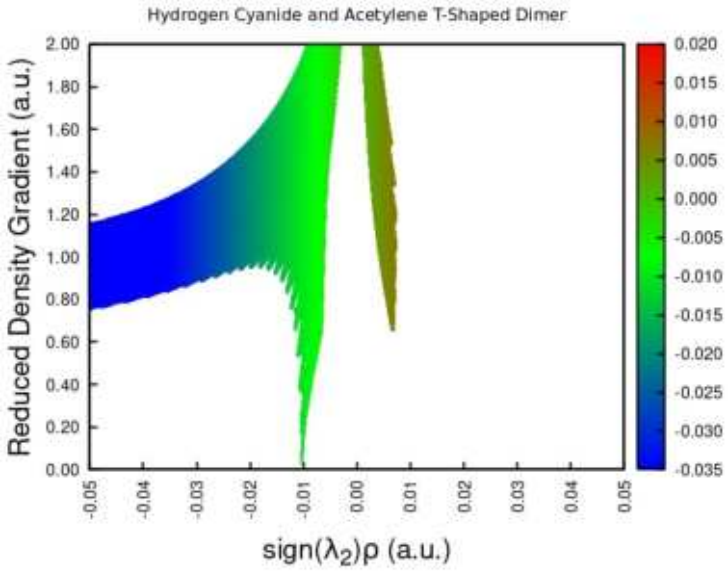
Figure 4.25: Absolute Percentage Error (APE%) of intermolecular distance parameter of CAM-B3LYP, LC-ωPBE, DSD-PBEP86 and rev-DSD-PBEP86 with appropriated triple-ζ basis sets.

4.4.3 Topology Analysis

Fig. 4.26a display the RDG isosurface among the NCI path, indicating the π-electrons interaction from the acetylene to electron-poor hydrogen atom of HCN molecule. In Fig. 4.26b are shown a Van der Waals interaction at -0.01 a.u., stronger than T-Shaped acetylene dimer, due to hydrogen cyanide electron polarization, but slightly weaker than HCN- - -H_2C_2 linear dimer due to a more conducive σ-bond and a better polarization of acetylene molecule.



(a)



(b)

Figure 4.26: Isosurfaces (RDG(ρ) = 0.5) and Reduced Density Gradient vs sign(λ_2)·ρ plot of the HCN and H_2C_2 linear dimer.

4.5 Binding Energy vs AIM descriptors Analysis

The nature of non-covalent interactions can be studied by the analysis of QTAIM descriptors and binding energy. The calculations have been carried out with fc-CCSD(T)/aug-cc-pVTZ method and analyzed through Multiwfn software. The parameters at the bond critical point, such as electron density, potential and kinetic energy, correspond at some interesting proprieties characterizing the complex. All the descriptors are reported in atomic units because we are interested in the trend of these variables. The first descriptor analysed is electron density at Bond Critical Point. Figure 4.27 shows that at higher electron density values there is a stronger interaction between the two monomers, *i.e.* the binding energy is more negative. As explained before, the linear dimers are the stable ones, due to the arrangement in the space that favours interaction. This is consistent with the increase of electrons in the area where the interaction is taking place, the more electrons are involved in the bond, the more stable

the complex will be. Despite the few measurements it is interesting to observe the linear correlation of these plots. A coefficient R^2 equal to 0.98 means that there is a very good correlation between the two variables. So it is possible to say that dimer binding energy and electron density are correlated.

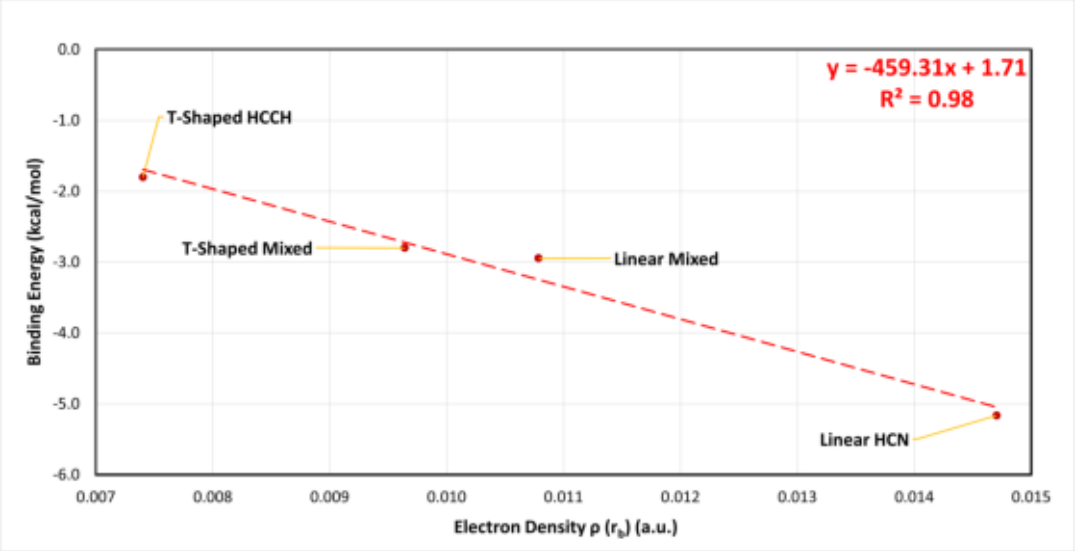


Figure 4.27: Plot of binding energy vs electron density on Bond Critical Point of each dimer.

The regression plots of BEs against $H(r_b)$, $H(r_b)/\rho(r_b)$ and $|V(r_b)|/G(r_b)$ at the BCP for the entire set of dimers are given in figures 4.28, 4.29 and 4.30. Unfortunately, none of them correlates well with BE values.

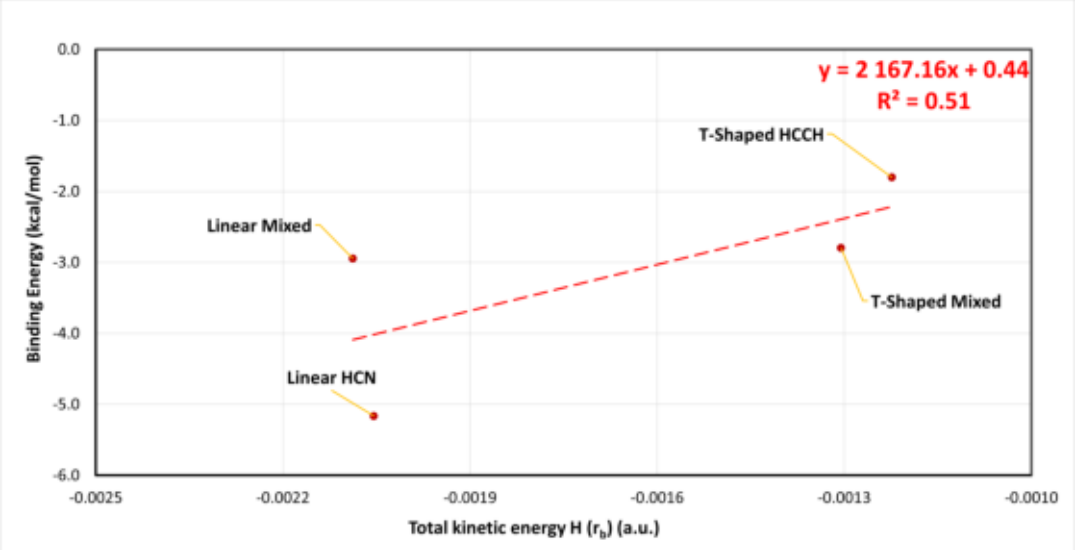


Figure 4.28: Plot of binding energy vs total kinetic energy density on Bond Critical Point of each dimer.

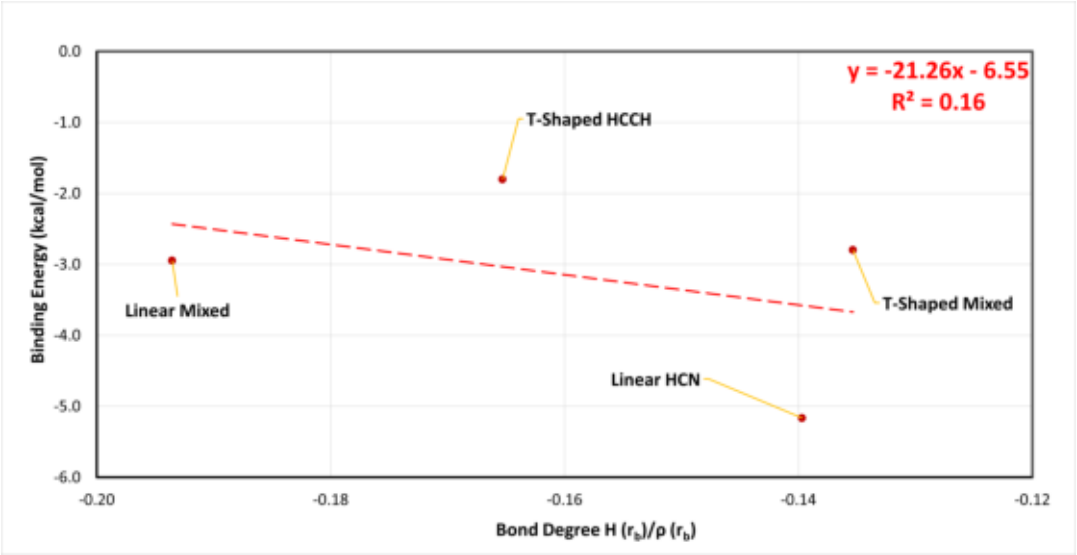


Figure 4.29: Plot of binding energy vs energy density per electron (bond degree) on Bond Critical Point of each dimer.

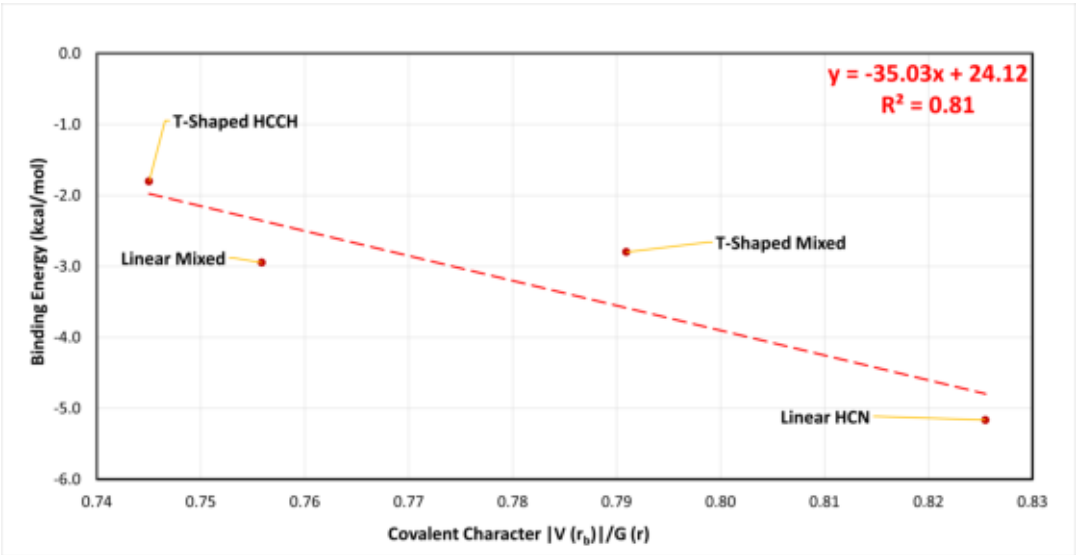


Figure 4.30: Plot of binding energy vs covalent character on Bond Critical Point of each dimer.

4.6 Mean Absolute Percentage Error (MAPE) comparison

In Tab. 4.9 are listed the MAPE% values calculated from the benchmark performed on the four dimers. Most of the methods used perform well in the calculation of the Van der Waals bond length. As far as Grimme's semi-parameterized methods are concerned, due to their low resources cost, the B97-3c proves to be the most accurate one to implement in larger system calculations.

Table 4.9 Mean Absolute Percentage Error (MAPE%) of the different methods used in the benchmark.

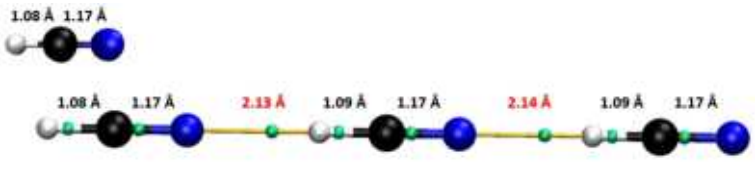
Functional	**Basis set**	**Bond length MAPE%**
fc-CCSD(T)	aug-cc-pVTZ	1.1
B2PLYP-D3BJ	aug-cc-pVTZ	**1.2**
	jun-cc-pVTZ	1.3
	may-cc-pVTZ	1.3
	maug-cc-pVTZ(-dH)	1.2
B3LYP-D3BJ	aug-cc-pVTZ	1.1
	jun-cc-pVTZ	1.1
	may-cc-pVTZ	1.1
	maug-cc-pVTZ(-dH)	1.1
	SNSD	**1.1**
M06-2X	aug-cc-pVTZ	4.2
	jun-cc-pVTZ	3.7
	may-cc-pVTZ	3.5
	maug-cc-pVTZ(-dH)	**3.3**
xTB	STO-G	2.0
PBEh-3c	def2-TZVP	3.9
HF-3c		5.1
B97-3c		**2.7**
DSD-PBEP86-D3BJ	def2-TZVPP	1.4
revDSD-PBEP86-D3BJ		1.6
CAM-B3LYP-D3BJ	aug-pc-3	1.0
LC-ωPBE-D3BJ	aug-pc-3	2.4

4.7 $(HCN)_m$ clusters series

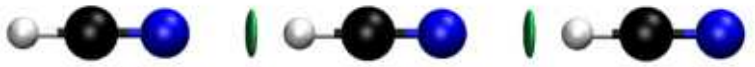
4.7.1 Linear clusters

Hydrogen cyanide is a highly polarized molecule which in the gaseous phase can establish many non-covalent interactions. Considering the dimer $(HCN)_2$ only one structure was plausible, namely the linear one, which consists of the interaction between the negative and positive poles between the two dipoles. The in-line arrangement of molecules allows for the best interaction, it also induces the polarization of the units that are chained, favouring the in-line arrangement of additional molecules.

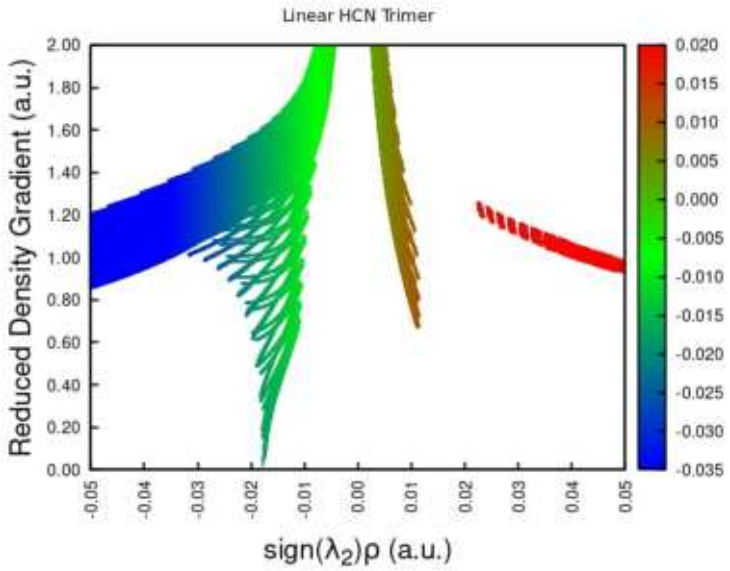
The series of linear clusters propagates up to eight HCN units. Considering the relatively small size of the system under consideration, it was thought to implement revDSD-PBEP86 level calculations with basis set aug-cc-pVDZ. The calculations led to bond lengths similar to those of monomers (1.07 *Å* for H–C bond and 1.15 *Å* for C≡N bond result from the semi-empirical method, the results from the DSD-type functional are displayed in Fig. 4.31), except for the unit in the middle of the structure. Due to intermolecular interactions that cause a slight extension of the H-C bond and a slight shortening of the C≡N bond. The intermolecular length value calculated at both levels is about 2.13-2.14 *Å*. In Fig. 4.31a are reported only the geometry parameters calculated at revDSD-PBEP86/aug-cc-pVDZ level, retained as the most accurate hybrid DFT. Despite the different degrees of freedom of the structure, the calculations show an angle of 180° between the molecules. The results of revDSD-PBEP86/aug-cc-pVDZ calculations determined binding energy value of -48.2 kJ/-mol (-11.5 kcal/mol). After zero-point energy correction the values passed to a higher value of -39.7 kJ/mol (-9.5 kcal/mol). Since with larger clusters it will not be possible to continue calculations with high-level methods, calculations have been performed with the semi-empirical B97-3c/def2-TZVP method for comparison with the following calculations. The intramolecular bond lengths are in line with other calculation methods, 1.07 *Å* for the H-C bond and 1.15 *Å* for the C≡N bond. Intermolecular distances range from 2.18 *Å* to 2.19 *Å*. The binding energy calculated at this level is -41.8 kJ/mol (-10 kcal/mol), -34.6 kJ/mol (-8.3 kcal/mol) with the ZPVE correction. Unfortunately, it is not possible to compare the results obtained for trimers either because of the absence of experimental data or because of the impracticability of accurate calculation methods. The fact remains that the results are quite consistent with what we assume about these systems. QTAIM analysis reveals the intermolecular interactions between each molecule, illustrated by the green BCPs and RDG isosurfaces. The RDG vs $\text{sign}(\lambda_2) \cdot \rho$ plot identified only one type of Van der Waals force at -0.02 corresponding to hydrogen bonds.



(a)



(b)



(c)

Figure 4.31: QTAIM analysis, RDG isosurfaces (RDG($\boldsymbol{\rho}$) = 0.5) and Reduced Density Gradient vs sign(λ_2)·ρ plot of the HCN linear trimer.

The series continues up to four to eight hydrogen cyanide molecules aligned to give linear structures. The linear tetramer was optimized at B97-3c/def2-TZVP level, which leads to bond length average values of 1.08 *Å* for H–C bond and 1.15 *Å* for C≡N bond, consistent with the smaller linear clusters. In this structure, there are three intermolecular bonds between nitrogen and hydrogen. Meanwhile, the hydrogen bonds formed by the external molecules are about 2.16 *Å*, the central interaction is quite shorter (2.11 *Å*). This is probably due to a compression of the electronic cloud by the external molecules on the internal ones, as shown in Fig. 4.32.

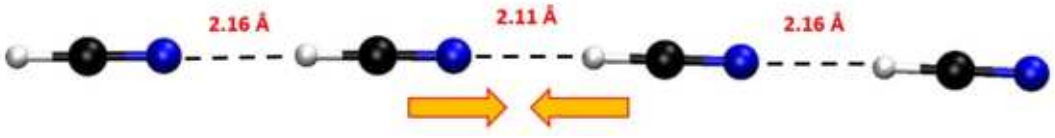


Figure 4.32: HCN linear tetramer contraction due to external molecules.

The contraction of intermolecular distances between the molecules arranged internally, also recurs in clusters with more than four units. The consolidation of internal interactions explains the stability of this structure. Not only is dipole-dipole interaction very effective, but with each added hydrogen cyanide molecule, the non-covalent bonds become stronger. All structures with more than five units have been optimized only at level B97-3c/def2-TZVP,

due to the increase of nuclei and consequently the calculation times. Moreover, for linear structures starting from the hexamer, it was not possible to continue with the calculation of harmonic frequencies. Due to the high rotational motions of linear structures, the calculations have not reached convergence. Therefore, binding energy could not be corrected for zero-point vibrational energy. Linear structures maintain an angle of 180° up to eight molecules aligned. The lengths of intramolecular bonds remain on average similar to those of the monomer, with slight variations for molecules that interact on both sides. Images of linear structures with more than five units are not reported, because of their size.

The energy contribution for each HCN molecule addition in the linear structure was analyzed. In Tab. 4.10 are reported the energy values of each cluster and the contributions for each HCN molecule addition to increasing the linear cluster. The contribution of each monomer is not constant from time to time, as it leads to an extra-stabilization of the complex, probably due to the compression of the central units which intensifies intermolecular interactions. Fig. 4.33 illustrates the energy contribution of linear cluster series.

Table 4.10 Energy contribution per monomer of hydrogen cyanide linear structure cluster.

Cluster	**Energy contribution per monomer (kJ/mol)**
$(HCN)_2$	$-9.3\,(+0.35\%)$
$(HCN)_3$	$-13.9\,(+0.53\%)$
$(HCN)_4$	$-16.7\,(+0.64\%)$
$(HCN)_5$	$-18.5\,(+0.70\%)$
$(HCN)_6$	$-19.8\,(+0.75\%)$
$(HCN)_7$	$-20.7\,(+0.79\%)$
$(HCN)_8$	$-21.4\,(+0.81\%)$

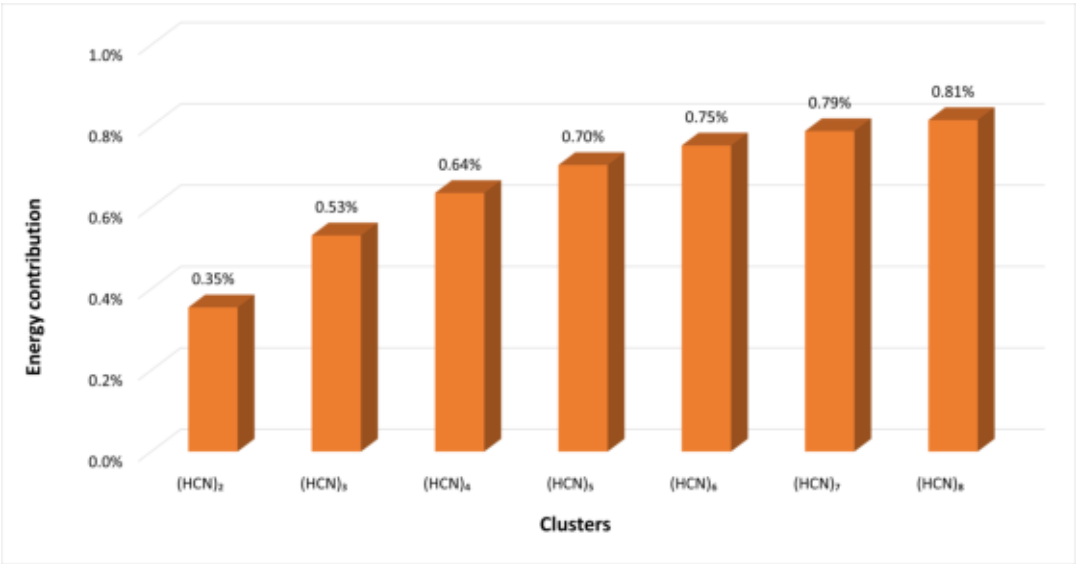
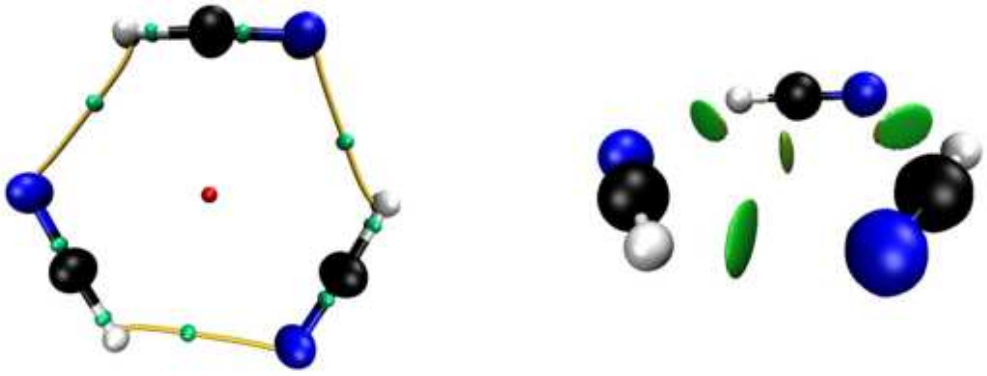


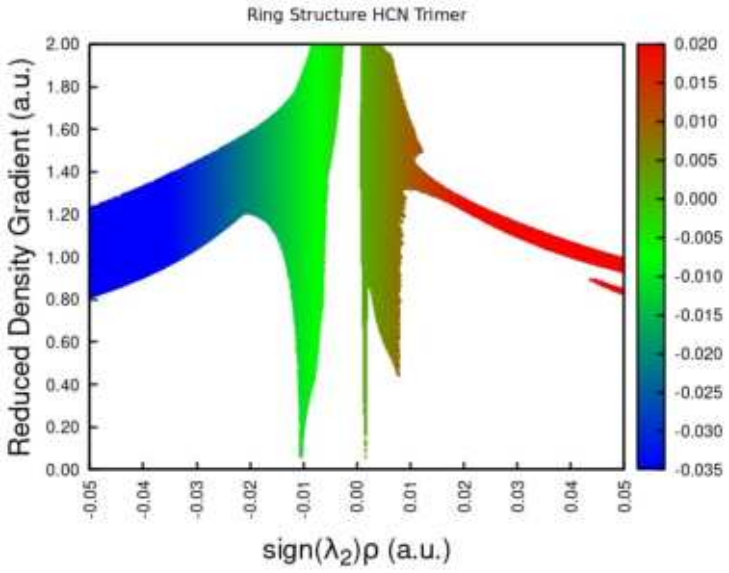
Figure 4.33: Percentage energy contributions histogram for hydrogen cyanide linear clusters.

4.7.2 Ring clusters

Another type of structure identified is the ring one. Three or more hydrogen cyanide molecules concatenate to form a closed structure. The intermolecular interaction is not as whole as the linear one, the orbitals overlap partially forming an angle different from 180° between one molecule and the other. Calculations at the revDSD-PBEP86/aug-cc-pVDZ level of the circular trimer reported values of intramolecular bond lengths equal to those of the monomers calculated at the same level. In addition, no difference is observed between the molecules within the cluster. Concerning intermolecular bond lengths, the same value has been obtained with the revDSD-PBEP86 method, 2.42 *Å*. The semi-empirical B97-3c method has obtained bond lengths of 1.07 *Å* and 1.15 *Å* for intramolecular bonds H-C and C≡N. 2.51 *Å* as regards the three non-covalent interactions. The internal angles of the structure have variable angles, averaging around 120°, the same value was observed for all structures optimized at different levels of theory. The binding energy values for the ring trimer calculated at revDSD-PBEP86/aug-cc-pVDZ and B97-3c/def2-TZVP are respectively -45.3 kJ/mol (-10.8 kcal/mol) and -42.9 kJ/mol (-10.3 kcal/mol). Zero-point vibrational energy correction led to higher binding energy values reported here for the two methods, -39.0 kJ/mol (-9.3 kcal/mol) and -36.9 kJ/mol (-8.8 kcal/mol). In Fig. 4.34 are shown the topology analysis of the structure Fig. 4.34a, the RDG isosurfaces 4.34b and the RDG plot 4.34c. The topology analysis detected three BCP in the NCI path and one RCP, which confirms the formation of a cyclical structure. As the RDG vs ρ plot shows, the Van der Waals interactions are very distinct and very intense, always due to the high polarization of the hydrogen cyanide molecule.



(a) (b)



(c)

Figure 4.34: QTAIM analysis, RDG isosurfaces (RDG($\boldsymbol{\rho}$) = 0.5) and Reduced Density Gradient vs sign(λ_2)·ρ plot of the HCN ring trimer.

The B97-3c calculations identified ring structures of up to six HCN molecules. Topological analysis shows that $(HCN)_7$ and $(HCN)_8$ ring clusters have no critical ring points at the centre of the structure. More than six units placed in a circle, despite establishing non-covalent interactions, possess too many degrees of freedom to form a stable ring. The B97-3c calculations reported intramolecular bond length values, for all the clusters up to three HCN molecules, equal to 1.08 $\AA$ for H–C bond and 1.15 $\AA$ for C≡N bond. As HCN molecules are added, intermolecular distances gradually decrease. Starting from the trimer that for the calculations B97-3c/def2-TZVP has a bond distance equal to 2.51 $\AA$, we have that the intermolecular distance inside the octamer is equal to 2.08 $\AA$. This led to suppose that the cluster may continue to exist up to the cluster $(HCN)_8$, but after six units the RCP is no longer visible in topological analysis.

The energy contribution for each HCN molecule addition in the ring structure was analyzed. In Tab. 4.11 are reported the energy values of each cluster and the contributions for each HCN molecule addition to increasing the circular cluster. The contribution of each

monomer is not constant from time to time, as it leads to an extra-stabilization of the complex. Figure 4.35 illustrates the energy contribution of the ring trimer.

Table 4.11 Energy contribution per monomer of hydrogen cyanide ring structure cluster.

Cluster	Energy contribution per monomer (kJ/mol)
$(HCN)_3$	$-14.3\,(+0.54\%)$
$(HCN)_4$	$-18.8\,(+0.72\%)$
$(HCN)_5$	$-21.3\,(+0.81\%)$
$(HCN)_6$	$-22.8\,(+0.87\%)$
$(HCN)_7$	$-23.8\,(+0.91\%)$
$(HCN)_8$	$-24.4\,(+0.93\%)$

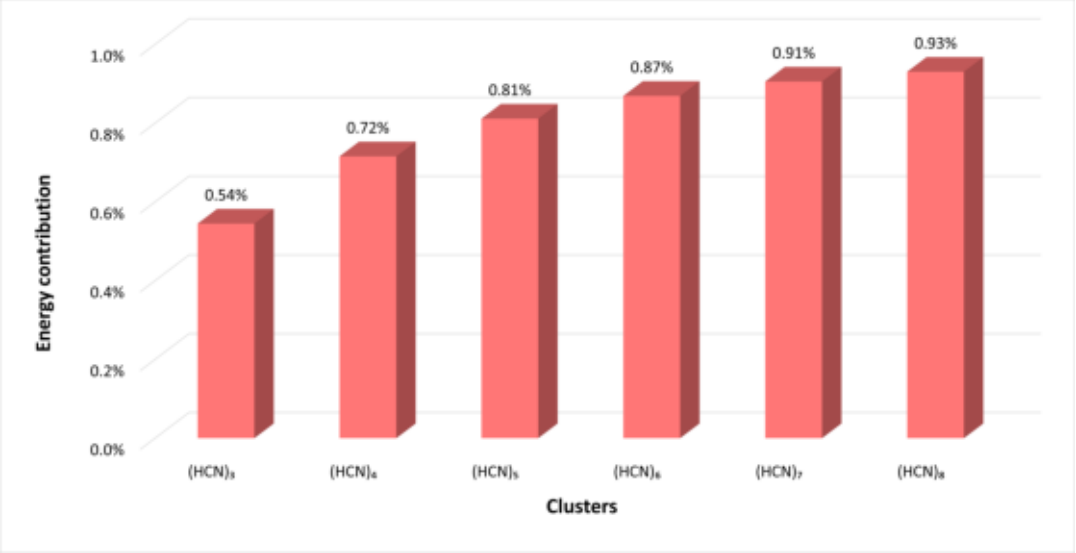


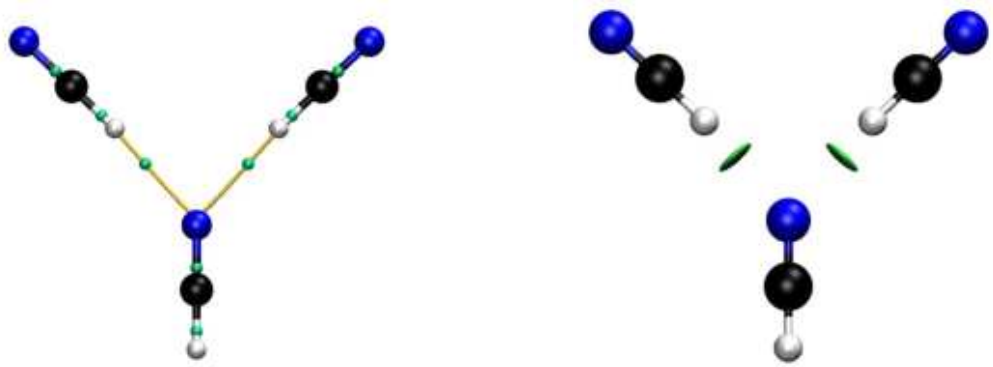
Figure 4.35: Percentage energy contributions histogram for hydrogen cyanide ring structure clusters.

4.7.3 Flux Capacitor clusters

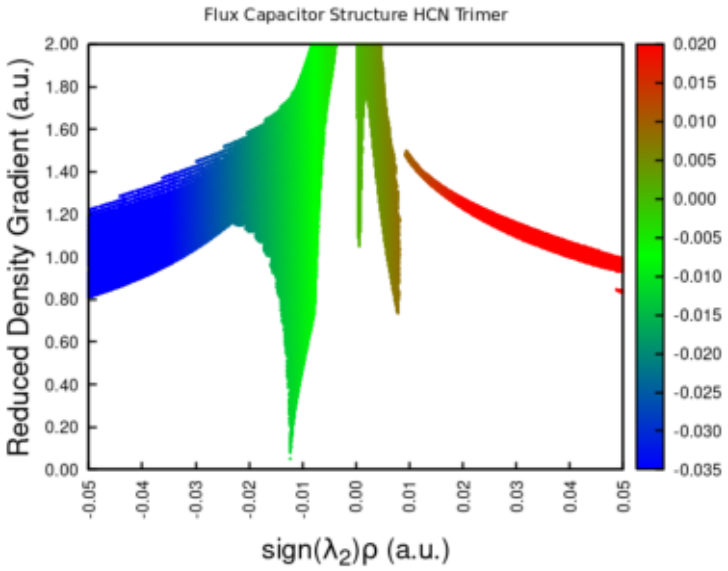
During the optimization of the trimers has been characterized a structure from we called *"Flux Capacitor"*, name inspired by the movie "Back to the Future", written by Robert Zemeckis and Michael Robert Gale. In Fig. 4.36 is shown the flux capacitor from the movie.

Figure 4.36: *"That was the day I invented time-travel. I remember it vividly. I was standing on the edge of my toilet hanging a clock, the porcelain was wet, I slipped, hit my head on the sink, and when I came to I had a revelation! A vision! A picture in my head! A picture of this! This is what makes time travel possible: the flux capacitor!"* – Dr. Emmett Brown.

From the calculations carried out on the other HCN clusters this type of structure has not been presented anymore. So to determine whether the Flux Capacitor structure was stable even with four or more units, it was decided to manually build clusters and optimize them at the B97-3c/def2-TZVP level. The calculations have obtained the structure display in Fig. 4.37 in which the lengths of the intramolecular bonds are equivalent to those of the monomer structure (1.07 $\AA$ H–C and 1.15 $\AA$ C≡N). This means that non-covalent interactions are not particularly intense. Despite this, the length of the non-covalent bonds is less than those found in the ring trimer, with a value of 2.35 $\AA$. The internal angle between the two HCN molecules is about 84°. The binding energy value for this cluster is -25.2 kJ/mol (-6.0 kcal/mol), -20.5 kJ/mol (-4.9 kcal/mol) considering the ZPVE correction.



(a) (b)



(c)

Figure 4.37: QTAIM analysis, RDG isosurfaces (RDG($\boldsymbol{\rho}$) = 0.5) and Reduced Density Gradient vs sign(λ_2)·ρ plot of the HCN flux capacitor trimer.

The tetramer and the pentamer were constructed by adding one and two molecules of HCN to the tail of the structure, in which there is the hydrogen atom ready to accept electrons of the nitrogen atom. A structure with six HCN molecules has been also devised by adding a molecule to each docking point of the trimer as shown in Fig. 4.38, which display the geometric parameter values calculated at B97-3c level. As can be seen, the lengths of the intramolecular bonds are similar to those of the isolated monomers, except for the HCN molecules inside the structure which have a slight elongation of the H–C bond. The intermolecular bond lengths of linear forming molecules are slightly lower than those measured in linear HCN clusters, probably because of the double interaction of the nitrogen atom, which polarizes a lot the molecule. As for intermolecular bonds of non-aligned molecules, their length reaches values about 2.30 $\AA$. The binding energies ZPVE corrected are -43.7 kJ/mol (-10.4 kcal/mol) for the flux capacitor tetramer, -66.4 kJ/mol (-15.9 kcal/mol) for the flux capacitor pentamer and -79.8 kJ/mol (-19.1 kcal/mol) for the analogue hexamer. These many structures are shown in Fig. 4.38.

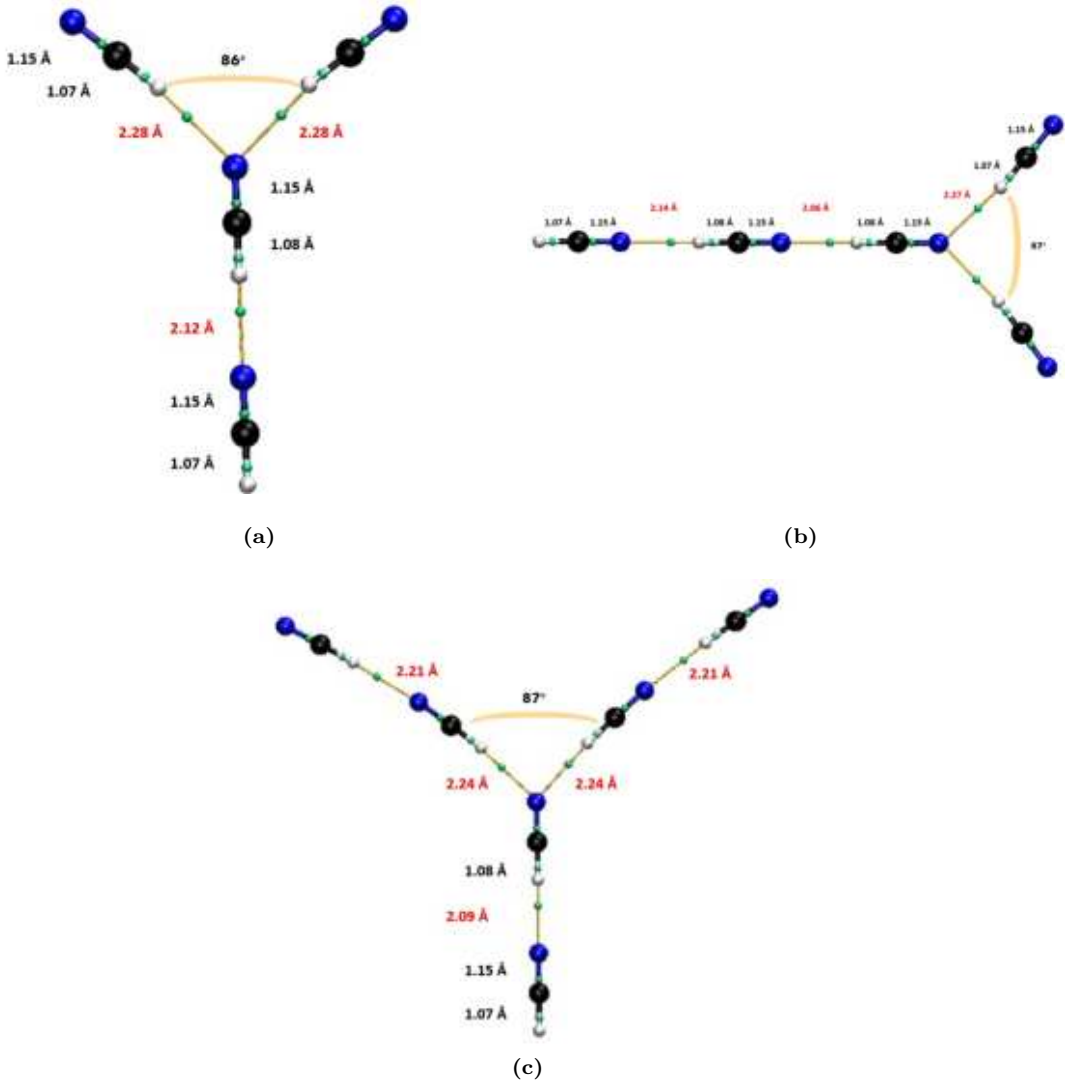


Figure 4.38: Flux capacitor structures build by adding hydrogen cyanide molecules to the flux capacitor trimer.

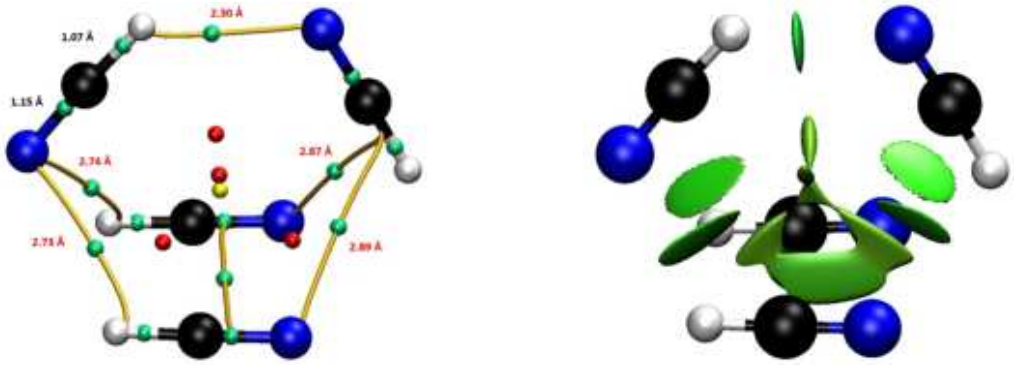
Unfortunately, calculations for flux capacitor $(HCN)_7$ and $(HCN)_8$ structures have not reached convergence and it has not been possible to optimize them.

4.7.4 Tamar Stein tetramer

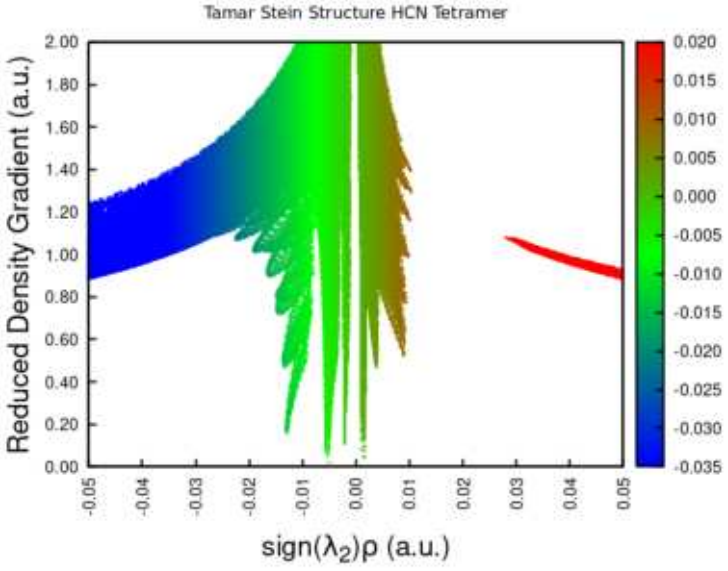
In 2020, Stein T. and *et al.* [18] studied clusters of acetylene and hydrogen cyanide. In their work, they identified some types of structures also present in this thesis work, which we also identified and presented in this thesis. The only exception is the tetramer of HCN that we decided to optimize to verify its stability compared to other optimized structures. This tetramer was given the name of *"Tamar Stein Structure"* as it was already studied by the group of Dr. Tamar Stein. The structure has been optimized at B97-3c/def2-TZVP level, the frequencies calculations have resulted in a binding energy ZPVE corrected value of -47.5 kJ/mol (-11.3 kcal/mol). Less stable than ring or linear structures, but more stable than the flux capacitor tetramer.

Figure 4.39a shows the geometric values of the tetramer and the different critical points as results from the QTAIM analysis. The four HCN molecules form a three-dimensional structure identified by the CCP at the center. In addition, four "faces" identified by the four RCPs are also visible. The lengths of intramolecular bonds have the same values as the isolated monomer and do not appear to be affected by the presence of NCIs. Analysis of the RDG

isosurface shows the numerous interactions between all four HCN molecules. The electronic doublet of the nitrogen atom above interacts with two hydrogen atoms. Vice versa, on the other side. The two hydrogen cyanide molecules placed side by side below repel slightly due to their polarization. For that reason, the intermolecular distances between the two molecules below and the two above are greater than normal. The figure 4.39 are illustrated the topology analysis of the structure Fig. 4.39a, the RDG isosurfaces 4.39b and the RDG plot 4.39c.



(a) **(b)**



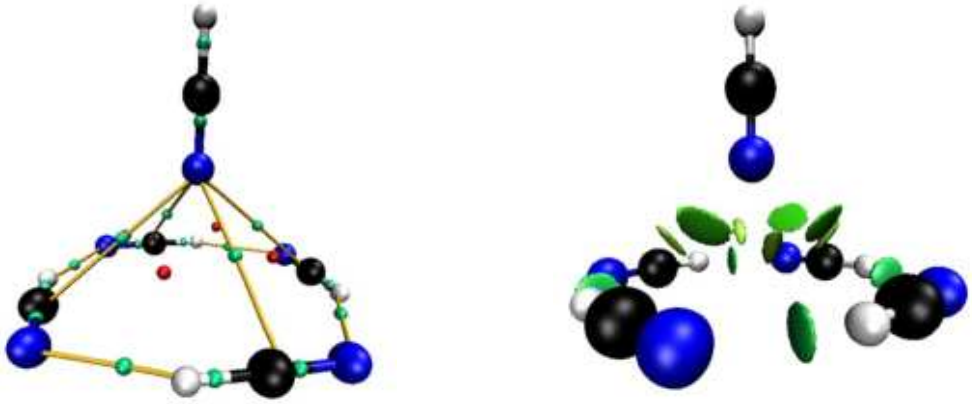
(c)

Figure 4.39: QTAIM analysis, RDG isosurfaces (RDG($\boldsymbol{\rho}$) = 0.5) and Reduced Density Gradient vs sign(λ_2)·ρ plot of the HCN Tamar Stein tetramer.

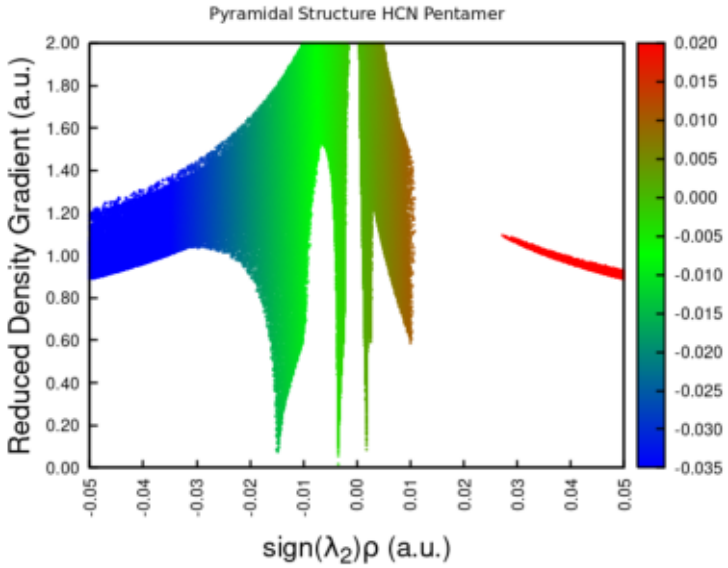
4.7.5 Pyramidal clusters

From the xTB calculations several structures have been obtained, which where subsequently optimized at B97-3c level; some of them have a ring with four HCN molecules and one or more molecules aligned with the centre of the ring. The ring structure is identical to the one discussed above and all lengths of intramolecular bonds are equal to those of the isolated monomer. The distance between the ring centre and the HCN molecule is equal to 2.16 $\AA$. The binding energy calculated for this structure is equal to -75 kJ/mol (-17.9 kcal/mol),

such stability is possible to explain it through the figures 4.40 illustrating the numerous interactions of the HCN molecule with the ring units. Fig. 4.40c detected three Van der Waals interactions, one of which pretty intense at -0.015 of electron density, generated from the interaction of the single HCN molecule with the ring structure.



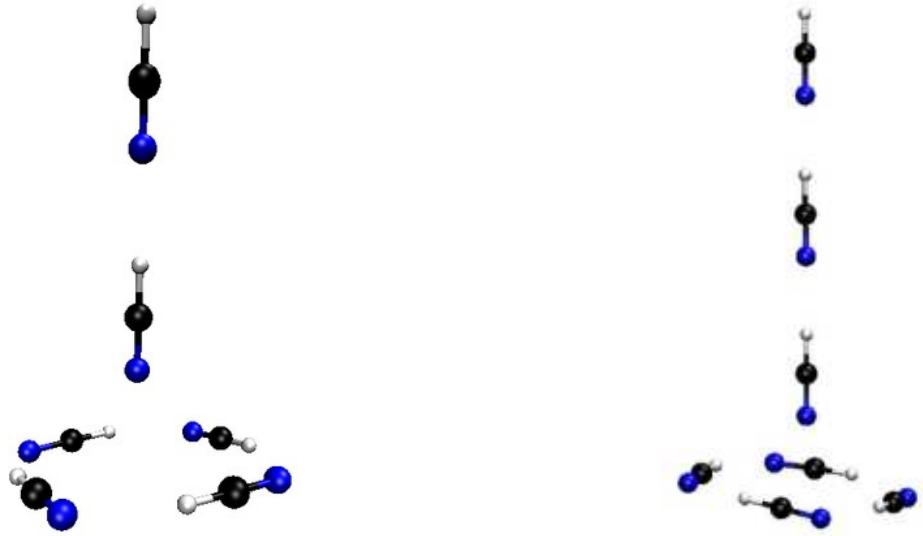
(a) **(b)**



(c)

Figure 4.40: QTAIM analysis, RDG isosurfaces (RDG($\boldsymbol{\rho}$) = 0.5) and Reduced Density Gradient vs sign(λ_2)·ρ plot of the HCN pyramidal pentamer.

This type of structure was also identified during the study of hexamers and heptamers and are shown in Fig. 4.41. The union between two of the most stable structures, the ring one and the linear one, gives rise to a regular and stable cluster pyramidal-shaped with one, two or three HCN aligned with the centre of the ring. The same structure with eight HCN molecules has not been observed. In addition, the ring with four HCN molecules seems to be the most stable one for the pyramidal structure. Five molecules (or more) arranged in a ring are too far away to interact well with another external from the centre.



(a) (b)

Figure 4.41: Pyramidal hexamer and heptamer cluster.

4.7.6 Clusters energy comparison

The figure 4.42 compares the binding energy of the different hydrogen cyanide clusters studied in this work. The series of ring clusters is the most stable. Initially, the energy of the ring trimer is similar to the analogue of the linear series, but any addition of monomer leads to greater stabilization of the ring complex compared to the linear structure, as was also noted in Tab. 4.11. The second place is given at the linear clusters thanks to the high polarization of HCN units. The greater stability of ring complexes can be explained by the fact that all the components of the cluster interact in the same way and with the same intensity. The series of pyramidal clusters with a base of four HCN molecules ring arranged, turns out to be at the third place regarding the stability of the complexes. Meanwhile, the flux capacitor series results as the least stable. Concerning the $(HCN)_4$ clusters, the Tamar Stein structure is just below the analogue of the flux capacitor series. The reported binding energy does not present the ZPVE correction, as not all clusters have reached convergence in the calculation of harmonic frequencies.

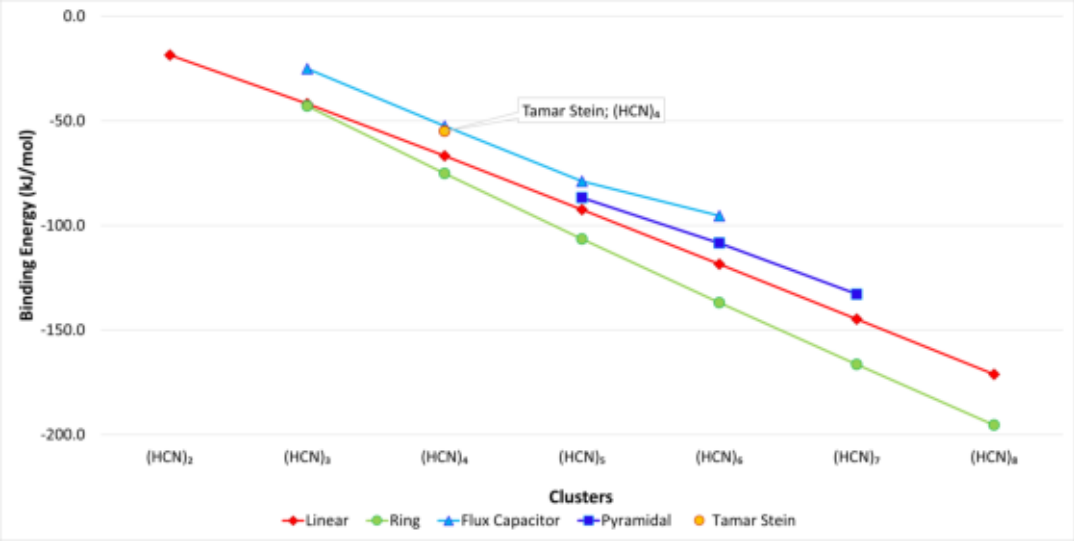
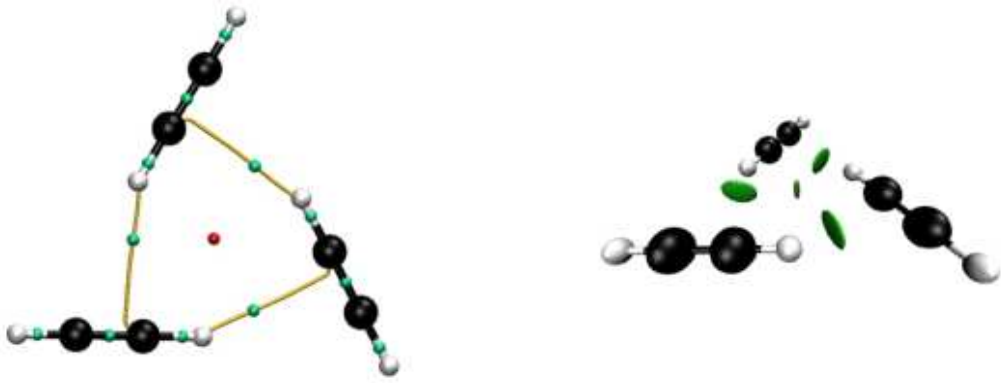


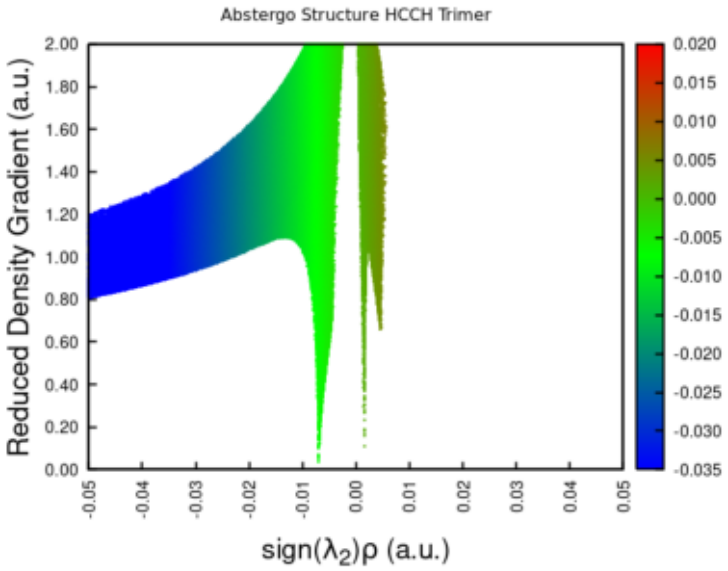
Figure 4.42: Binding energy comparison between the different $(HCN)_m$ clusters.

4.8 $(H_2C_2)_n$ clusters series

The $(H_2C_2)_n$ series of clusters has a smaller variety of structures than the $(HCN)_m$ series, therefore this section of the thesis will be analyzed with a different approach than the previous one. This is due to the type of forces involved, whereas in hydrogen cyanide clusters dipole-dipole interactions prevailed, in the acetylene cluster dispersion forces prevailed, which establish weaker induced dipole-induced dipole interactions type. Only trimers have more than one structure. The most stable is the structure called *"Abstergo"*, i.e. three ring-arranged acetylene molecules, since the intermolecular interaction takes place between hydrogen and the electron system π at the centre of the molecule, the structure is staggered and takes the shAPE% shown in the figure 4.43. The calculation of the geometric parameters gives values of bond lengths equal to those of the isolated monomer optimized at the same level. For both B2PLYP/may-cc-pVTZ(-dH) and B97-3c/def2-TZVP calculations, the H-C bonds length is 1.06 *Å*, while the C≡C bond is 1.20 *Å*. Calculations at revDSD-PBEP86/aug-cc-pVDZ gave slightly larger values, for the precision 1.08 *Å* for the H-C bond and 1.22 *Å* for the C≡C bond, similar lengths to the monomer optimized at the same level. The intermolecular bond length is approximately 2.70 *Å* for all computational methods used. The binding energy values for the B2PLYP/may-cc-pVTZ(-dH), the revDSD-PBEP86/aug-cc-pVDZ and the B97-3c/def2-TZVP methods are -19.8 kJ/mol (-4.7 kcal/mol), -17.3 kJ/mol (-4.1 kcal/-mol) and -21.0 kJ/mol (-5.0 kcal/mol) respectively. The ZPVE correction leads to upper BE values, specifically the B2PLYP ZPVE BE is about -2.2 kJ/mol (-0.5 kcal/mol), -11.4 kJ/mol (-2.7 kcal/mol) for the revDSD-PBEP86 functional and -16.3 kJ/mol (-3.9 kcal/mol) for the B97-3c functional. The QTAIM analysis shows the presence of an RCP and the interactions between the hydrogen atoms and the π system. Fig. 4.43c shows two types of interactions, the steric repulsion at the centre and the Van der Waals interactions between the molecules. The RDG plot is quite similar to the one for the ring $(HCN)_3$ cluster, with the difference that in this case, the VdW signal falls above -0.01 a.u. of electron density, indicating a weaker interaction.



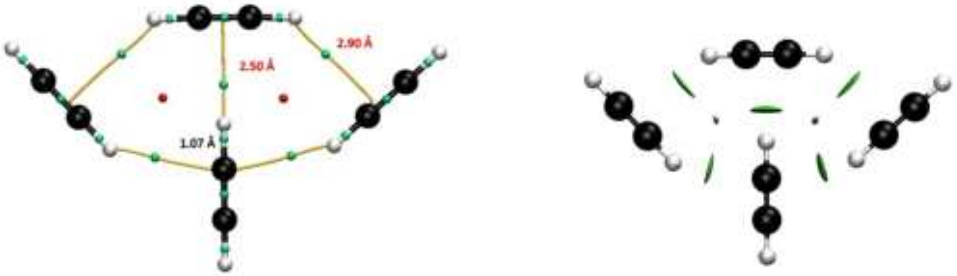
(a) (b)



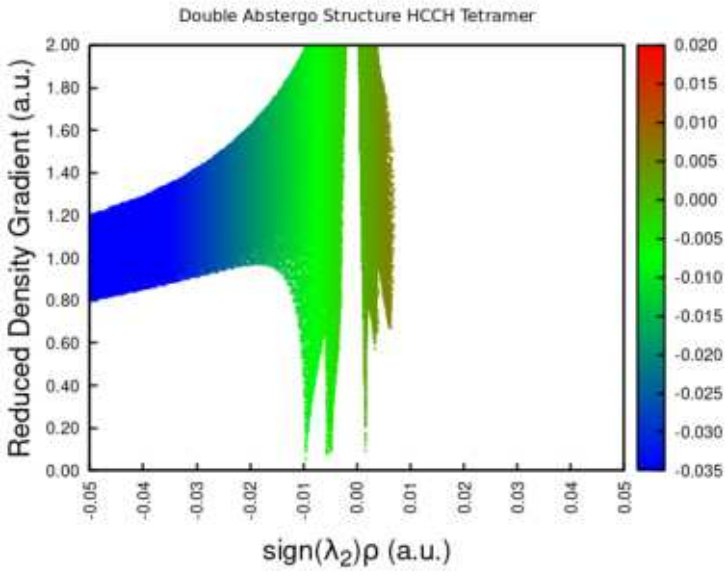
(c)

Figure 4.43: QTAIM analysis, RDG isosurfaces (RDG($\boldsymbol{\rho}$) = 0.5) and Reduced Density Gradient vs sign(λ_2)·ρ plot of the H_2C_2 abstergo trimer.

This structure recurs in the other more complex clusters, probably because it is the most stable space arrangement that three acetylene molecules can have. Two other trimer structures have been optimized, but are less stable than the abstergo cluster. Fig. 4.44 shows the acetylene tetramer, calculated at B97-3c/def2-TZVP level, which consists of a planar structure where the four molecules form a double abstergo structure. The topology analysis detected two RCPs which confirm the formation of two rings. The central molecules are the ones with more non-covalent interactions so their intramolecular bonds are quite perturbed. In particular, the C–H bond highlighted in fig. 4.44a, is slightly longer than the isolated monomer. The two central acetylene molecules have a stronger interaction, the intermolecular bond length is about 2.50 *Å*, while the other NCI lengths are almost 3.00 *Å*. The ZPVE corrected binding energy for this structure is -24.2 kJ/mol (-5.8 kcal/mol).



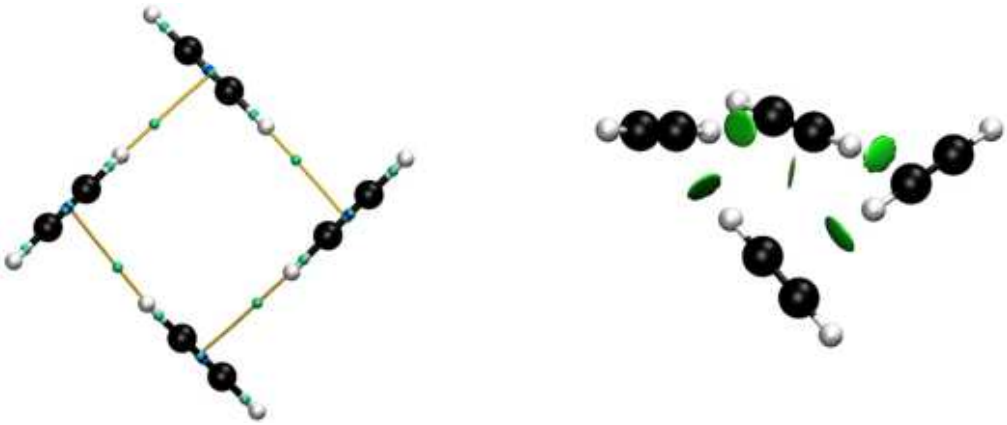
(a) **(b)**



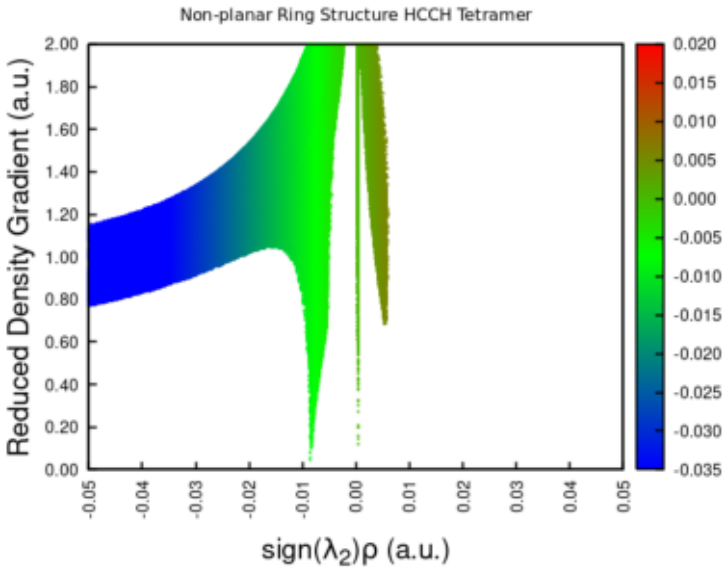
(c)

Figure 4.44: QTAIM analysis, RDG isosurfaces (RDG($\boldsymbol{\rho}$) = 0.5) and Reduced Density Gradient vs sign(λ_2)·ρ plot of the H_2C_2 double abstergo tetramer.

A cyclic tetramer was also found during the B97-3c optimization and it is isoenergetic to the double abstergo one. This structure consists in four acetylene molecules that interact to form a non-planar square. The intramolecular bonds have the same length as the monomer one, except for the C–H bond which interacts with the π-electrons of another acetylene, which is a bit longer. The intermolecular length is about 2.60 $\AA$ for all the four NCIs. As shown in Fig. 4.45a there is not a ring critical point in the structure centre, that because the structure is not planar. However, the RDG analysis detects a thin isosurface within the structure, due to a small steric effect. This structure will also be repeated in some more complex clusters, such as $(H_2C_2)_6$ and $(H_2C_2)_7$.



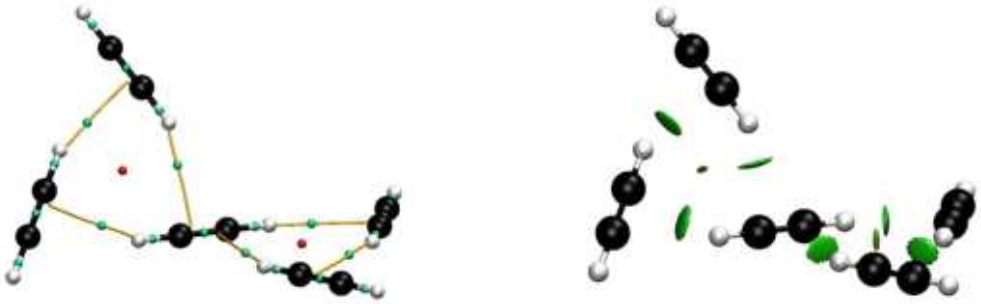
(a) (b)



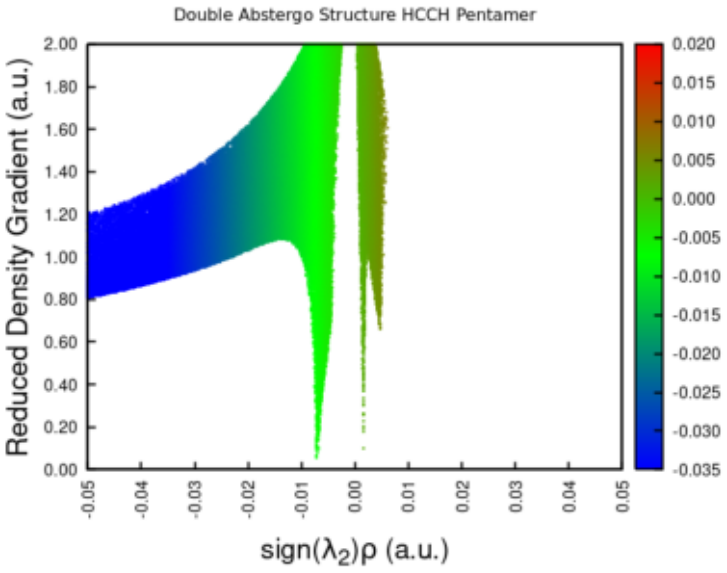
(c)

Figure 4.45: QTAIM analysis, RDG isosurfaces (RDG($\boldsymbol{\rho}$) = 0.5) and Reduced Density Gradient vs sign(λ_2)·ρ plot of the H_2C_2 non-planar square tetramer.

Among the pentamers, only one structure has been identified, consisting of two abstergo structures out of the plan on the opposite side, which share one unit of acetylene. The structure was called *"Butterfly"* due to the two abstergo structures which looks like the wings and the central shared acetylene molecule, which looks like the bug body. The geometric parameters are the same as the abstergo structure itself, the binding energy value is twice the ones for the abstergo structure -42.6 kJ/mol (-10.2 kcal/mol) and -33.3 kJ/mol (-7.9 kcal/mol) the ZPVE corrected value.



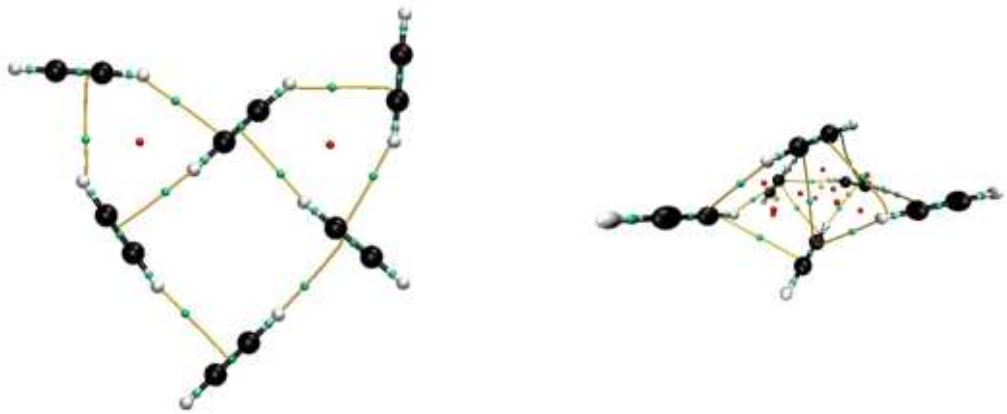
(a) (b)



(c)

Figure 4.46: QTAIM analysis, RDG isosurfaces (RDG($\boldsymbol{\rho}$) = 0.5) and Reduced Density Gradient vs sign(λ_2)·ρ plot of the H_2C_2 butterfly pentamer.

In the case of $(H_2C_2)_6$ cluster, we found two structures. One planar, which consists in two abstergo type structure and one four terms cyclic structure condensed together, and one 3D structure without a regular pattern. It consists of three cages formed by the six acetylene molecules and in numerous NCIs. We notice in the planar hexamer, that the cyclic structure is on the plan established by two abstergo structures, unlike the tetramer analogue. This is probably due to better optimization of hexamer clusters. About the planar hexamer optimized, the values of bond lengths are quite similar to the analogue structures described previously. This structure has a binding energy of -55.5 kJ/mol (-13.3 kcal/mol) and -44.1 kJ/mol (-10.5 kcal/mol) for the ZPVE corrected value. The tridimensional structure has the same intramolecular bond lengths values of the isolated monomer, and there are many different NCIs between the molecules which goes from 2 $\AA$ to almost 4 $\AA$. This structure results in a little bit more stable than the planar ones, -59.7 kJ/mol (-14.3 kcal/mol) and -47.2 kJ/mol (-11.3 kcal/mol) for the ZPVE corrected value.



(a) **(b)**

Figure 4.47: QTAIM analysis and RDG isosurfaces (RDG($\boldsymbol{\rho}$) = 0.5) analysis of the $(H_2C_2)_6$ clusters.

The study of seven acetylene molecules leads to only one optimized structure. This one is formed by one abstergo structure and two square type structures condensed together. Abstergo and one of the square structures are arranged on the plan, while the other square structure is out of the plan. The bond lengths are quite similar to the other structures described previously. The binding energy value for are about -67.5 kJ/mol (-16.1 kcal/mol) and -53.3 kJ/mol (-12.8 kcal/mol) for the ZPVE corrected.

(a) **(b)**

Figure 4.48: QTAIM analysis and RDG isosurfaces (RDG($\boldsymbol{\rho}$) = 0.5) analysis of the $(H_2C_2)_7$ cluster.

Finally, the $(H_2C_2)_8$ cluster optimization led to one tridimensional structure, with apparently two abstergo structures staggered one above the other, and two acetylene molecules randomly distributed around as illustrated in Fig. 4.49. The binding energy of this cluster is about -86.9 kJ/mol (-20.8 kcal/mol), while the ZPVE corrected binding energy value is -69.2 kJ/mol (-16.5 kcal/mol).

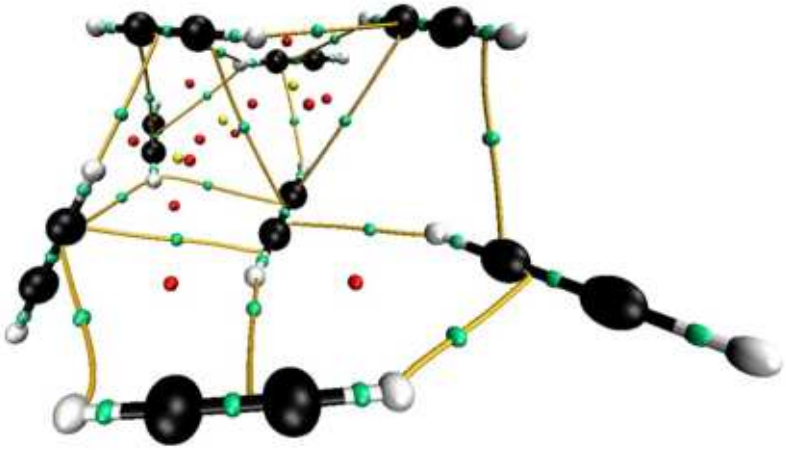


Figure 4.49: QTAIM analysis analysis of the $(H_2C_2)_8$ cluster.

Table 4.12 Energy contribution per monomer of acetylene cluster considering the planar tetramer at Fig. 4.44 and the planar hexamer at Fig. 4.47.

Cluster	**Energy contribution per monomer (kJ/mol)**
HCCH	–
$(HCCH)_2$	−3.5 (+0.13%)
$(HCCH)_3$	−7.0 (+0.27%)
$(HCCH)_4$	−7.7 (+0.29%)
$(HCCH)_5$	−8.5 (+0.32%)
$(HCCH)_6$	−9.9 (+0.38%)
$(HCCH)_7$	−9.6 (+0.37%)
$(HCCH)_8$	−10.9 (+0.41%)

Figure 4.50: Percentage energy contributions histogram for acetylene clusters with Fig. 4.44 tetramer and Fig. 4.47 planar hexamer.

Table 4.13 Energy contribution per monomer of acetylene cluster considering the tetramer at Fig. 4.45 and the planar hexamer at Fig. 4.47.

Cluster	**Energy contribution per monomer (kJ/mol)**
HCCH	–
$(HCCH)_2$	$-3.5\,(+0.13\%)$
$(HCCH)_3$	$-7.0\,(+0.27\%)$
$(HCCH)_4$	$-8.0\,(+0.30\%)$
$(HCCH)_5$	$-8.5\,(+0.32\%)$
$(HCCH)_6$	$-9.9\,(+0.38\%)$
$(HCCH)_7$	$-9.6\,(+0.37\%)$
$(HCCH)_8$	$-10.9\,(+0.41\%)$

Figure 4.51: Percentage energy contributions histogram for acetylene clusters with Fig. 4.45 tetramer and Fig. 4.47 planar hexamer.

Table 4.14 Energy contribution per monomer of acetylene cluster considering the planar tetramer at Fig. 4.44 and the non-planar hexamer at Fig. 4.47.

Cluster	**Energy contribution per monomer (kJ/mol)**
HCCH	–
$(HCCH)_2$	$-3.5\,(+0.13\%)$
$(HCCH)_3$	$-7.0\,(+0.27\%)$
$(HCCH)_4$	$-7.7\,(+0.29\%)$
$(HCCH)_5$	$-8.5\,(+0.32\%)$
$(HCCH)_6$	$-9.3\,(+0.35\%)$
$(HCCH)_7$	$-9.6\,(+0.37\%)$
$(HCCH)_8$	$-10.9\,(+0.41\%)$

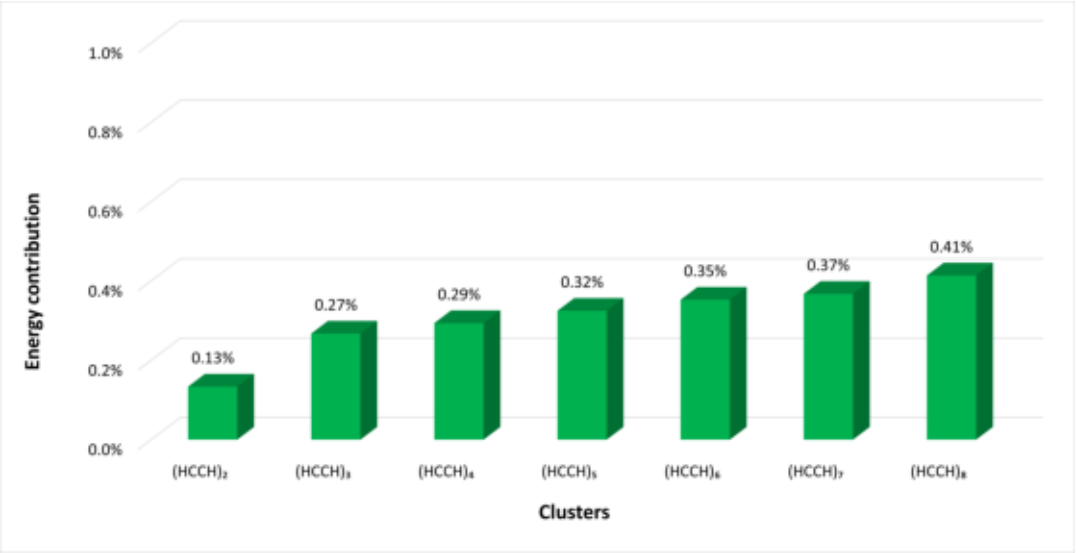


Figure 4.52: Percentage energy contributions histogram for acetylene clusters with Fig. 4.44 tetramer and Fig. 4.47 non-planar hexamer.

Table 4.15 Energy contribution per monomer of acetylene cluster considering the tetramer at Fig. 4.45 and the non-planar hexamer at Fig. 4.47.

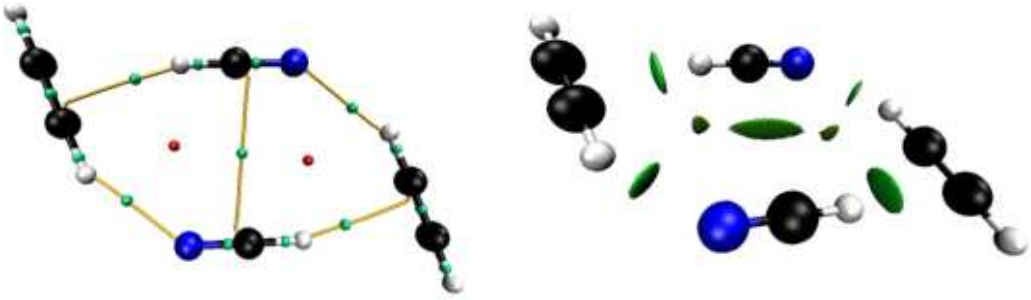
Cluster	**Energy contribution per monomer (kJ/mol)**
HCCH	–
$(HCCH)_2$	$-3.5\,(+0.13\%)$
$(HCCH)_3$	$-7.0\,(+0.27\%)$
$(HCCH)_4$	$-8.0\,(+0.30\%)$
$(HCCH)_5$	$-8.5\,(+0.32\%)$
$(HCCH)_6$	$-9.3\,(+0.35\%)$
$(HCCH)_7$	$-9.6\,(+0.37\%)$
$(HCCH)_8$	$-10.9\,(+0.41\%)$

Figure 4.53: Percentage energy contributions histogram for acetylene clusters with Fig. 4.45 tetramer and Fig. 4.47 non-planar hexamer.

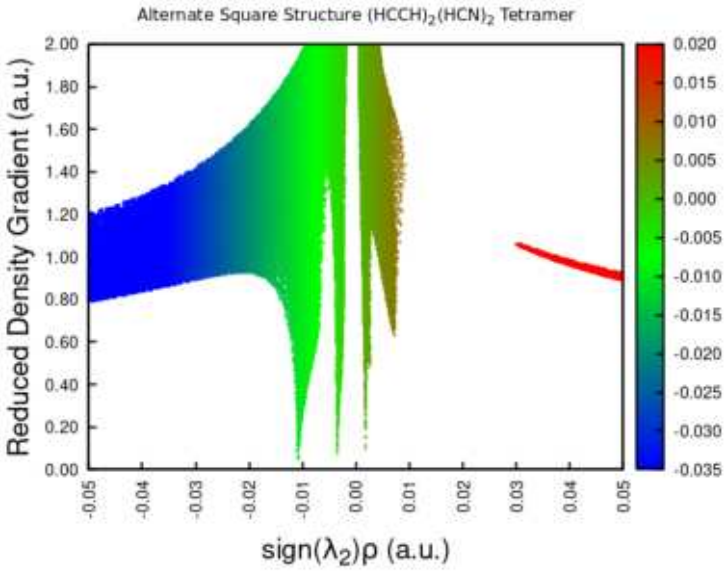
4.9 $(H_2C_2)_2 \cdot (HCN)_2$ clusters series

An optimization at xTB level of $(H_2C_2)_2 \cdot (HCN)_2$ clusters led to three stable structures, successively optimized at B97-3c level. Two circle tetramer were found, one with the two acetylene molecules on the same side and one which alternates the acetylene and hydrogen cyanide molecules in the circle. The third structure consists of two acetylene molecules side-by-side and two HCN molecules, one above and one below the acetylene couple. The first two structures are called *"Square"* and *"Alternate Square"*, while the last one is called *"Parallelepiped"*, they are shown in the figures below.

The alternate square structure is the most stable $(H_2C_2)_2 \cdot (HCN)_2$ tetramer, due to the formation of two rings and a dipole-dipole interaction between the HCN molecules, as illustrated from the isosurfaces between HCN molecules in Fig. 4.54b. The RDG plot versus electron density shows two different VdW interactions, the strong one at -0.01 of electron density is due to hydrogen bonds, the other near 0.00 of electron density indicate the dipole-dipole interaction. The intramolecular bond lengths are equal to the one of isolated monomer, while the intermolecular bond lengths between HCN hydrogen atom and H_2C_2 π-electrons are about 2.50 $\AA$, and between HCN nitrogen atom and H_2C_2 hydrogen, 2.40 $\AA$. The ZPV binding energy of this cluster is -44.3 kJ/mol (-10.6 kcal/mol).



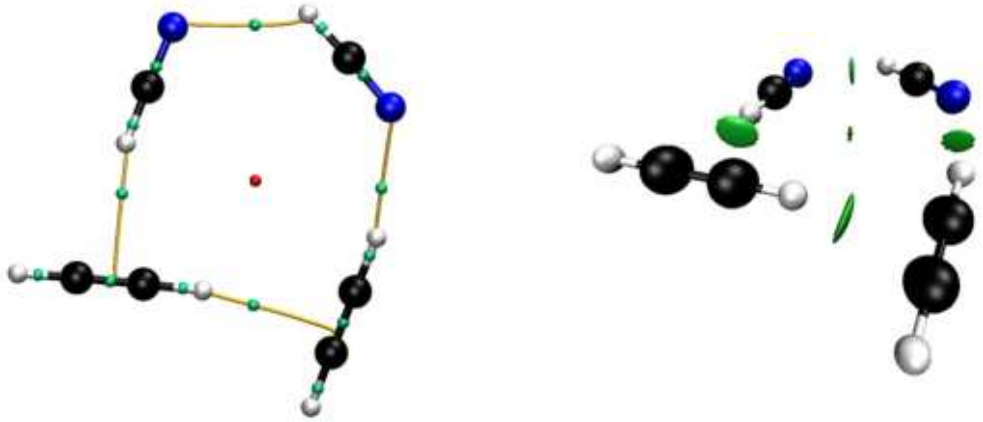
(a) (b)



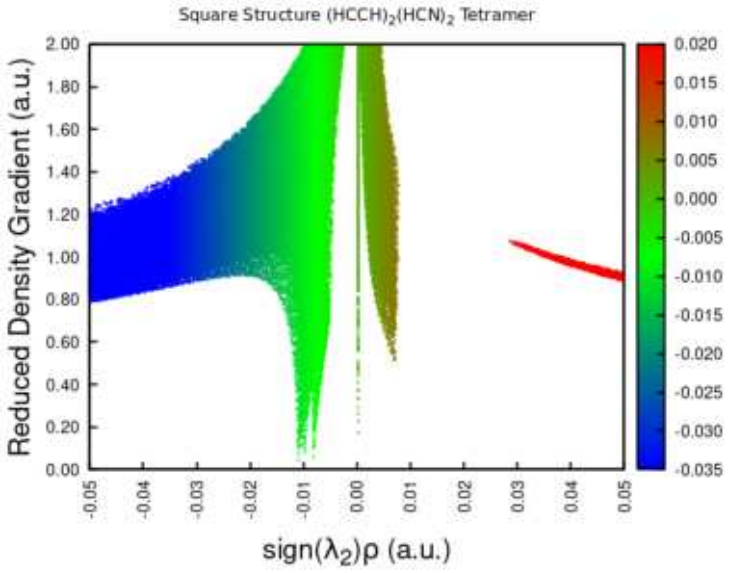
(c)

Figure 4.54: QTAIM analysis, RDG isosurfaces (RDG($\boldsymbol{\rho}$) = 0.5) and Reduced Density Gradient vs sign(λ_2)·ρ plot of the alternate square $(H_2C_2)_2$·$(HCN)_2$ tetramer.

QTAIM analysis detected an RCP and several BCPs in the square structure as shown in Fig. 4.55c there are many Van der Waals interactions and one steric repulsion due to the ring structure. The intramolecular bond lengths are quite similar to the one of isolated monomer, while the intermolecular distance between the two HCN molecules is the shortest (2.43 $\mathring{A}$), the bonds between the HCN and the H_2C_2 molecules have intermediate values of 2.40-2.50 $\mathring{A}$ and the bond between the two acetylene molecules is the longest (2.63 $\mathring{A}$). This is the second more stable structure of these clusters the zero-point binding energy value is about -38.8 kJ/mol (-9.3 kcal/mol).



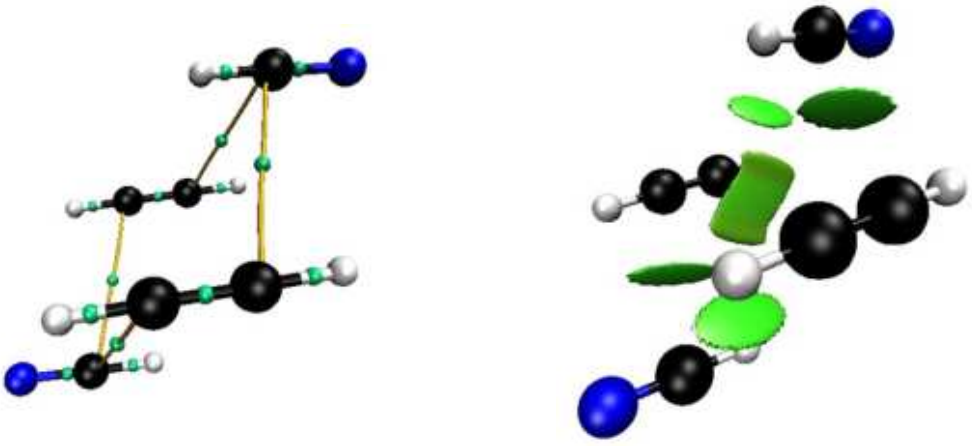
(a) (b)



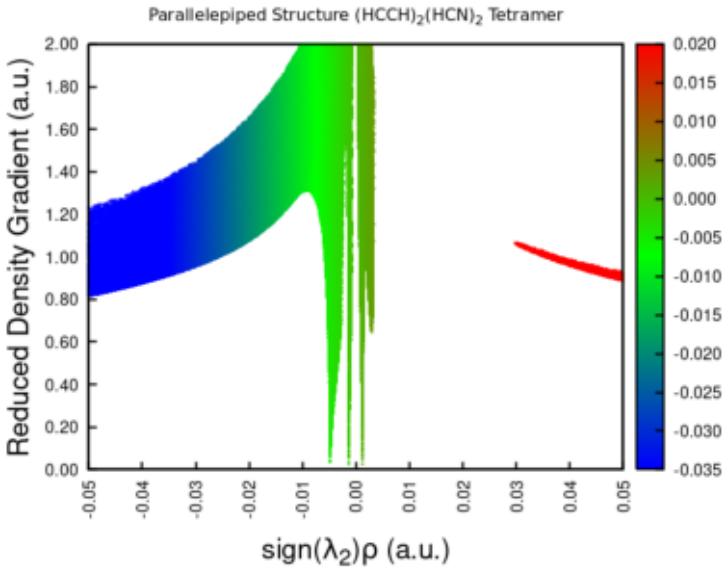
(c)

Figure 4.55: QTAIM analysis, RDG isosurfaces (RDG($\boldsymbol{\rho}$) = 0.5) and Reduced Density Gradient vs sign(λ_2)$\cdot\rho$ plot of the square $(H_2C_2)_2 \cdot (HCN)_2$ tetramer.

The last structure analyzed was the parallelepiped type and as shown in Fig. 4.56a, there are four NCIs between the HCN carbon atoms and the H_2C_2 carbon atoms. In this structure, the dispersion forces are predominant thus is the less stable between $(H_2C_2)_2 \cdot (HCN)_2$ tetramers. The RDG analysis lighted four dipole-induced dipole interactions and one steric repulsion between the two acetylene molecules, as shown in Fig. 4.56c. The intramolecular bond lengths are quite similar to the one of isolated monomer, while intermolecular distances between each molecule are quite similar, 3.33 $\AA$, and almost 5 $\AA$ between the acetylene molecules. The binding energy ZPV corrected value is about -26.4 kJ/mol (-6.3 kcal/mol).



(a) **(b)**



(c)

Figure 4.56: QTAIM analysis, RDG isosurfaces (RDG($\boldsymbol{\rho}$) = 0.5) and Reduced Density Gradient vs sign(λ_2)·ρ plot of the parallelepiped $(H_2C_2)_2$·$(HCN)_2$ tetramer.

4.10 $(H_2C_2)\cdot(HCN)_{m-1}$ clusters series

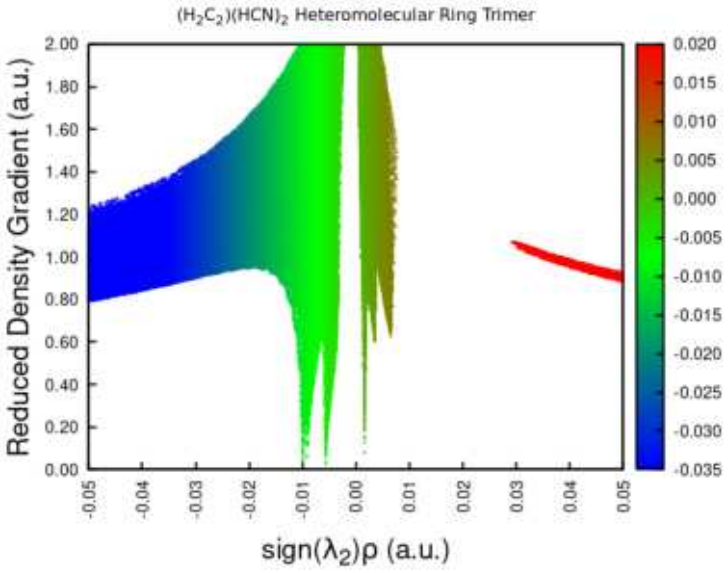
Following the results obtained from the optimization of homomolecular clusters and $(H_2C_2)_2\cdot(HCN)_2$, it was decided to try to optimize the type $(H_2C_2)\cdot(HCN)_{m-1}$, where m indicates the number of units present in the cluster. This study aims to observe how different intermolecular interactions collaborate to form stable Van der Waals clusters. The structures $(H_2C_2)\cdot(HCN)_{m-1}$ resulting from the optimization demonstrate the existence of several interactions between the single acetylene molecule and the surrounding hydrogen cyanide molecules. In particular, the acetylene molecule interacts through the triple bond or polarizes to give dipole-dipole interaction.

4.10.1 Ring clusters

The xTB algorithm generate many structures of the mixed trimer $(H_2C_2)\cdot(HCN)_2$, but only three of them were stable enough for further analysis. Unfortunately, the revDSD-PBEP86/aug-cc-pVDZ calculations have not been performed, due to the Galileo Cluster shutdown. Thus for the $(H_2C_2)\cdot(HCN)_2$ trimer only the B97-3c/def2-TZVP results are available.

The most stable structure optimized at B97-3c/def2-TZVP level is a ring structure, illustrated in Fig. 4.57. Figure 4.57a displays an RCP at the centre of the structure, confirming the ring formation. The intramolecular bond lengths are quite similar to the monomer one, optimized at the same level of theory. The intermolecular distance between the π-electrons and the HCN hydrogen atom is about 2.53 $\AA$. Meanwhile, the hydrogen bond between the HCN nitrogen and the HCCH hydrogen is about 2.43 $\AA$, a little bit shorter due to the more covalent character of nitrogen electron doublet. Hydrogen bond length between the two hydrogen cyanide molecules is about 2.87 $\AA$, longer than the other intermolecular interactions. As shows by Fig. 4.57b, the interaction which the two HCN molecules establish is quite weak, since it occurs between the carbon atom and the triple bond C≡N, and not between the hydrogen atom and the nitrogen non-bonding electrons. The binding energy value calculated at B97-3c/def2-TZVP level results to be -33.0 kJ/mol (-7.9 kcal/mol). After the zero-point vibrational energy correction, the binding energy value results to be -27.1 kJ/mol (-6.5 kcal/mol).

(a) (b)

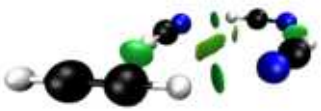


(c)

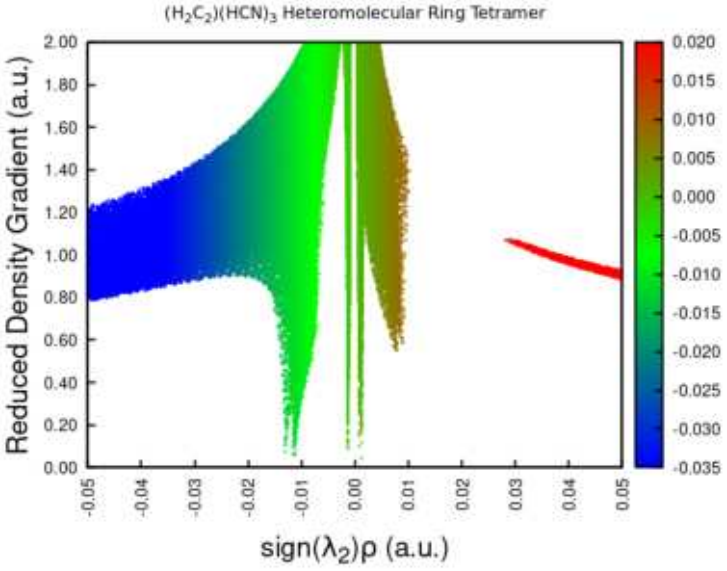
Figure 4.57: QTAIM analysis, RDG isosurfaces (RDG($\boldsymbol{\rho}$) = 0.5) and Reduced Density Gradient vs sign(λ_2)·ρ plot of the ring $(H_2C_2)\cdot(HCN)_2$ trimer.

The ring series keep going with $(H_2C_2)\cdot(HCN)_3$ tetramers complexes, as the most stable structure identified. The intramolecular bond lengths are the same of the monomers, while the π- - -HCN distance is about 2.46 $\mathring{A}$ and the others NCIs have an average value of 2.33 . As displayed in Fig. 4.58a, the QTAIM analysis detected a ring critical point which confirm the structure. The binding energy value, calculated at B97-3c level is about -62.4 kJ/mol (-14.9 kcal/mol), -53.1 kJ/mol (-12.7 kcal/mol) after the ZPVE correction.

(a)



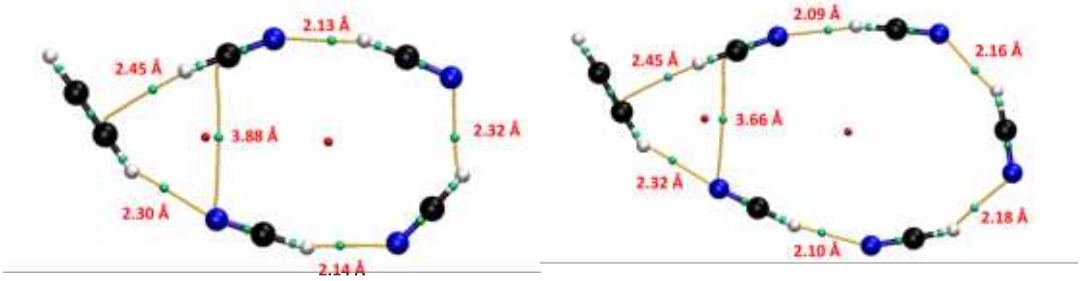
(b)



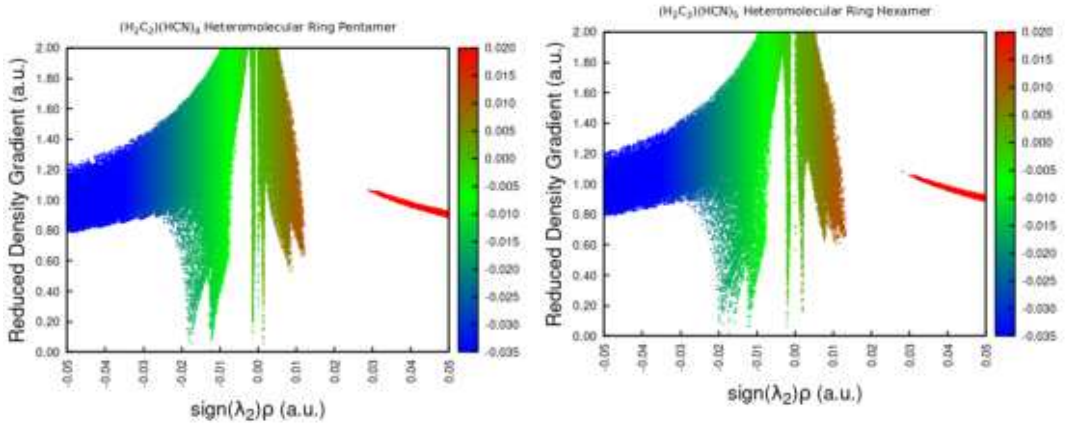
(c)

Figure 4.58: QTAIM analysis, RDG isosurfaces (RDG($\boldsymbol{\rho}$) = 0.5) and Reduced Density Gradient vs sign(λ_2)$\cdot\rho$ plot of the ring $(H_2C_2)\cdot(HCN)_3$ tetramer.

The ring structures found from the pentamer and hexamer calculations seems to be quite similar to the previous tetramer. The difference arise from the topology analysis, where a second ring was detected inside the structure. The acetylene molecule together with two HCN molecules forms a second ring, identified by an RCP in Fig. 4.59a and 4.59b. The RDG isosurface figures are not reported since they are quite similar to the ones shown in Fig. 4.58b. The bond lengths are outlined in the pictures below, except the intramolecular ones because are the same of the isolated monomers. The binding energy value is about -91.6 kJ/mol (-21.9 kcal/mol) and -78.3 kJ/mol (-18.7 kcal/mol) with the ZPVE correction for the $(H_2C_2)\cdot(HCN)_4$ complex, -120.2 kJ/mol (-28.7 kcal/mol) and -103.4 kJ/mol (-24.7 kcal/mol) with the ZPVE correction for the $(H_2C_2)\cdot(HCN)_5$ complex. The same structure was also found in the $(H_2C_2)\cdot(HCN)_7$ octamer analysis, which doesn't reach the optimization.



(a) **(b)**



(c) **(d)**

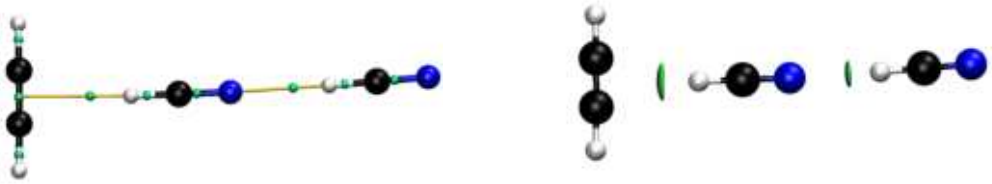
Figure 4.59: QTAIM analysis and Reduced Density Gradient vs sign(λ_2)·ρ plot of the ring complexes.

Figures 4.59c and 4.59d show several NCIs between the HCN and H_2C_2 molecules. The two peaks near $\rho = 0$ identified the two rings, while the peaks between -0.01 a.u. and -0.02 a.u. $sign(\lambda_2)\rho$ show the presence of Van der Waals interactions with high covalent character, *e.g.* hydrogen bond. The NCIs between HCN molecules are stronger when the donor nitrogen atom and the acceptor hydrogen atom are aligned, thus the peaks at -0.02 a.u. identified this type of interaction. While the peak at -0.01 a.u. shows the presence of weaker Van der Waals interactions, such as the ones between H_2C_2 and HCN molecules. Figure 4.59d shows more NCIs in the -0.02 a.u. $< sign(\lambda_2)\rho <$ -0.01 a.u. range, due to insertion of one HCN molecule in the ring, which enhance the cluster NCIs.

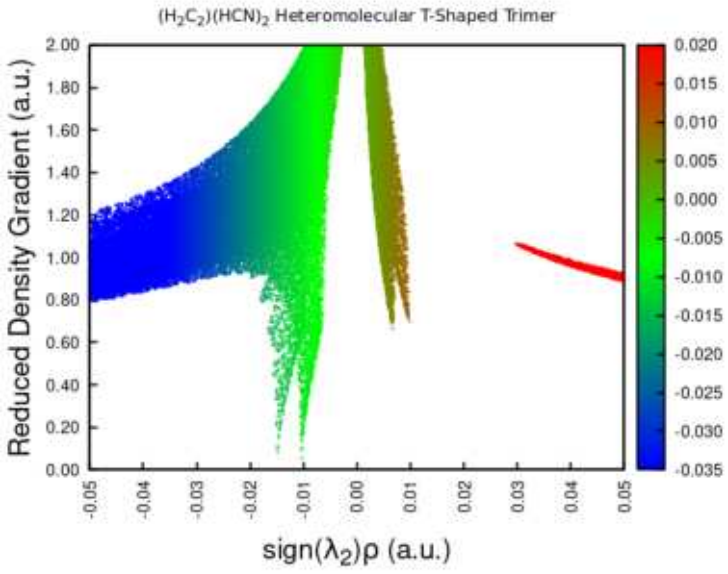
4.10.2 T-Shaped and Linear clusters

A $(H_2C_2)\cdot(HCN)_2$ trimer with an identical structure to the fourth dimer with one more HCN molecule was optimized. The intramolecular lengths are the same of the HCN and HCCH monomers. The intermolecular bond between the π-electrons has a value of 2.50 $\AA$, while the one between the two HCN molecules is 2.21 $\AA$. From Fig. 4.60b it's clear that the HCN- - -HCN interaction is stronger than the other, due to dipole-dipole forces and the

better polarization of HCN molecules thanks to π-electrons interaction. In fact, two peaks are illustrated in Fig. 4.60c, one at -0.01 a.u., associated at the π- - -HCN interaction, and the other at -0.015 a.u. of sign(λ_2)$\cdot\rho$, associated at the HCN- - -HCN interaction. The latter has a greater covalent character that the former. This complex has a binding energy value of -31.9 kJ/mol (-7.6 kcal/mol), after the ZPVE correction the binding energy value is about -26.0 kJ/mol (-6.2 kcal/mol). This structure was also found in the tetramer analysis, where three HCN molecules are aligned together instead of two. The π- - -HCN NCI length is about 2.48 *Å*, while the internal HCN- - -HCN interaction is 2.14 *Å* long and the other HCN- - -HCN bond length is about 2.18 *Å*. The binding energy of this tetramer complex is -55.8 kJ/mol (-13.3 kcal/mol) and -46.4 kJ/mol (-11.1 kcal/mol) with the ZPVE correction.



(a) **(b)**



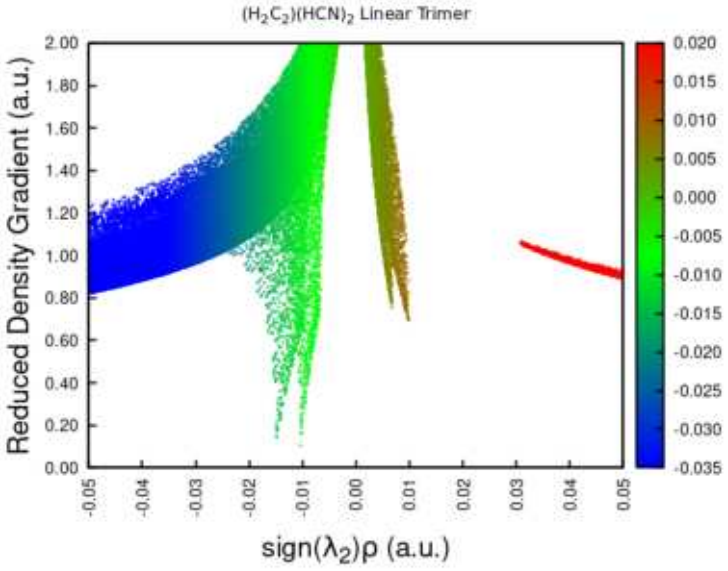
(c)

Figure 4.60: QTAIM analysis, RDG isosurfaces (RDG($\boldsymbol{\rho}$) = 0.5) and Reduced Density Gradient vs sign(λ_2)$\cdot\rho$ plot of the T-Shaped $(H_2C_2)\cdot(HCN)_2$ trimer.

The linear $(H_2C_2)\cdot(HCN)_2$ trimer optimized consist in all three molecules align together, as illustrated in Fig. 4.61. The intramolecular bonds are quite similar to the ones of monomers, except for the acetylene C-H bond, which interact with HCN molecule which is a little bit

longer. HCN- - -HCN intermolecular interaction length is about 2.21 $\AA$, while the HCN- - -HCCH distance is about 2.38 $\AA$. The BE value is -30.3 kJ/mol (-7.2 kcal/mol), while the ZPVBE is -24.3 kJ/mol (-5.8 kcal/mol). The $(H_2C_2)\cdot(HCN)_3$ linear tetramer was also found where the intermolecular lengths between the HCN molecules are about 2.16 $\AA$ and the one between the HCN and H_2C_2 molecules is equal to 2.36 $\AA$. The binding energy value is -54.2 kJ/mol (-13.0 kcal/mol) and the ZPVBE value is -44.6 kJ/mol (-10.7 kcal/mol).

(a) **(b)**



(c)

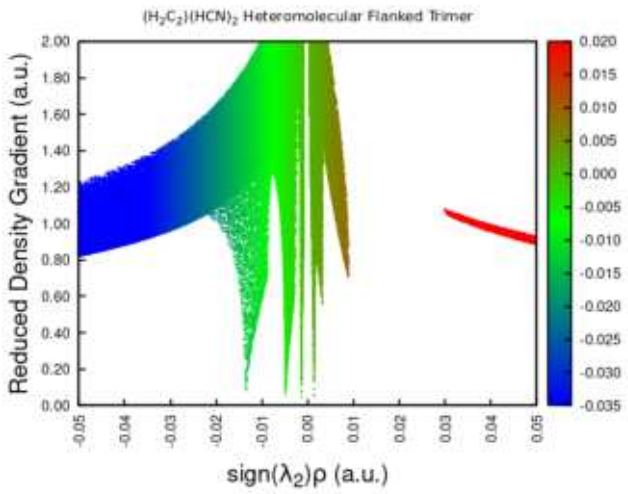
Figure 4.61: QTAIM analysis, RDG isosurfaces (RDG($\boldsymbol{\rho}$) = 0.5) and Reduced Density Gradient vs sign(λ_2)·ρ plot of the linear $(H_2C_2)\cdot(HCN)_2$ trimer.

Both T-Shaped and linear complexes show as the intermolecular interactions between HCN molecules, change from the trimer (two HCN molecules aligned) and the tetramer (three HCN molecules aligned). The dominant structure in both clusters is the linear disposition of HCN and as showed in section 4.7.1, the intermolecular lengths become shorter with the insertion of more hydrogen cyanide molecules.

4.10.3 Flanked clusters

As seen above, the linear structure of hydrogen cyanide is particularly stable. Among the mixed clusters, several structures have been identified in which the acetylene molecule is flanked to a linear cluster of HCN. Among the different clusters, this is the least stable, due to the poor interaction between the H_2C_2 molecule and the remaining HCN. The intramolecular bond lengths calculated at B97-3c/def2-TZVP level are the same as the monomers, while the HCN- - -HCN Van der Waals interaction is 2.25 *Å* long and the distance, between the acetylene and the hydrogen cyanide molecules, is over 3 *Å*. The BE value is -25.4 kJ/mol (-6.1 kcal/mol), while the ZPVBE value is -20.7 kJ/mol (-5.0 kcal/mol). Fig. 4.62a illustrates the topology analysis which confirm the presence of an NCI between acetylene and hydrogen cyanide, and also the RDG isosurface analysis, Fig. 4.62b, shows two surfaces among the HCN linear cluster and the H_2C_2 molecule, but only one is identified as a strong interaction, the other surface probably identified as a steric repulsion. The RDG plot in Fig. 4.62c shows many NCIs, the peak at most negative value refer to HCN- - -HCN interaction, while the one at -0.015 a.u. is the Van der Waals interaction between the acetylene and the hydrogen cyanide molecule.

(a) **(b)**

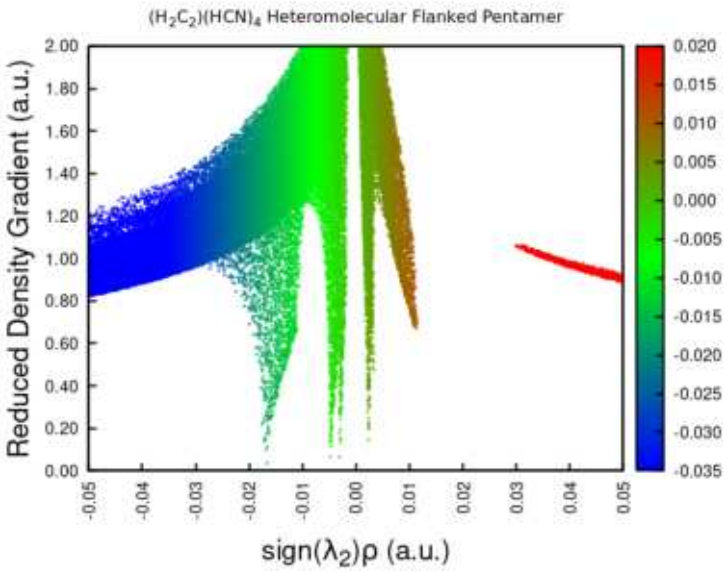


(c)

Figure 4.62: QTAIM analysis, RDG isosurfaces (RDG($\boldsymbol{\rho}$) = 0.5) and Reduced Density Gradient vs sign(λ_2)·ρ plot of the flanked $(H_2C_2)\cdot(HCN)_2$ trimer.

Among the $(H_2C_2)\cdot(HCN)_4$ clusters a similar structure was optimized, where a H_2C_2 molecules stands alongside a linear $(HCN)_4$ cluster. In this case, the acetylene molecule interact in a better way thanks the bending of $(HCN)_4$ cluster, which lost the linear arrangement. In Fig. 4.63a shows the topology analysis where a RCP is identified between two HCN molecules and the acetylene molecule. Intermolecular distance between the acetylene molecule and the $(HCN)_4$ cluster is quite long and the RDG plot illustrate two intense Van der Waals interactions, besides the peak near -0.02 a.u. which refer to HCN- - -HCN interactions. The BE of this cluster is -73.7 kJ/mol (-17.6 kcal/mol) and the ZPVBE is -61.9 kJ/mol (-14.8 kcal/mol).

(a) **(b)**

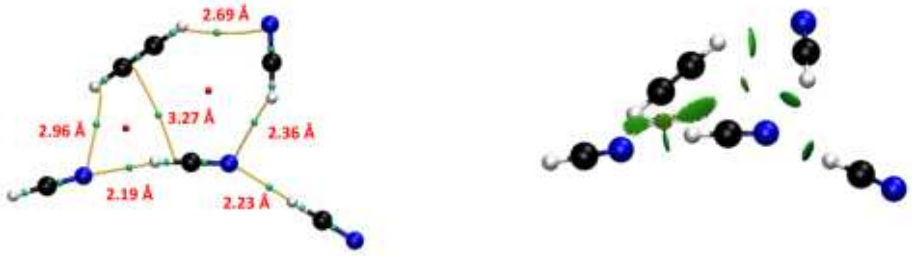


(c)

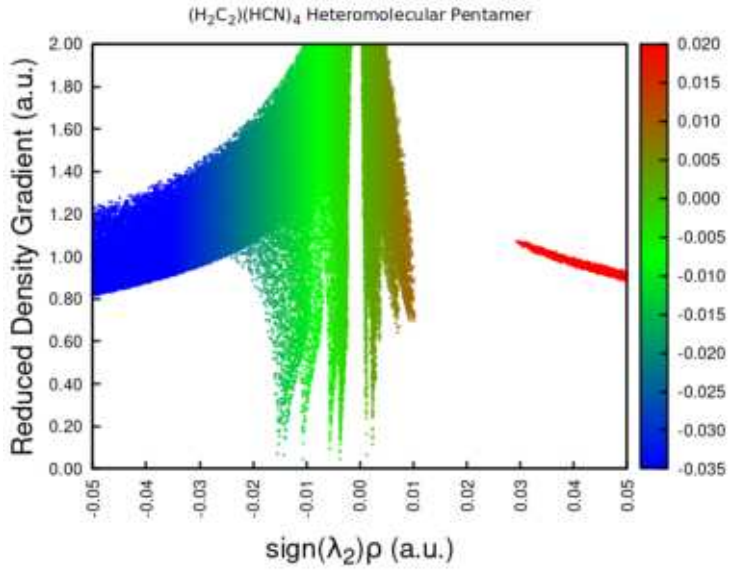
Figure 4.63: QTAIM analysis, RDG isosurfaces (RDG($\boldsymbol{\rho}$) = 0.5) and Reduced Density Gradient vs sign(λ_2)·ρ plot of the flanked $(H_2C_2)\cdot(HCN)_4$ pentamer.

An evolution of this complex could be the following pentamer, found during the xTB calculation and optimized at B97-3c level, where an HCN molecule and H_2C_2 molecule interact together and with an $(HCN)_3$ linear cluster. Many NCIs occur among the molecules those form this complex, as shown from figure 4.64c. Two peaks over the 0.00 a.u. identified the

RCPs, while the other peaks a negative value of sign(λ_2)ρ refer to the many Van der Waals interactions displayed in Fig. 4.64a. The BE value is -65.5 kJ/mol (-15.7 kcal/mol) and the ZPVBE value is -54.8 kJ/mol (-13.1 kcal/mol).



(a) **(b)**



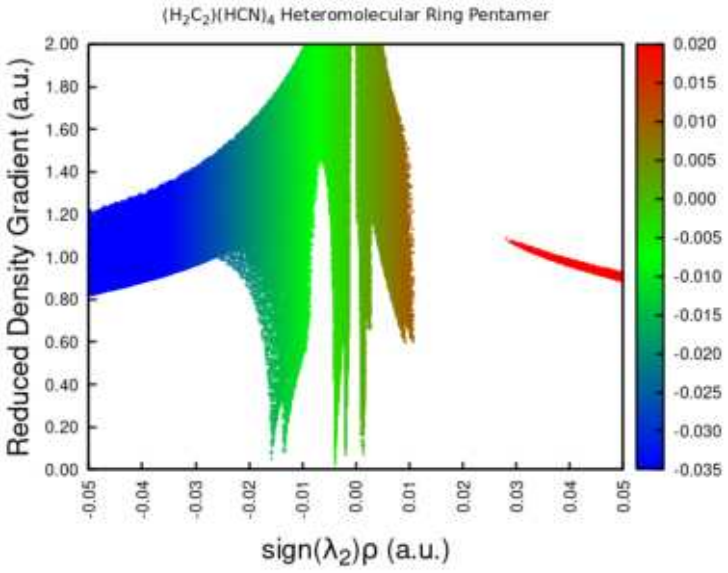
(c)

Figure 4.64: QTAIM analysis, RDG isosurfaces (RDG($\boldsymbol{\rho}$) = 0.5) and Reduced Density Gradient vs sign(λ_2)$\cdot\rho$ plot of $(H_2C_2)\cdot(HCN)_4$ pentamer.

4.10.4 $(HCN)_4$ Ring cluster + H_2C_2

A cluster formed by a ring of four HCN molecules with an external acetylene molecule interacting with them, were optimized at B97-3c level. The analysis of the pentamer result in a $(HCN)_4$ ring similar to the one saw in section 4.7.2 with a H_2C_2 molecule on top of the ring centre, as illustrated in figure 4.65a. The distance between the ring centre of mass and the one of acetylene molecule, is 2.88 *Å*. Meanwhile, the intermolecular lengths HCN- - -HCN are 2.26 *Å* long where the nitrogen atom also interact with the acetylene hydrogen, 2.19 *Å* the other two hydrogen bonds. The BE value is -86.1 kJ/mol (-20.6 kcal/mol) and the ZPVBE value is -74.4 kJ/mol (-17.8 kcal/mol).

(a) **(b)**

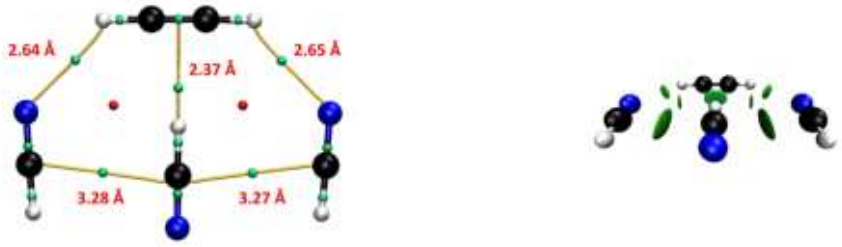


(c)

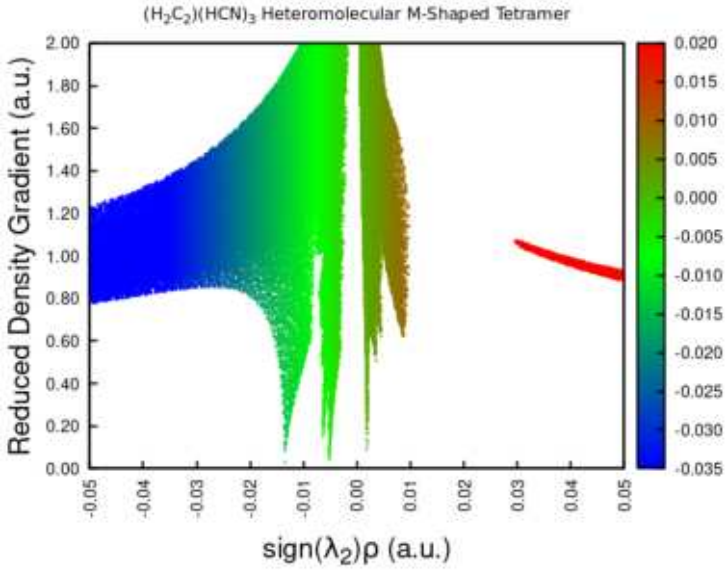
Figure 4.65: QTAIM analysis, RDG isosurfaces (RDG($\boldsymbol{\rho}$) = 0.5) and Reduced Density Gradient vs sign(λ_2)·ρ plot of $(HCN)_4$ ring + H_2C_2.

4.10.5 Other interesting clusters

During the tetramer analysis an interesting structure was found, where an acetylene molecule form two rings with three different hydrogen cyanide molecules, with a pretty strong interaction between π-electrons and the hydrogen atom. Figures 4.66 show the QTAIM and RDG analysis of the complex, the three HCN molecules are quite polarized and interact through transversal dipole-dipole forces. Fig. 4.66c illustrates the different interactions, at positive sign$(\lambda_2)\cdot\rho$ values there are peaks which refer to RCPs, the green peaks refer to the Van der Waals interactions between the HCN molecules and the peak at -0.015 a.u. is associated to the strong interaction between H_2C_2 and HCN. The BE value is -48.6 kJ/mol (-11.6 kcal/mol), while the ZPVBE value is -40.9 kJ/mol (-9.8 kcal/mol).



(a) **(b)**



(c)

Figure 4.66: QTAIM analysis, RDG isosurfaces (RDG($\boldsymbol{\rho}$) = 0.5) and Reduced Density Gradient vs sign$(\lambda_2)\cdot\rho$ plot of $(H_2C_2)\cdot(HCN)_3$ tetramer.

Another interesting structure was found during the pentamer analysis and subsequently optimized at B97-3c/def2-TZVP level. The structure illustrated in figures 4.67 was called *"Snail Shell"* due to molecules spiral arrangement. As shown by Fig. 4.67c, there are several

NCIs occurring in this complex, especially VdW attractive forces. The BE value is -72.2 kJ/mol (-18.4 kcal/mol), while the ZPVBE value is -65.8 kJ/mol (-15.7 kcal/mol).

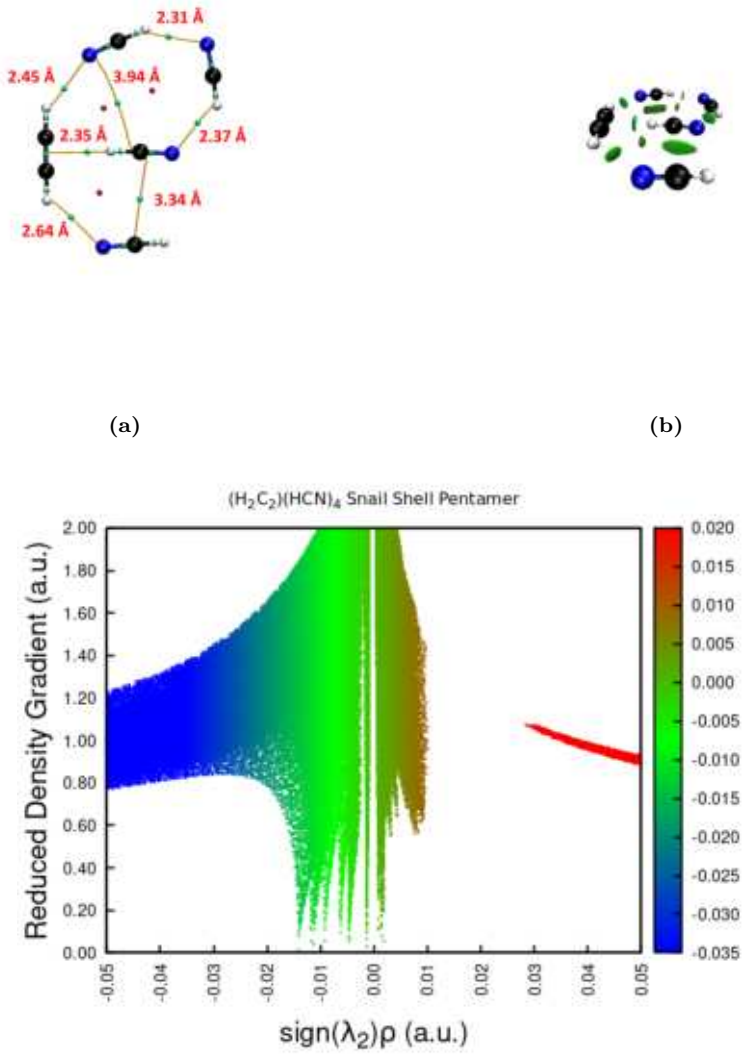
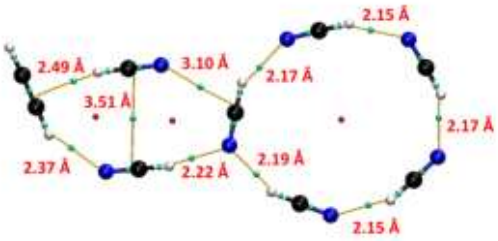
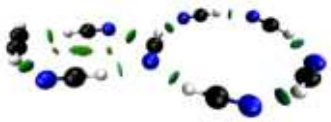


Figure 4.67: QTAIM analysis, RDG isosurfaces (RDG($\boldsymbol{\rho}$) = 0.5) and Reduced Density Gradient vs sign(λ_2)·ρ plot of $(H_2C_2)\cdot(HCN)_4$ pentamer.

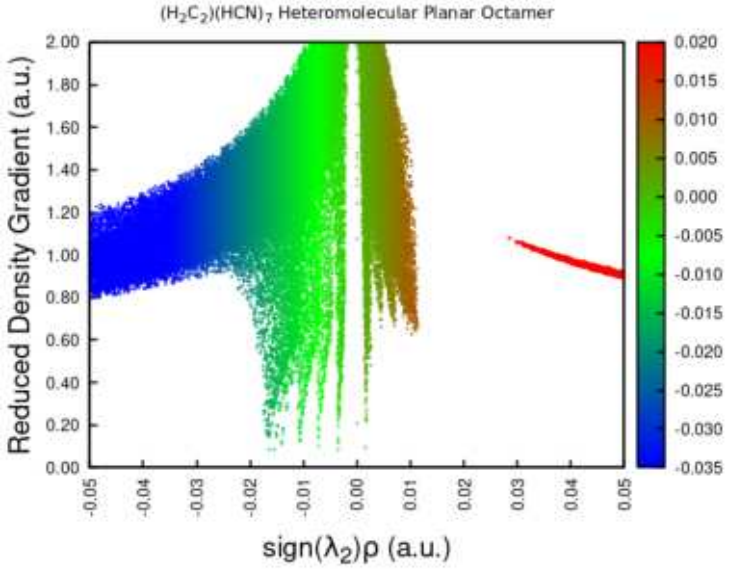
Among the octamers clusters was found a planar structure composed by a $(HCN)_5$ ring, a $(HCN)_3$ ring and a $(H_2C_2)\cdot(HCN)_2$ ring, condensed together as show in Fig. 4.68. The BE value is -158.5 kJ/mol (-37.9 kcal/mol), while the ZPVBE value is -136.3 kJ/mol (-32.6 kcal/mol).



(a)



(b)



(c)

Figure 4.68: QTAIM analysis, RDG isosurfaces (RDG($\boldsymbol{\rho}$) = 0.5) and Reduced Density Gradient vs sign(λ_2)$\cdot\rho$ plot of $(H_2C_2)\cdot(HCN)_7$ octamer.

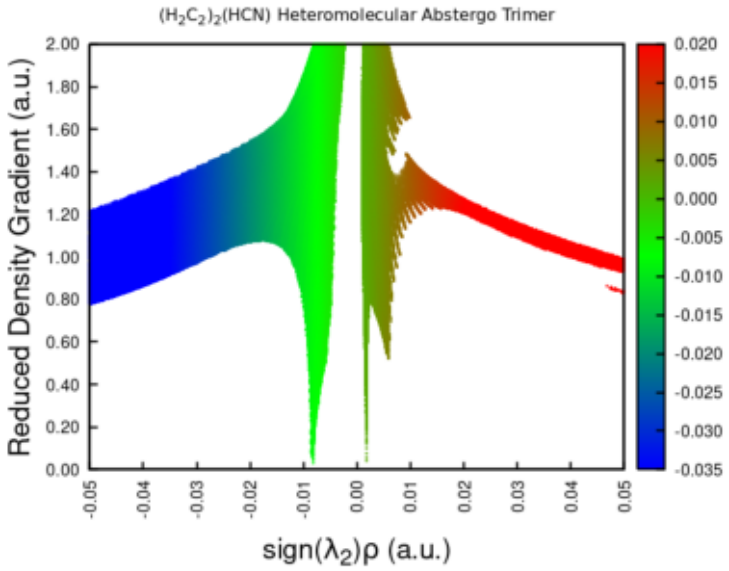
4.11 $(H_2C_2)_{n-1}\cdot(HCN)$ clusters series

The $(H_2C_2)_{n-1}\cdot(HCN)$ series, where n indicates the number of units present in the cluster, have a poor structure diversity, due to the weakness of dispersion forces. However, the presence of one hydrogen cyanide molecule leads to obtain many interesting complexes from the point of view of Van der Waals interactions.

4.11.1 Simil Abstergo cluster

As seen for the $(H_2C_2)_n$ cluster series, the abstergo structure reoccurs in many $(H_2C_2)_{n-1}\cdot(HCN)$ clusters optimized and analyzed in this thesis. The heteromolecular abstergo trimer is constituted by two acetylene molecules, which interact through H- - -π-electrons, and one hydrogen cyanide molecule. Figures 4.69 show the formation of a ring as seen in $(H_2C_2)_3$ abstergo cluster. The calculation at revDSD-PBEP86/aug-cc-pVDZ level result in binding energy value of -22.6 kJ/mol (-5.4 kcal/mol) and ZPVBE value of -17.0 kJ/mol (-4.1 kcal/mol). The geometry optimization result in intramolecular bond length equal to the ones of isolated monomers, 1.08 $\AA$ (H-C) and 1.22 $\AA$ (C-C) for acetylene monomer, 1.08 $\AA$ (H-C) and 1.17 $\AA$ (C≡N) for hydrogen cyanide monomer. The intermolecular distance of N- - -H interaction is 2.51 $\AA$, while the other two NCIs are quite similar and have a bond length value of 2.63 $\AA$. Fig. 4.69c shows only two peaks, one which refer to attractive Van der Waals interactions between the three molecules and one, at positive $\text{sign}(\lambda_2)\cdot\rho$ value, which refer to ring critical point.

(a) (b)

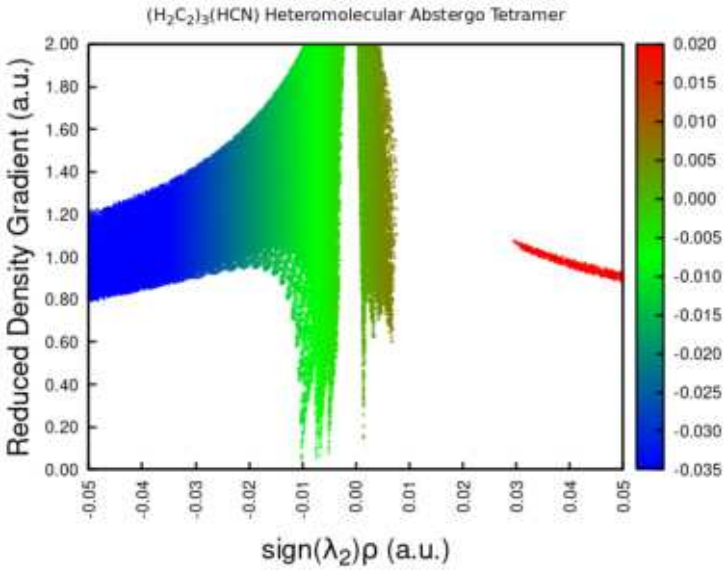


(c)

Figure 4.69: QTAIM analysis, RDG isosurfaces (RDG($\boldsymbol{\rho}$) = 0.5) and Reduced Density Gradient vs sign(λ_2)·ρ plot of heteromolecular abstergo $(H_2C_2)_2\cdot(HCN)$ trimer.

The same structure with an additional acetylene molecule was found during the tetramer analysis, after the calculation at B97-3c/def2-TZVP level the geometry optimization shows intermolecular bond lengths of 2.53 *Å* between π- - -HCN, 2.76 *Å* between HCN- - -HCCH and 2.72 *Å* the π- - -HCCH, while the non-covalent interaction with the external acetylene molecule is 2.66 *Å* long. Fig. 4.70 present the structure and the topology analysis. Fig. 4.70c shows the establishment of many NCIs, several dipole-dipole and dispersion forces and repulsive VdW due to ring formation. The BE value calculated at B97-3c/def2-TZVP level is -33.5 kJ/mol (-8.0 kcal/mol), while the ZPVBE value is -26.5 kJ/mol (-6.3 kcal/mol).

(a) (b)

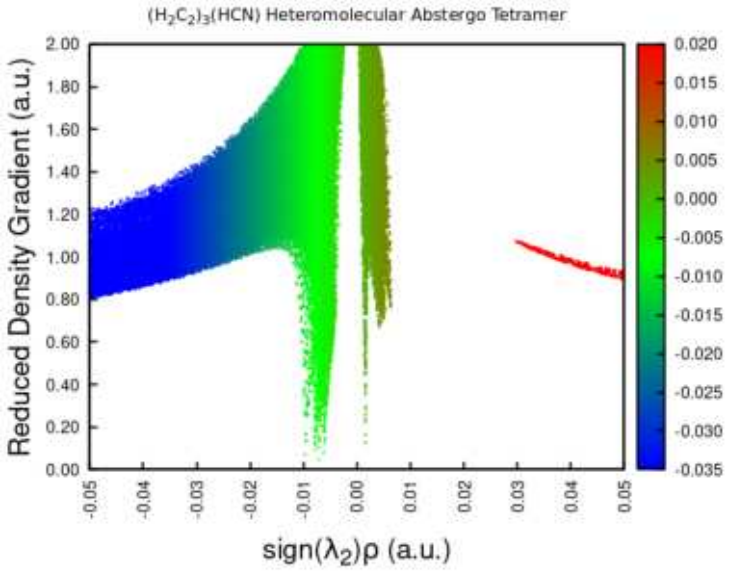


(c)

Figure 4.70: QTAIM analysis, RDG isosurfaces (RDG($\boldsymbol{\rho}$) = 0.5) and Reduced Density Gradient vs sign(λ_2)·ρ plot of heteromolecular abstergo $(H_2C_2)_3$·(HCN) tetramer.

Another similar tetramer structure consist of an $(H_2C_2)_3$ abstergo structure with an HCN molecule interacting from aside. The B97-3c/def2-TZVP calculations show an average value of NCI inside the abstergo ring of 2.70 *Å* as seen in the isolated abstergo cluster. While the intermolecular distance between HCN and H_2C_2 molecule is 2.40 *Å*. Fig. 4.71 shows the structure and the topology analysis of the tetramer, figure 4.71c illustrate the different NCIs occurring in this cluster, there are many attractive Van der Waals interactions due to dispersion and dipole-dipole forces and a repulsive one due to the ring formation. The BE value is -31.9 kJ/mol (-7.6 kcal/mol) and -24.4 kJ/mol (-5.8 kcal/mol) after the ZPVE correction.

(a) (b)



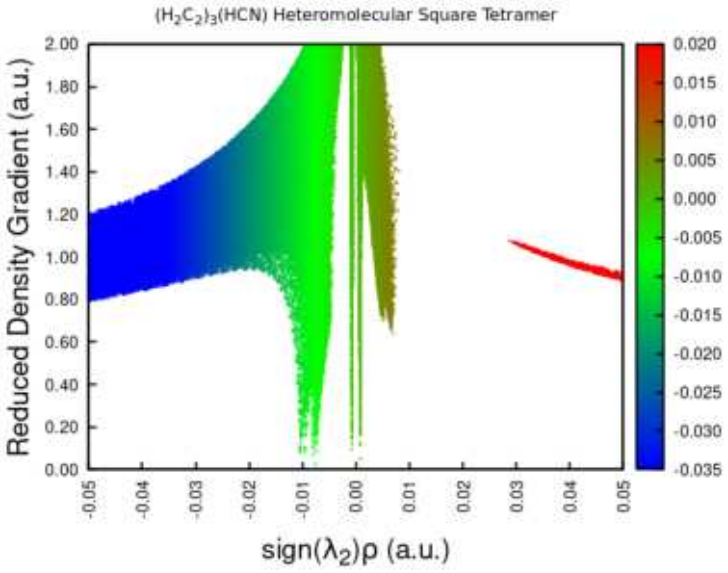
(c)

Figure 4.71: QTAIM analysis, RDG isosurfaces (RDG($\boldsymbol{\rho}$) = 0.5) and Reduced Density Gradient vs sign(λ_2)·ρ plot of $(H_2C_2)_3$·(HCN) tetramer.

4.11.2 Other $(H_2C_2)_{n-1}$·(HCN) clusters

Other clusters optimized and analyzed at B97-3c/def2-TZVP level during this work, are illustrated below with a brief commentary since there are no significant observations. Fig. 4.72 shows a non-planar structure of $(H_2C_2)_3$·(HCN) tetramer. The topology analysis doesn't detect an RCP but the RDG analysis in figure 4.55b shows a repulsive interaction at the centre. The interactions between HCN and HCCH molecules are quite strong, as observed from the bond length. The BE value calculated is -39.2 kJ/mol (-9.4 kcal/mol) and the ZPVBE value is -31.0 kJ/mol (-7.4 kcal/mol).

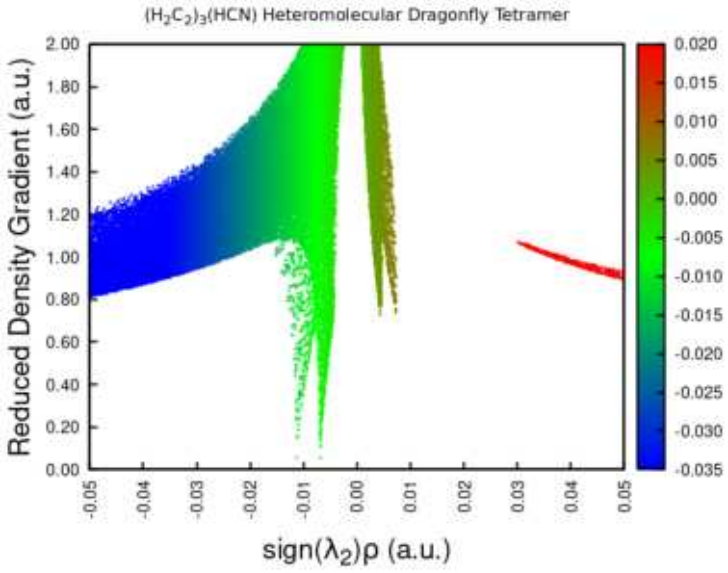
(a) **(b)**



(c)

Figure 4.72: QTAIM analysis, RDG isosurfaces (RDG($\boldsymbol{\rho}$) = 0.5) and Reduced Density Gradient vs sign(λ_2)·ρ plot of squared $(H_2C_2)_3$·(HCN) tetramer.

(a) (b)

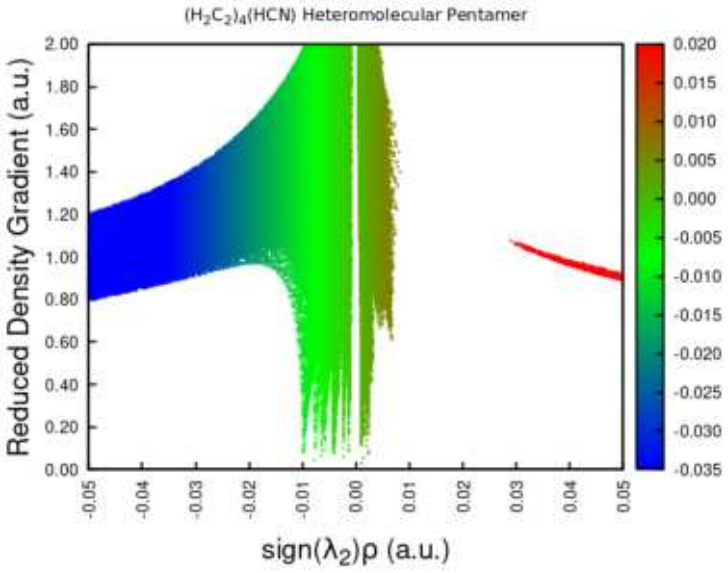


(c)

Figure 4.73: QTAIM analysis, RDG isosurfaces (RDG($\boldsymbol{\rho}$) = 0.5) and Reduced Density Gradient vs sign(λ_2)$\cdot\rho$ plot of dragonfly $(H_2C_2)_3\cdot(HCN)$ tetramer.

During the optimization of $(H_2C_2)_4\cdot(HCN)$, $(H_2C_2)_5\cdot(HCN)$ and $(H_2C_2)_6\cdot(HCN)$ clusters, at B97-3c/def2-TZVP level, only one structure each was found. Several NCIs occur between the molecules, which are arranged three-dimensionally in space without a specific pattern. These structures are illustrated and briefly described below. The pentamer cluster consists of a cage of acetylene molecules and one hydrogen cyanide molecule, where the interactions are mostly London forces. Fig. 4.74 display the structure and the different non-covalent bonds occurring in the complex, the HCN- - -HCCH interaction length is 2.41 $\mathring{A}$ and it's the strongest among the others, which have distances between 2.50 and 4.00 $\mathring{A}$. Figure 4.74c shows a repulsive Van der Waals interaction due to the cage formation. As seen from Fig. 4.74c, there are only weak Van der Waals interactions. The BE value is -54.8 kJ/mol (-13.1 kcal/mol) and the ZPVBE value is -43.8 kJ/mol (-10.5 kcal/mol).

(a) (b)

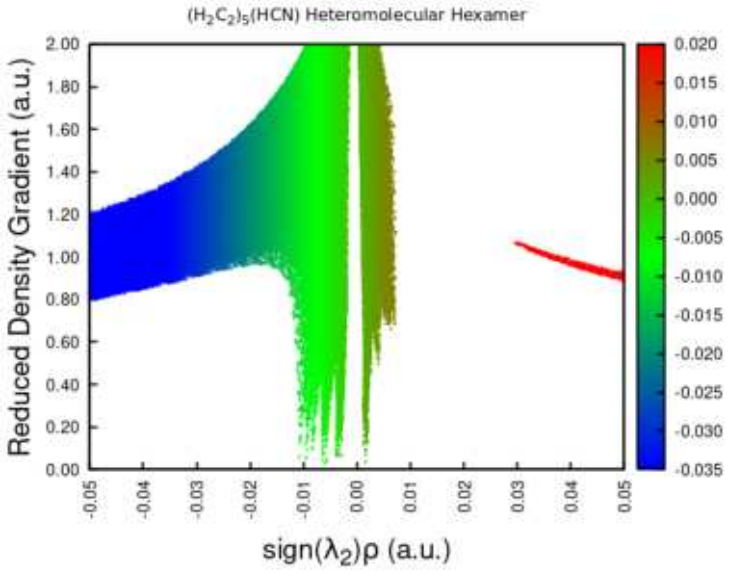


(c)

Figure 4.74: QTAIM analysis, RDG isosurfaces (RDG($\boldsymbol{\rho}$) = 0.5) and Reduced Density Gradient vs sign(λ_2)·ρ plot of $(H_2C_2)_4$·(HCN) pentamer.

Regarding the hexamer cluster, the molecules of acetylene and hydrogen cyanide assemble two cages and many rings as shown from figure 4.75. Fig. 4.75c illustrates many peaks due to several VdW interactions, which are mainly attractive. The HCN- - -HCCH interaction length is 2.36 $\AA$ and it's the strongest among the others, which have distances between 2.50 and 4.00 $\AA$. The BE value for this structure is -70.1 kJ/mol (-16.7 kcal/mol), while the ZPVBE value is -56.6 kJ/mol (-13.5 kcal/mol).

(a) (b)

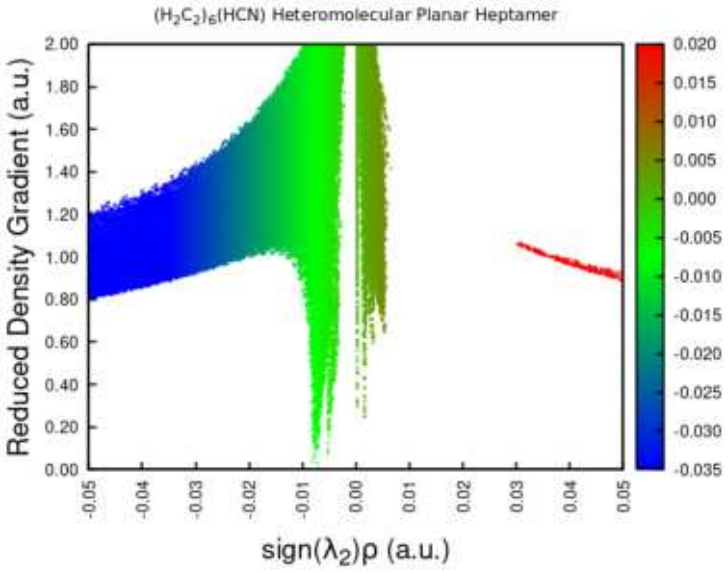


(c)

Figure 4.75: QTAIM analysis, RDG isosurfaces (RDG($\boldsymbol{\rho}$) = 0.5) and Reduced Density Gradient vs sign(λ_2)·ρ plot of $(H_2C_2)_5$·(HCN) hexamer.

The heptamer cluster found during the optimization process is a highly ordered structure, where seven H_2C_2 molecules form two planar square cycles. The HCN molecule stands out of the acetylene cluster and assembles a three-member ring, identified by topology analysis in Fig. 4.76a. The topology analysis doesn't recognise the two squares, but two small RDG isosurface were detected by the software in the RDG analysis. Fig. 4.76c shows fewer NCIs than the previous clusters, the green peaks refer to attractive Van der Waals interactions between the molecules that compose the complex, and the red peaks refer to the RCP. The non-covalent bond lengths between acetylene molecules have an average value of 2.65 $\AA$, while the HCN- - -HCCH length is 2.50 $\AA$ long and π- - -HCN length is 2.98 $\AA$ long. The BE value for the planar heptamer is -80.5 kJ/mol (-19.2 kcal/mol), while the ZPVBE value is -64.9 kJ/mol (-15.5 kcal/mol).

(a) **(b)**



(c)

Figure 4.76: QTAIM analysis, RDG isosurfaces (RDG($\boldsymbol{\rho}$) = 0.5) and Reduced Density Gradient vs sign(λ_2)·ρ plot of $(H_2C_2)_6$·(HCN) heptamer.

Unfortunately, none of the octamer structures of the mixed $(H_2C_2)_{n-1}$·(HCN) series reached a minimum stationary point during the optimization.

Chapter 5

Conclusions

The first aim of this work was the building up of a benchmark for the hydrogen cyanide and acetylene dimers. Most common DFT functionals and modified versions of them were tested, including the tight-binding semi-empirical method xTB. Also, the "gold-standard" fc-CCSD(T)/aug-cc-pVTZ method was implemented as the highest theory method. Our focus was on the analysis of non-covalent interactions (NCIs) length and strength and their correlation with the dimers binding energy, with particular attention to "low-cost" methods. During the benchmark performance, the atoms and molecules were able to move freely in space. The accuracy of the tested methods was determined by comparison with the dimers geometrical parameters obtained by extrapolation scheme developed by Barone *et al.* [56], which it's not affected by Basis Set Superposition Error (BSSE). The analysis of non-covalent bonds Absolute Percentage Error (APE%) respect the extrapolated value, allow us to determine the best computational approach for the $(HCN)_m$, $(H_2C_2)_n$, $(H_2C_2)\cdot(HCN)_{m-1}$ and $(H_2C_2)_{n-1}\cdot(HCN)$ systems. Below are reported the benchmark and the extrapolation scheme results for the dimers, and their topology analysis for a better description of Van der Waals complexes.

- **Dimers benchmark and extrapolation scheme results:**

  The linear $(HCN)_2$ dimer structure obtained from the extrapolation procedure shows intramolecular bonds length with a slight deviation concerning the monomer values. The intermolecular bond length HCN- - -HCN present a value of 2.17 *Å*. From the benchmark result that many functionals return good APE% values for NCIs, in particular among the "low-cost" methods, xTB and B97-3c result pretty accurate. The QTAIM analysis detected a strong dipole-dipole interaction between the two HCN molecules.

  The T-Shaped $(H_2C_2)_2$ dimer extrapolated structure shows intramolecular bonds length equal to the monomer ones, while the intermolecular bond length result 2.62 *Å* long. In this case, the B97-3c functional results as the most accurate for the geometry calculation, but also more expensive methods return good values of NCI length. The QTAIM analysis detected a strong VdW interaction due to dispersion forces, occurring between the π-electrons and the temporally dipole of the acetylene molecule. This NCI turns out weaker than the HCN- - -HCN one.

  The linear $(HCN)\cdot(H_2C_2)$ dimer extrapolated structure shows intramolecular values slightly different concerning the monomers, while the intermolecular bond length value is 2.31 *Å*. For this system, the benchmark results have shown good APE% values, especially the B2PLYP and the B3LYP functionals. Concerning the "low-cost" methods, only the xTB model give a good result, but the B97-3c one is also acceptable. The QTAIM analysis detected a strong heteromolecular interaction occurring between the nitrogen unshared electrons and the temporally polarized acetylene molecule. This interaction is weak than the one of $(HCN)_2$ dimer but stronger than the one in the $(H_2C_2)_2$ dimer.

  The T-Shaped $(H_2C_2)\cdot(HCN)$ dimer structure obtained from the extrapolation scheme shows intramolecular bond lengths slightly different from the monomers, while the intermolecular bond length value results to be 2.50 *Å*. In this case, only the PBEh-3c

and HF-3c have high APE% values, the other functional tested turns out pretty accurate. The QTAIM analysis shows an NCI occurring between the π-electrons and the positive pole of the HCN molecule, its strength is comparable to the linear mixed dimer.

The second aim of this thesis was the optimization of larger clusters of hydrogen cyanide and acetylene, which require more computational resources. First of all, the trimers clusters were analyzed by the revDSD-PBEP86/aug-cc-pVDZ method, but not all the calculations were possible, due to the Galileo Cluster shutdown. The high-level revDSD-PBEP86 method seems to go well for acetylene and hydrogen cyanide systems with more than one NCIs, but its computational demand makes it inconvenient for complexes larger than trimers. After the dimers benchmark, the composite functional B97-3c with the def2-TZVP basis set turns out as an excellent method for the best evaluation of NCIs in larger clusters, also concerning its velocity. The B97-3c/def2-TZVP was the choice for the optimization of clusters of up to three molecules. Optimization of clusters to heptamers did not present problems, while some structures between octamers did not achieve convergence. Most of the structures presented and discussed in this work have never been reported in the literature, especially for complexes with more than four units. Below are illustrated the results for the more stable structures of $(H_2C_2)_n \cdot (HCN)_m$ clusters optimized at B97-3c/def2-TZVP level and their topology analysis.

- **Optimization and analysis of larger clusters:**

  The B97-3c/def2-TZVP calculations of $(HCN)_m$ clusters, returns several structure types such as the linear series, the ring series, and others, due to the dipole-dipole interactions of the HCN molecule. Some of the structures optimized in this work were already optimized by Stein T. *et al.* [18], only a specific tetramer was not found in our analysis, which was reproduced and optimized at B97-3c/def2-TZVP level. The QTAIM analysis confirmed the strong interactions occurring between the HCN molecules, which help to stabilize the structures of the clusters.

  During the $(H_2C_2)_m$ clusters series optimization at B97-3c/def2-TZVP level was found a stable trimer which structure reappears in many clusters. The structures in this series are predominated by the abstergo arrangement of acetylene molecules, from the trimer to the octamer. The weak dispersion forces allow the formation only of a few structures, and the QTAIM analysis confirmed the weakness of those interactions.

  The analysis of $(H_2C_2)_2 \cdot (HCN)_2$ clusters series returns some stable structures, optimized at B97-3c/def2-TZVP level. Two square structures were found, one with an alternate arrangement of acetylene and hydrogen cyanide molecules and one with acetylene and hydrogen cyanide molecules paired together. The last structure optimized it's stabilized by the formation of dipole-induced dipole interactions between HCN and H_2C_2 molecules. The QTAIM analysis shows the formation of a ring structure for the first two clusters, while for the third there is no sign of RCP or CCP, the interactions are only between the hydrogen cyanide and acetylene molecules.

  The optimization of $(H_2C_2) \cdot (HCN)_{m-1}$ clusters series shows many hydrogen cyanide structures, where one acetylene molecule was trying to insert in the HCN complex. Several HCN ring or linear clusters were deformed by an acetylene molecule on the outside. Other structures incorporate the acetylene molecule and form a stable heteromolecular cluster, *e.g.* the one in Fig. 4.66 or 4.67. The QTAIM analysis shows many strong NCIs between the HCN molecules and the H_2C_2 molecule.

  The $(H_2C_2)_{n-1} \cdot (HCN)$ cluster series shows up again the abstergo structure, among the trimers between two acetylene molecules and one of hydrogen cyanide. During the tetramers analysis, also the pure acetylene abstergo structure was found. As shown by the RDG plot versus the $\text{sign}(\lambda_2) \cdot \rho$, one HCN molecule considerably increase the number of VdW interactions, due to the polarization of the acetylene molecules around.

5.0.1 Possible continuation

This work will provide the basis for further studies related to these clusters. In particular, it will be possible to:

- Perform more accurate calculation methods than previously tested, even for clusters above dimers, *i.e.* revDSD-PBEP86 and fc-CCSD(T);
- Perform an extrapolation scheme for the most stable trimers clusters analyzed in this work;
- Perform the analysis through the symmetry-adapted perturbation theory (SAPT) binding energy decomposition for the dimers, illustrated by Tian Lu *et al.* [59], for a better understanding of BE nature;
- Study the reactivity and possible mechanisms of reactions between acetylene or hydrogen cyanide radicals, which can lead to the formation of precursor molecules of biological systems, e.g. nucleic acids, through Internal Redundant Coordinates (IRC) calculation.

Acknowledgements

My heartfelt thanks go to those who have always supported me in my choices.
First of all, I would like to thank the research group *"Team Gamma"* which hosted me during these months of internship. I'm glad that in addition to the hard work there were also a lot of laughs, for this, I sincerely thank my supervisors Andrea Charmet Pietropolli and Paolo Stoppa. Here I recite the phrase quoted in the movie "L'armata Brancaleone", that Prof. Pietropolli used to recite, which has accompanied me throughout my thesis internship:

"Oh, gioveni! Quando vi dico 'sequitemi miei pugnaci', dovete sequire et pugnare! Poche fotte! Se no qui stemo a prenderci per le natiche." – Brancaleone da Norcia, L'armata Brancaleone.

A thank you also to the DSMN, for better or worse, which contributed to my growth.
A huge thank you goes to my family for all the support, both emotional and economic, but especially the last one. Thanks to my brother Samir, who has read this thesis and fixed my bad English.

Many friends supported me during the crowning of my studies, special thanks to those who have been closest to me as Sara *"Oto"* Liuzzi and Andrea *"Athos"* Valenti, which have always been able to advise and guide me. Many thanks to my favourite little 🦀 SaraH Delfina Piccolo, for being a good friend and helping me to fight against the devilish wasps.
A big thank you goes to my dear university friend Alberto Palù for being a good friend, and his family for their kindness and generosity.
Thanks to my classmate Silvia Pianesi, who has gone through this journey with me and shared distress and successes.
Finally, I want to thanks my second family, my brothers and sisters, my team ***T&S*** Conegliano, with whom I have shared every second of my life in the last eight years.

Sola vox, sola lux, sola urbs, Colimglane!

Bibliography

(1) *"The Cologne Database for Molecular Spectroscopy, CDMS" Last Update to Data Content: February 2021*, https://cdms.astro.uni-koeln.de/.

(2) Y. Oba, Y. Takano, H. Naraoka, Y. Furukawa, D. P. Glavin, J. P. Dworkin and S. Tachibana, *Nat. Commun.*, **11**, 6243.

(3) M. Biczysko, J. Bloino and C. Puzzarini, *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 2018, **8**, e1349.

(4) L. H. Aller and S. J. Czyzak, *ApJS*, **51**, 211.

(5) M. Peña, G. Stasińska and M. G. Richer, *A&A*, **476**, 745–758.

(6) M. Hack, *"L'universo alle soglie del 2000"*, Rizzoli, 1992.

(7) J. L. Cranford, *"From Dying Stars to the Birth of Life: The new science of astrobiology and the search for life in the universe"*, Nottingham University Press, 2011.

(8) C. W. Allen, *"Astrophysical quantities"*, Athlone Press, London, 1973.

(9) S. Yamamoto, *"Introduction to Astrochemistry"*, Springer Japan, 2017.

(10) I. W. M. Smith, C. S. Cockell and S. Leach, *"Astrochemistry and Astrobiology"*, Springer Science & Business Media, 2012.

(11) B. T. Draine, *"Physics of the Interstellar and Intergalactic Medium"*, Princeton University Press, 2010.

(12) L. E. Snyder and D. Buhl, *ApJ*, 1971, **163**, L47.

(13) J. Cernicharo, M. Agúndez, C. Kahane, M. Guélin, J. Goicoechea, N. Marcelino, E. De Beck and L. Decin, *A&A*, 2011, **529**, L3.

(14) S. T. Ridgway, D. N. Hall, S. G. Kleinmann, D. A. Weinberger and R. S. Wojslaw, *Nature*, 1976, **264**, 345–346.

(15) J. H. Lacy, N. J. Evans, II, J. M. Achtermann, D. E. Bruce, J. F. Arens and J. S. Carr, *ApJ*, **342**, L43.

(16) L. Bizzocchi, D. Prudenzano, V. M. Rivilla, A. Pietropolli-Charmet, B. M. Giuliano, P. Caselli, J. Martin-Pintado, I. Jimenez-Serra, S. Martin, M. A. Requena-Torres et al., *A&A*, 2020, **640**, A98.

(17) B. A. McGuire, R. A. Loomis, A. M. Burkhardt, K. L. K. Lee, C. N. Shingledecker, S. B. Charnley, I. R. Cooke, M. A. Cordiner, E. Herbst, S. Kalenskii et al., *Science*, 2021, **371**, 1265–1269.

(18) T. Stein, P. P. Bera, T. J. Lee and M. Head-Gordon, *Phys. Chem. Chem. Phys.*, 2020, **22**, 20337–20348.

(19) J. Cui, R. Yelle, V. Vuitton, J. Waite Jr, W. Kasprzak, D. Gell, H. Niemann, I. Müller-Wodarg, N. Borggren, G. Fletcher et al., *Icarus*, 2009, **200**, 581–615.

(20) E. Papajak, J. Zheng, X. Xu, H. R. Leverentz and D. G. Truhlar, *J. Chem. Theory Comput.*, **7**, 3027–3034.

(21) E. Papajak and D. G. Truhlar, *J. Chem. Theory Comput.*, **6**, 597–601.

(22) T. H. Dunning Jr, *J. Chem. Phys.*, 1989, **90**, 1007–1023.

(23) P. J. Stephens, F. J. Devlin, C. F. Chabalowski and M. J. Frisch, *J. Phys. Chem.*, **98**, 11623–11627.

(24) S. Grimme, *J. Phys. Chem.*, 2006, **124**, 034108.

(25) Y. Zhao and D. G. Truhlar, *J. Phys. Chem. A*, 2006, **110**, 13126–13130.

(26) S. Grimme, C. Bannwarth and P. Shushkov, *Journal of chemical theory and computation*, 2017, **13**, 1989–2009.

(27) C. Bannwarth, S. Ehlert and S. Grimme, *J. Chem. Theory Comput.*, **15**, 1652–1671.

(28) J. G. Brandenburg, C. Bannwarth, A. Hansen and S. Grimme, *J. Chem. Phys.*, 2018, **148**, 064104.

(29) S. Grimme, J. G. Brandenburg, C. Bannwarth and A. Hansen, *J. Chem. Phys.*, 2015, **143**, 054107.

(30) R. Sure and S. Grimme, *J. Comput. Chem.*, 2013, **34**, 1672–1685.

(31) S. Grimme, J. Antony, S. Ehrlich and H. Krieg, *J. Chem. Phys.*, 2010, **132**, 154104.

(32) H. Kruse and S. Grimme, *J. Chem. Phys.*, 2012, **136**, 04B613.

(33) R. F. Bader, *Acc. Chem. Res.*, 1985, **18**, 9–15.

(34) M. Bortoluzzi, *"Approccio qualitativo alla chimica computazionale"*, ARACNE editrice S.r.l., 2009.

(35) F. Jensen, *"Introduction to Computational Chemistry"*, John Wiley & Sons, 2017.

(36) C. J. Cramer, *"Essentials of Computational Chemistry: Theories and Models"*, John Wiley & Sons, 2013.

(37) C. F. Matta and R. J. Boyd, *"The Quantum Theory of Atoms in Molecules: From Solid State to DNA and Drug Design"*, John Wiley & Sons, 2007.

(38) K. W. Jucks and R. E. Miller, *Chem. Phys. Lett.*, **147**, 137–141.

(39) K. Bläsing, J. Bresien, R. Labbow, A. Schulz and A. Villinger, *Angew. Chem. Int. Ed.*, **57**, 9170–9175.

(40) A. Karpfen, *J. Phys. Chem. A*, 1999, **103**, 11431–11441.

(41) M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman and D. J. Fox, *Gaussian˜16 Revision C.01*, 2016.

(42) F. Neese, *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 2012, **2**, 73–78.

(43) T. Lu and F. Chen, *J. Comput. Chem.*, **33**, 580–592.

(44) R. Boto, F. Peccati, nLaplazaRub, A. Carbone, J.-P. Piquemal, Y. Maday et al., 2020, DOI: 10.26434/chemrxiv.9831536.v2.

(45) J. Contreras-García, E. R. Johnson, S. Keinan, R. Chaudret, J.-P. Piquemal, D. N. Beratan and W. Yang, *J. Chem. Theory Comput.*, **7**, 625–632.

(46) *"NCIPLOT: Program for revealing non-covalent interactions"*, https://github.com/juliacontrerasgarcia/nciplot.

(47) W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graph.*, **14**, 33–38.

(48) M. Piccardo, E. Penocchio, C. Puzzarini, M. Biczysko and V. Barone, *J. Phys. Chem. A*, **119**, 2058–2082.

(49) S. Grimme, J. G. Brandenburg, C. Bannwarth and A. Hansen, *J. Chem. Phys*, **143**, 054107.

(50) C. Bannwarth, E. Caldeweyher, S. Ehlert, A. Hansen, P. Pracht, J. Seibert, S. Spicher and S. Grimme, *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, **11**, e1493.

(51) S. Kozuch and J. M. L. Martin, *J. Comput. Chem.*, **34**, 2327–2344.

(52) S. Kozuch and J. M. L. Martin, *Phys. Chem. Chem. Phys.*, **13**, 20104–20107.

(53) G. Santra, N. Sylvetsky and J. M. Martin, *J. Phys. Chem. A*, 2019, **123**, 5129–5143.

(54) J. P. Perdew and K. Schmidt, *AIP Conf. Proc.*, **577**, 1–20.

(55) J. M. L. Martin and G. Santra, *Isr. J. Chem.*, **60**, 787–804.

(56) S. Alessandrini, V. Barone and C. Puzzarini, *J. Chem. Theory Comput.*, 2019, **16**, 988–1006.

(57) A. Halkier, T. Helgaker, P. Jørgensen, W. Klopper and J. Olsen, *Chem. Phys. Lett.*, **302**, 437–446.

(58) T. Helgaker, W. Klopper, H. Koch and J. Noga, *J. Chem. Phys*, **106**, 9639–9646.

(59) S. Emamian, T. Lu, H. Kruse and H. Emamian, *J. Comput. Chem.*, **40**, 2868–2881.

(60) J. Chen, Y. Zheng, A. Melli, L. Spada, T. Lu, G. Feng, Q. Gou, V. Barone and C. Puzzarini, *Phys. Chem. Chem. Phys.*, **22**, 5024–5032.

(61) A. Malloum, J. J. Fifen and J. Conradie, *Int. J. Quantum Chem.*, 2020, **120**, e26221.

(62) *"Computational Chemistry Comparison and Benchmark DataBase" Last Update to Data Content: August 2020*, https://cccbdb.nist.gov/diplistx.asp#NSRDS-NBS10.